ICAROB 2024

PROCEEDINGS OF THE 2024 INTERNATIONAL CONFERENCE ON ARTIFICIAL LIFE AND ROBOTICS

February 22 to 25, 2024
J:COM HorutoHall, Oita, Japan
29th AROB International Meeting Series

Editor-in-Chief
Masanori Sugisaka
Editors: Yingmin Jia, Takao Ito, Ju-Jang Lee
ISBN 978-4-9913337-0-5

Proceedings of The 2024 International Conference on

ARTIFICIAL LIFE AND ROBOTICS

(ICAROB2024)

February 22 to 25, J:COM HorutoHall, Oita, Japan, 2024
29th AROB International Meeting Series

Editor-in-Chief
Masanori Sugisaka
Editors: Yingmin Jia, Takao Ito, Ju-Jang Lee
ISBN 978-4-9913337-0-5

Contents

SPONSERED

ALife Robotics Corporation Ltd.

ORGANIZED BY

International Steering Committee of International Conference on Artificial Life and Robotics (ICAROB)

SUPPORTED BY

IEEE Fukuoka Section (Japan)

CO-ORGANIZED BY

Chinese Association for Artificial Intelligence (CAAI, P. R. China)

ADVISORY COMMITTEE CHAIRS

Kazuo Ishii (Kyushu Institute of Technology, Japan)
Kensuke Harada (Osaka University, Japan)
Takashi Kohno (University of Tokyo, Japan)
Eiji Hayashi (Kyushu Institute of Technology, Japan)
Jeffrey Johnson (The Open University, UK)
Katia Passerini (Seton Hall University, USA)

ADVISORY COMMITTEE

Bruce Eisenstein (Drexel University, Former IEEE President, USA)
Hidenori Kimura (RIKEN & Wasada University, Japan)
Joshua M. Epstein (The Johns Hopkins University, USA)
Ken-ichi Tanaka (Executive Fellow, Mitsubishi Electric Corporation, Japan)
Masato Nakagawa (Executive Fellow, DENSO CORPORATION, Visiting Professor of Hiroshima University, Japan)
Moshe Kam (New Jersey Institute of Technology, Former IEEE President, USA)
Paul Kalata (Drexel University, USA)
Paul Oh (Drexel University, USA)
Peter Herczfeld (Drexel University, USA)

GENERAL CHAIR

Masanori Sugisaka
(Alife Robotics Corporation Ltd., Japan)

CO-GENERAL CHAIRS

Yingmin Jia (Beihang Univesty, P. R. China)
Takao Ito (Hiroshima University, Japan)
Ju-Jang Lee (Honorary professor, KAIST, Korea)

VICE GENERAL CHAIR

Ang, Chun Kit (University of UCSI, Malaysia)
Henrik. H. Lund (Technical University of Denmark, Denmark)
John. L. Casti (International Institute for Applied Systems Analysis, Austria)
Luigi Pagliarini (Technical University of Denmark, Denmark)
(Academy of Fine Arts of Macerata, Italy)
Mohd Rizon bin Juhari (University of UCSI, Malaysia)
Norrima Mokhtar (University of Malaya, Malaysia)

PROGRAM CHAIRMAN

Makoto Sakamoto (University of Miyazaki, Japan)

CO-PROGRAM CHAIR

Marion Oswald (Technische Universität Wien, Austria)

INTERNATIONAL ORGANIZING COMMITTEE

Akira Nakamura (Saitama Institute of Technology, Japan)
Akinori Abe (Chiba University, Japan)
Caoli Wang (University of Shanghai for Science and Technology, P. R. China)
Chan Gook Park (Seoul National University, Korea)
Evgeni Magid (Kazan Federal University, Russia)
Fuzhong Wang (Henan Polytechnic University, P. R. China)
Hazry Desa (University of Malaysia, Perlis, Malaysia)
Hidehiko Yamamoto (Gifu University, Japan)
Hiroki Tamura (The University of Miyazaki, Japan)
Hiroshi Kage (Mitsubishi Electric Corporation, Japan)
Hiroshi Matsuno (Yamaguchi University, Japan)
Jiwu Wang (Beijing Jiaotong University, P. R. China)
Jovana Jovic (CNRS-AIST JRL, Japan, France)
Junping Du (Beijing University of Posts and Telecommunications, P. R. China)
Katsunori Shimohara (Doshisha University, Japan)
Kazuo Ishii (Kyushu Institute of Technology, Japan)
Kenichi Tanaka (Nagasaki Institute of Applied Science, Japan)
Kenji Hashimoto (Waseda University, Japan)
Kunikazu Kobayashi (Aichi Prefectural University, Japan)
Kuo-Hsien Hsia (Far East University, Taiwan)
Kuo-Lan Su (National Yunlin University of Science and Technology, Taiwan)
Kyungho Park (U.S. ARMY, Japan and USA)
Masao Kubo (National Defense Academy of Japan, Japan)
Maxim Sokolov (Innopolis University, Russia)
Mehta Rajiv (New Jersey Institute of Technology, USA)
Minoru Kumano (University of Miyazaki, Japan)
Peter Sapaty (Ukrainian Academy of Science, Ukraine)
Pierre Parrend (University of Strasbourg, France)
Qiang Cai (Beijing Technology and Business University, P. R. China)
Qu Yanbin (Harbin Institute of Technology, P. R. China)
Singo Mabu (Yamaguchi University, Japan)
Takashi Kohno (The University of Tokyo, Japan)
Takashi Ogata (Iwate Prefectural University)
Tetsuro Hattori (Kagawa University, Japan)

Thi Thi Zin (University of Miyazaki, Japan)
Thomas S. Ray (University of Oklahoma, USA)
Toru Yamamoto (Hiroshima University, Japan)
Victor Berdonosov (Komsomolsk-on-Amur State University of Technology, Russia)
Yasunari Yoshitomi (Kyoto Prefectural University, Japan)
Yi Chai (Chongqing University, P. R. China)
Yoshifumi Morita (Nagoya Institute of Technology, Japan)
Yoshiro Imai (Kagawa University, Japan)

INTERNATIONAL PROGRAM COMMITTEE

Abdul Rahman bin Dullah (Universiti Teknikal Malaysia Melaka, Malaysia)
Akinori Abe (Chiba University, Japan)
Akihiro Yamaguchi (Fukuoka Institute of Technology, Japan)
Akihito Kanai (Hosei University, Japan)
Ali Selamat (University of Technology of Malaysia, Malaysia)
Aminurafiuddin bin Zulkifli (Multimedia University, Malaysia)
Amornphun Phunopas (King Mongkut's University of Technology, North Bankok, Thailand)
Andre Rosendo (ShanghaiTech University, P. R. China)
Anne Jeannin-Girardon (University of Strasbourg, France)
Anton Shiriaev (Norwegian University of Science and Technology, Norway)
Artur Sagitov (Kazan Federal University, Russia)
Ashi Gautam (The University of Tokyo, Japan)
Atsuya Tange (The University of Tokyo, Japan)
Ayumu Tominaga (Kitakyushu National College of Technology, Japan)
Bin Zhang (Beijing University of Posts and Telecommunications, P. R. China)
Cecilia Zanni-Merk (INSA-Rouen, France)
Chaoli Wang (University of Shanghai for Science and Technology, P. R. China)
Chia-Nan Ko (Nan kai University of Technology, Taiwan)
Chung-Wen Hung (National Yunlin University of Science & Technology, Taiwan)
Congdao Han (Shanghai Institute of Technology, P. R. China)
Cui Zhang (Beihang University, P. R. China)
Cynthia Matuszek (University of Maryland, USA)
Donglian Qi (Zhejiang University, P. R. China)
Duangiai Jitkongchuen (Dhurakij Pundit University, Thailand)
Dunwei Gong (China University of Mining and Technology, P. R. China)
Jiao Jia (Beihang University, P. R. China)
Endra Joelianto (Bandung Institute of Technology, Indonesia)
Fabio Guigou (University of Strasbourg, France)
Fei Hao (Beihang University, P. R. China)
Fuzhong Wang (Henan Polytechnic University, P. R. China)
Haibo Li (Royal Institute of Technology, Sweden)
Haisheng Li (Beijing Technology and Business University, P. R. China)

Haruhisa Okuda (Mitsubishi Electric Corporation, Japan)
Haruka Tsuboi (The University of Miyazaki, Japan)
Haruna Matsushita (Kagawa University, Japan)
Heeje Kim (Pusan National University, Korea)
Hidetsugu Suto (Muroran Institute of Technology, Japan)
Hiroyuki Iizuka (Osaka University, Japan)
Hongbo Li (Tsinghua University, P. R. China)
Hongji Ma (Shandong University of Science and Technology, P. R. China)
Hongjiu Yang (Yanshan University, P. R. China)
Hongpeng Yin (Chongqing University, P. R. China)
Hussein Abbass (University of New South Wales, and ADFA, Australia)
I-Hsien Liu (National Cheng Kung Univrsiity, Taiwan)
Ilya Afanasyev (Innopolis University, Russia)
Istvan Harmati (Budapest Institute of Technology and Economics, Hungary)
Ivan Tanev (Doshisha University, Japan)
Jiandong Zhao (Beijing Jiaotong University, P. R. China)
JJ Merelo (University of Granada, Spain)
Joono Cheong (Korea University, Korea)
Julio Navarro Lara (University of Strasbourg, France)
Jun Kobayashi (Kyushu Institute of Technology, Japan)
Jung-Shian Li (National Cheng Kung University, Taiwan)
Keiji Kamei (Nishinippon Institute of Technology, Japan)
Keisuke Watanabe (Tokai University, Japan)
Keita Honda (Gifu University, Japan)
Ke Zhang (Chongqing University, P. R. China)
Kenichi Tanaka (Nagasaki Institute of Applied Science, Japan)
Kensuke Ogata (Osaka City University Japan)
Khairul Salleh bin Mohamed Sahari (Universiti Tenaga Nasional, Malaysia)
Khoyratee Farad (The University of Tokyo, Japan)
Kouichi Takeuchi (Okayama University, Japan)
Konstantin Yakovlev (Russian Academy of Sciences, Higher School of Economics, Russia)
Kui Xiang (Wuhan University of Technology, P. R. China)
Leonid Freidovich (Umea University, Sweden)
Levi Timothée (The University of Tokyo, Japan)
Liming Chen (Beihang University, P. R. China)
Lin Li (University of Shanghai for Science and Technology, P. R. China)
Lin Zhao (Qingdao University, P. R. China)
Mamoru Yoshimura (The University of Miyazaki, Japan)
Manabu Yamada (Nagoya Institute of Technology, Japan)
Masahide Ito (Aichi Prefectural University, Japan)
Masahiro Ushio (Kyushu Institute of Technology, Japan)
Masahiro Yokomichi (The University of Miyazaki, Japan)
Masamichi Hori (The University of Miyazaki, Japan)
Masanori Takahashi (Tokai University, Japan)

Masayoshi Kano (Chukyou University, Japan)
Masayoshi Tabuse (Kyoto Prefectural University, Japan)
Masaomi Hatakeyama (university of Zurich, Switzerland)
Max Talanov (Kazan Federal University, Russia)
Meng Duan (Beihang University, P. R. China)
Moeko Tominaga (Kyushu Institute of Technology, Japan)
Mohammad Al-Shabi (University of Sharjah, United Arab Emirates)
Mohammad Biglarbegian (University of Guelph, Canada)
Mou Chen (Nanjing University of Aeronautics and Astronautics, P. R. China)
Nan Xiao (Northeastern University at Qinhuangdao, P. R. China)
Nicolas Monmarché (university of Tours, France)
Noritaka Sato (Nagoya Institute of Technology, Japan)
Norrima Mokhtar (University of Malaya, Malaysia)
Palakorn Tantrakool (King Mongkut's Institute of Technology, North Bankok, Thailand)
Ping Wang (North China Electric Power University, P. R. China)
Pierre David (University of Strasbourg, France)
Pierre Willaume (University of Strasbourg, France)
Rizauddin bin Ramli (Universiti Kebangsaan Malaysia, Malaysia)
Roman Lavrenov (Kazan Federal University, Russia)
Ruztamreen bin Jenal (Universiti Teknikal Malaysia Melaka, Malaysia)
Ryohei Anshi (Kyushu Institute of Technology, Japan)
R.K.P.S. Ranaweera (University of Moratuwa, Sri Lanka)
Satoshi Ikeda (The University of Miyazaki, Japan)
Sanjay S. Joshi (University of California, USA)
Seong-Ik Han (Pusan National University, Korea)
Shahriman Abu Bakar (Universiti Malaysia Perlis, Malaysia)
Shanbi Wei (Chongqing University, P. R. China)
Shihao Sun (Beihang University, P. R. China)
Shin-ichi Asakawa (Tokyo Woman's Christian University, Japan)
Shin Wakitani (Hiroshima University)
Shumin Fei (Southeast University, P. R. China)
Shyi-Ming Chen (National Taichung University of Education, Taiwan)
Stephen Wilkerson (York College of Pennsylvania, USA)
Takashi Kuremoto (Nippon Institute of Technology, Japan)
Takayoshi Yamada (Gifu University, Japan)
Takuya Fuginaga (Fukuoka University, Japan)
Takuya Nanami (The University of Tokyo. Japan)
Taishiro Kishimoto (Keio University, Japan)
Taisuke Akimoto (Kyushu Institute of Technology, Japan)
Tarik bin Abd Latef (Universiti Malaya, Malaysia)
Taro Shibanoki (Okayama University, Japan)
Tetsuro Katayama (The University of Miyazaki, Japan)
Thomas Noel (University of Strasbourg, France)
Thunyaseth Sethaput (Thammasat University, Thailand)

Tianping Zhang (Yangzhou University, P. R. China)
Tomohiko Takagi (Kagawa University, Japan)
Toru Hiraoka (University of Nagasaki, Japan)
Toshihiro Inukai (DENSO Wave Incorporated, Japan)
Toshinori Nawata (Kumamoto National College of Technology, Japan)
Tsunehiro Yoshinaga (Tokuyama National College of Technology, Japan)
Ussath Martin (The University of Potsdam, Germany)
Wan Khairunizam Wan Ahmad (Universiti Malaysia Perlis, Malaysia)
Weicun Zhang (University of Science and Technology Beijing, P. R. China)
Wenhao Zheng (Beihang University, P. R. China)
Wenlin Li (Beihang University, P. R. China)
Wisanu Jiviriya (King Mongkut's University of Technology North Bankok, Thailand)
Xiaocan Wang (Technical University Munich, Germany)
Xiaofeng Su (Beijing Institute of Astronautical Systems Engineering, P. R. China)
Xiaoyan Chen (Tianjin University of Science and Technology, P. R. China)
Xiaoyan Fu (Capital Normal University, P. R. China)
Xuemei Ren (Beijing Institute of Technology, P. R. China)
Xuhui Lu (Beihang University, P. R. China)
Yan Cui (Shanxi Normal University, P. R. China)
Yasunori Takemura (Nishinippon Institute of Technology, Japan)
Yo Horikawa (Kagawa University, Japan)
Yongqiang Qi (China University of Mining and Technology, P. R. China)
Yoshihiro Kita (University of Nagasaki, Japan)
Youji Kawamura (Kindai University, Japan)
Yu-an Zhang (Qinghai University, P. R. China)
Yue Lin (Beijing Institute of Control Engineering, P.R. China)
Yueqian Liang (China Academy of Elecronics and Information Technology, P. R. China)
Yuji Minami (National Institute of Technology, Ube College, Japan)
Yunju Chen ((Shiga University, Japan)
Yunzhong Song (Henan Polytechnic University, P. R. China)
Zacharie Mbaitiga National Institute of Technology, Okinawa College, Japan)
Zakri bin Ghazali (Universiti Malaysia Pahang, Malaysia)
Zengqiang Chen (Nankai University, P. R. China)
Zhao Peng (Huazhong University of Science and Technology, P. R. China)
Zhengquan Yang (Civil Aviation University of China, P. R. China)
Zhongxin Liu (Nankai University, P. R. China)
Zuradzman bin Mohamad Razlan (Universiti Malaysia Perlis, Malaysia)

LOCAL ARRANGEMENT COMMITTEE

Makoto Sakamoto (University of Miyazaki, Japan)
Masanori Sugisaka (ALife Robotics Corporation Ltd., Japan)
Takao Ito (Hiroshima University, Japan)

HISTORY

The International Conference on Artificial Life and Robotics (ICAROB) resulted from the AROB-symposium (International Symposium on Artificial Life and Robotics) whose first edition was held in 1996 and the eighteenth and last edition in 2013. The AROB symposium was annually organized by Oita University and ALife Robotics Corporation Ltd., under the sponsorship of the Science and Technology Policy Bureau, the Ministry of Education, Science, Sports, and Culture (Monbusho), presently, the Ministry of Education, Culture, Sports, Science, and Technology (Monkasho), Japanese Government, Japan Society for the Promotion of Science (JSPS), the Commemorative Organization for the Japan World Exposition ('70), Air Force Office of Scientific Research, Asian Office of Aerospace Research and Development (AFOSR/AOARD), USA. I would like to express my sincere thanks to not only Monkasho (annually fund support from 1996 to 2013) but also JSPS, the Commemorative Organization for the Japan World Exposition ('70), and various other Japanese companies for their repeated support. The old symposium (this symposium has been held every year at B-Con Plaza, Beppu, Oita, Japan except in Oita, Japan (AROB 5th '00) and in Tokyo, Japan (AROB 6th '01).) was organized by the International Organizing Committee of AROB and was co-operated by the Santa Fe Institute (USA), RSJ, IEEJ, ICASE (Now ICROS) (Korea), CAAI (P. R. China), ISCIE, IEICE, IEEE (Japan Council), JARA, and SICE. The old AROB-symposium expanded much by absorbing much new knowledge and technologies into it. This history and character of the former AROB symposiums are passed on the current ICAROB conference and to these journals, Journal of Robotics, Networking and Artificial Life (JRNAL)(vol1-8) indexed by SCOPUS & ESCI and Journal of Robotics, Networking and Artificial Life (JRNAL)(vol9-) indexed by SCOPUS & ESCI and Journal of Advances in Artificial Life Robotics (JAALR). From now on, ALife Robotics Corporation Ltd. is in charge of management of both the conference and the journals. The future of the ICAROB is brilliant from a point of view of yielding new technologies to human society in the 21st century. We also expect to establish an international research institute on Artificial Life and Robotics in the future with the help of Japanese Government and ICAROB. This conference invites you all.

AIMS AND SCOPE

The objective of this conference is the development of new technologies for artificial life and robotics which have been recently born in Japan and are expected to be applied in various fields. This conference presents original technical papers and authoritative state-of-the-art reviews on the development of new technologies concerning robotics, networking and artificial life and, especially computer-based simulation and hardware for the twenty-first century. This conference covers a broad multidisciplinary field, including areas such as:

Artificial intelligence & complexity
Artificial living
Artificial mind research
Artificial nervous systems for robots
Artificial sciences
Bipedal robot
Brain science and computing
Chaos
Cognitive science
Computational Molecular biology
Computer graphics
Data mining
Disasters robotics
DNA computing
Empirical research on network and MOT
Environment navigation and localization
Evolutionary computations
Facial expression analysis, music recommendation and augmented reality
Foundation of computation and its application
Fuzzy control
Genetic algorithms
Human-welfare robotics
Image processing
Insect-like aero vehicles
Intelligence in biological systems
Intelligent control
Management of technology
Medical surgical robot
Micro-machines
Multi-agent systems
Nano-biology
Nano-robotics
Networking
Neural circuits
Neuro-computer

Neuromorphic Systems
Neuroscience
Pattern recognition
Quantum computing
Reinforcement learning system & genetic programing
Robotics
Software development support method
System cybernetics
Unmanned underwater vehicles
Unmanned Aerial Systems Technologies
Unmanned Aerial Systems designing, controls and navigation
Unmanned Aero vehicles
Virtual reality
Visualization
Hardware-oriented submissions are particularly welcome. This conference will discuss new results in the field of artificial life and robotics

COPYRIGHTS

Accepted papers will be published in the proceeding of The 2024 International Conference on Artificial Life and Robotics (ICAROB2024) by ALife Robotics Corp. Ltd. Copyright belongs to ALife Robotics Corp. Ltd. Some of high-quality papers in the proceeding will be requested to re-submit their papers for the consideration of publication in Journal of Robotics, Networking and Artificial Life indexed by SCOPUS & ESCI (JRNAL)(vol.1-8) & Journal of Robotics, Networking and Artificial Life(vol.9-) indexed by SCOPUS & ESCI and Journal of Advances in Artificial Life Robotics under agreement of both Editor-in- Chief Dr. Masanori Sugisaka and 3 reviewers. All correspondence related to the conference should be addressed to ICAROB Office.

MESSAGES

Masanori Sugisaka
General Chair
(President, ALife Robotics Corp.,
Ltd, Japan)

Masanori Sugisaka

General Chair of ICAROB

It is my great honor to invite you all to The 2024 International Conference on Artificial Life and Robotics (ICAROB 2024) to be held at J:COM HorutoHall, Oita, Japan, 2024. This Conference is changed as the old symposium from the first (1996) to the Eighteenth (2013) annually which were organized by Oita University and ALife Robotics Corporation Ltd. under the sponsorship of the Science and Technology Policy Bureau, the Ministry of Education, Science, Sports, and Culture (Monbusho), presently, the Ministry of Education, Culture, Sports, Science, and Technology (Monkasho), Japanese Government, Japan Society for the Promotion of Science (JSPS), The Commemorative Organization for the Japan World Exposition ('70), Air Force Office of Scientific Research, Asian Office of Aerospace Research and Development (AFOSR/AOARD), USA. I would like to express my sincere thanks to not only Monkasho (annually fund support from 1996 to 2013) but also JSPS, the Commemorative Organization for the Japan World Exposition ('70), Japanese companies for their repeated support.

The old symposium was organized by International Organizing Committee of AROB and was co-operated by the Santa Fe Institute (USA), RSJ, IEEJ, ICASE (Now ICROS) (Korea), CAAI (P. R. China), ISCIE, IEICE, IEEE (Japan Council), JARA, and SICE. The old AROB symposium was growing up by absorbing many new knowledge and technologies into it. This history and character was inherited also from ICAROB2014(The 2014 International Conference on Artificial Life and Robotics, included a series of ICAROB proceedings indexed by SCOPUS and CPCI-Web of Science now. From now on, ALife Robotics Corporation Ltd. is in charge of management. This year we have The 2024 International Conference on Artificial Life and Robotics (ICAROB2024) (29th AROB Anniversary). The future of The ICAROB is brilliant from a point of view of yielding new technologies to human society in 21st century. I have founded Robot Artificial Life Society in 2017/12/07 together with Professor at Hiroshima University Takao Ito and Professor at University of Miyazaki Makoto Sakamoto. I hope that fruitful discussions and exchange of ideas between researchers during Conference (ICAROB2024) will yield new merged technologies for happiness of human beings and, hence, will facilitate the establishment of an international joint research institute on Artificial Life and Robotics in future.

Yingmin Jia

Co-General Chair of ICAROB

It is my great pleasure to invite you to The 2024 International Conference on Artificial Life and Robotics (ICAROB *2024*), will be held at J:COM Horuto Hall, Oita, JAPAN, from February 22 to 25, your understanding and support will be the strongest driving force for us to organize the meeting well.

ICAROB develops from the AROB that was created in 1996 by Prof. Masanori Sugisaka and will celebrate her 29th birthday in 2024. So far many important results have been presented at the past meetings and have a profound impact on artificial life and robotics. Doubtless, it is really one of the most famous international conferences in the field of artificial intelligence and attract wide interests among scientist, researchers, and engineers around the world.

For a successful meeting, many people have contributed their great efforts to the ICAROB. Here, I would like to express my special thanks to all authors and speakers, and the meeting organizing team for their excellent works. Looking forward to seeing you at the ICAROB2024.

Yingmin Jia
Co-General Chair
(Professor, Beihang University, P.R. China)

Takao Ito

Co-General Chair of ICAROB

It is my great honor and pleasure to invite you all to the 2024 International Conference on Artificial Life and Robotics (ICAROB 2024).

The ICAROB has its long history. First launched in 1996 as ISAROB, this former organization of ICAROB, was developed under the strong leadership and yeoman efforts of the President—the internationally famous Professor Masanori Sugisaka, who is widely acknowledged as the father of our AROB conference. Our conference has brought together many research scholars, faculty members, and graduate students from all over the world, and published numerous manuscripts in high-quality proceedings as well as highly reputed journals every year.

Over the years, dramatic improvements have been made in the field of artificial life and its applications. The ICAROB has provided a foundation for unifying the exchange of scientific information on the studies of man-made systems that exhibit the behavioral characteristics of natural living systems, including software, hardware, and wetware. Our conference shapes the development of artificial life, extending our empirical research beyond the territory circumscribed by life-as-we-know-it and into the domain of life-as-it-could–be. It will provide us a good place to present our new research results, innovative ideas, and valuable information about artificial intelligence, complex systems theories, robotics, and management of technology.

The conference site is Horuto Hall, one of the most famous international convention centers in Oita city, Japan. You can find many fantastic scenic spots and splendid historical places in Oita city. Enjoy your stay!

I eagerly look forward to personally meeting you during the ICAROB 2024 and to sharing a most pleasant, interesting, and fruitful conference with you. Do come and make this conference a fruitful, productive as well as enjoyable event!

Takao Ito
Co-General Chair
(Professor Hiroshima University, Japan)

Ju-Jang Lee

Co-General Chair of ICAROB

The First International Conference on Artificial Life and Robotics (ICAROB) was held in Oita City, Oita, Japan from Jan. 11th to 13th, 2014. This year's Conference will be held amidst the high expectation of the increasingly important role of the new interdisciplinary paradigm of science and engineering represented by the field of artificial life and robotics that continuously attracts wide interests among scientist, researchers, and engineers around the globe.

Distinguished researchers and technologists from around the world are looking forward to attending and meeting at ICAROB. ICAROB is becoming the annual excellent forum that represents a unique opportunity for the academic and industrial communities to meet and assess the latest developments in this fast-growing artificial life and robotics field. ICAROB enables them to address new challenges, share solutions, discuss research directions for the future, exchange views and ideas, view the results of applied research, present and discuss the latest development of new technologies and relevant applications.

In addition, ICAROB offers the opportunity of hearing the opinions of well-known leading experts in the field through the keynote sessions, provides the bases for regional and international collaborative research, and enables to foresee the future evolution of new scientific paradigms and theories contributed by the field of artificial life and robotics and associated research area. The twenty-first century will become the century of artificial life and intelligent machines in support of humankind and ICAROB is contributing through wide technical topics of interest that support this direction.

It is a great honor for me as a Co-General Chair of the 11th ICAROB 2024 to welcome everyone to this important event. Also, I would like to extend my special thanks to all authors and speakers for contributing their research works, the participants, and the organizing team of the 11th ICAROB.

I'm looking forward to meeting you at the 11th ICAROB in on line and wishing you all the best.

Ju-Jang Lee
Co-General Chair
(Honorary professor, KAIST)

GENERAL SESSION TOPICS

GS1 Machine Learning & Neural Network & Artificial Life (8)	GS2 Image Processing I (5)
GS3 Image Processing II (3)	GS4 Robotics (6)
GS5 Applications I (4)	GS6 Applications II (4)
GS7 Applications III (6)	

ORGANIZED SESSION TOPICS

OS1 Intelligent Life and Cybersecurity (6)	OS2 Pattern Recognition and Robotics I (5)
OS3 Pattern Recognition and Robotics II (9)	OS4 Pattern Recognition and Robotics III (10)
OS5 Intelligent Life and Robotics (6)	OS6 Intelligence and Optimization (5)
OS7 Deep Learning and its Applications (4)	OS8 Intelligent Control (5)
OS9 Software Development Support Method (4)	OS10 Intelligent Life and Robotics I (6)
OS11 Intelligent Life and Robotics II (6)	OS12 Machine Learning and its Applications (4)
OS13 Robot Control (10)	OS14 Robotic Manipulation (3)
OS15 Artificial Intelligence for Embedded Systems and Robotics (6)	OS16 Industrial Artificial Intelligence Robotics (4)
OS17 Electronics and Kansei Engineering Based on ETT theory (5)	OS18 Computer and Information Engineering (12)
OS19 Natural Computing (3)	OS20 Advances in Field Robotics and Their Applications (15)
OS21 Human-Machine Interface (3)	OS22 Mathematical Informatics (8)
OS23 Industrial Revolution (5)	OS24 Robotics and Intelligent Casting (5)
OS25 Research Towards the Sustainable Development Goals (SDG's) (10)	OS26 Navigating the Digital Frontier: Innovations in the Age of Industry Revolution 4.0 (11)
OS27 Post-narratological Approaches to Cognition in Humans and Robots (5)	

TIME TABLE (2/22)   in Japan time

2/22(Thu.) 17:30-19:30	Welcome Party (Oita Century Hotel)
2/25(Sun.) 15:30-16:00	Farewell party (Room 406 ZOOM ID: 823 7614 7664)

TIME TABLE (2/23)

2/23(Fri.)	Room 403 on-site	Room 404 on-site	Room 405 on-site, online (ZOOM ID: 851 9389 0058)	Room 406 on-site, online (ZOOM ID: 823 7614 7664)	Room 407 online (ZOOM ID: 883 1147 2552)
9:40-	Registration (407)				
10:00-11:15	OS24 Robotics and Intelligent Casting (5) Chair: Jiwu Wang	OS9 Software Development Support Method (4) Chair: Tetsuro Katayama	GS2 Image Processing I (5) Chair: Yui Tanjo	OS23 Industrial Revolution (5) Chair: Hazry Desa on-site, online	OS13-1 Robot Control (5) Chair: Yizhun Peng online
11:15-11:30					
11:30-12:00	Chair: Marion Oswald (Room 302, 303) Opening ceremony			OS2 Pattern Recognition and Robotics I (5) Chair: Fengzhi Dai Will be end at 12:45 online	OS13-2 Robot Control (5) Chair: Yizhun Peng Will be end at 12:45
12:00-13:00	Lunch				
13:00-14:00	Chair: Takao Ito (Room 302, 303) Plenary Speech PS1 Tomoaki Ozaki				OS3-1 Pattern Recognition and Robotics II (5) Chair: Fangyan Li Will be end at 14:15
14:00-14:20					
14:20-15:20	Chair: Marion Oswald (Room 302, 303) Plenary Speech PS2 Hazry Desa				OS3-2 Pattern Recognition and Robotics II (4) Chair: Fangyan Li
15:20-15:40	Coffee break				
15:40-17:10	OS1 Intelligent Life and Cybersecurity (6) Chair: I-Hsien Liu	OS15 Artificial Intelligence for Embedded Systems and Robotics (6) Chair Hakaru Tamukoh	GS4 Robotics (6) Chair: Jiwu Wang onsite, online	GS3 Image Processing II (3) Chair: Seiji Ishikawa online, on-site	OS10 Intelligent Life and Robotics I (6) Chair: Evgeni Magid

16

TIME TEBLE (2/24)

2/24(Sat.)	Room 303 on-site	Room 403 on-site	Room 404 on-site	Room 405 on-site, online (ZOOM ID: 851 9389 0058)	Room 406 on-site, online (ZOOM ID: 823 7614 7664)
9:40-	Registration (407)				
10:00-11:15	OS20-1 Advances in Field Robotics and Their Applications (5) Chair: Kazuo Ishii	OS17 Electronics and Kansei Engineering Based on ETT theory (5) Chair: Tetsuo Hattori	GS1-1 Machine Learning & Neural Network & Artificial Life (5) Chair: Hiroki Tamura	OS19 Natural Computing (3) Chair Marion Oswald Will be end at 10:45	OS18-1 Computer and Information Engineering (5) Chair: Norrima Mokhtar online, onsite
11:15-11:30	Coffee break				
11:30-12:00	OS20-2 Advances in Field Robotics and Their Applications (2) Chair: Kazuo Ishii		GS1-2 Machine Learning & Neural Network & Artificial Life (3) Chair: Hiroki Tamura Will be end at 12:15	OS22 Mathematical Informatics (8) from 11:00 to 13:00 Chair: Takao Ito	OS18-2 Computer and Information Engineering (2) Chair: Norrima Mokhtar online, on-site
12:00-13:00	Lunch				
13:00-14:00	Chair: Yingmin Jia (Room 302, 303) Plenary Speech PS3 Haruhisa Okuda				
14:00-14:15	Coffee break				
14:15-15:30	OS20-3 Advances in Field Robotics and Their Applications (5) Chair: Kazuo Ishii	OS27 Post-narratological Approaches to Cognition in Humans and Robots (5) Chair: Jumpei Ono (ZOOM ID: 826 8397 1634)	OS8 Intelligent Control (5) Chair: Yingmin Jia	OS5-1 Intelligent Life and Robotics (4) Chair: Kuo-Hsien Hsia on-site, online	OS18-3 Computer and Information Engineering (5) Chair: Norrima Mokhtar online, on-site
15:30-15:40	Coffee break				
15:40-16:25	OS20-4 Advances in Field Robotics and Their Applications (3) Chair: Kazuo Ishii		OS14 Robotic Manipulation (3) Chair: Kensuke Harada onsite, online (ZOOM ID: 849 1028 4955)	OS5-2 Intelligent Life and Robotics (2) Chair: Kuo-Hsien Hsia on-site, online	OS21 Human-Machine Interface (3) Chair Norrima Mokhtar online, on-site
16:25-16:40	Coffee break				
16:40-17:40	OS7 Deep Learning and its Applications (4) online Chair Mastaneh Mokayef (ZOOM ID: 824 3203 3781)	OS12 Machine Learning and its Applications (4) Chair: Masato Nagayoshi	GS5 Applications I (4) Chair: onsite, online (ZOOM ID: 849 1028 4955)	OS16 Industrial Artificial Intelligence Robotics (4) Chair Eiji Hayashi on-site	GS6 Applications II (4) Chair Masayuki Fujiwara on-site
18:30-20:30	Banquet (Tokiwa Kaikan)				

17

TIME TEBLE (2/25)

2/25(Sun.)	Room 405 online (ZOOM ID: 851 9389 0058)	Room 406 online (ZOOM ID: 823 7614 7664)	Room 407 online (ZOOM ID: 883 1147 2552)
9:40-	Registration (407)		
10:00-11:30	OS6 Intelligence and Optimization (5) Chair Mastaneh Mokayef	OS26-1 Navigating the Digital Frontier: Innovations in the Age of Industry Revolution 4.0 (6) Chair Wei Hong Lim	OS25-1 Research Towards the Sustainable Development Goals (SDG's) (5) Chair Ammar A.M. Al Talib
11:30-12:30	Lunch		
12:30-13:45	OS4-1 Pattern Recognition and Robotics III (5) Chair Huahao Li	OS26-2 Navigating the Digital Frontier: Innovations in the Age of Industry Revolution 4.0 (5) Chair Wei Hong Lim	OS25-2 Research Towards the Sustainable Development Goals (SDG's) (5) Chair Ammar A.M. Al Talib
13:45-14:00	Coffee break		
14:00-15:30	OS4-2 Pattern Recognition and Robotics III (5) Chair Huahao Li	OS11 Intelligent Life and Robotics (6) Chair: Evgeni Magid	GS7 Applications III (6) Chair: Kasthuri Subaramaniam
15:30-16:00	Farewell Party (Room 406 ZOOM ID: 823 7614 7664)		

The 2024 International Conference on
ARTIFICIAL LIFE AND ROBOTICS
(ICAROB2024)

February 22 (Thursday)

17:30-19:30   Welcome Party (Oita Century Hotel)

February 23 (Friday)

Room 302, 303
11:30-12:00 Opening Ceremony
Chair: Marion Oswald (Vienna University of Technology, Austria)

Welcome Addresses

1. General Chairman of ICAROB   Masanori Sugisaka (ALife Robotics Co., Ltd., Japan)
2. Co-General Chairman of ICAROB   Yingmin Jia (Beihang University, China)
3. Co-General Chairman of ICAROB   TaKao Ito (Hiroshima University, Japan)
4. Vice General Chair of ICAROB   Katia Passerini (Seton Hall University, USA)
5. Vice General Chair of ICAROB   Norrima Mokhtar (University of Malaya, Malaysia)

February 24 (Sunday)

Banquet: Tokiwa Kaikan
18:30-20:30
Chair: Takao Ito (Hiroshima University, Japan)
Welcome Addresses
Prof. Yingmin Jia (Beihang University, P.R. China)
Dr. Norrima Mokhtar (University of Malaya, Malaysia)

TECHNICAL PAPER INDEX

February 23 (Friday)

9:40-Registration

Room 302, 303
11:30-12:00 Opening Ceremony
Chair: Marion Oswald (Vienna University of Technology, Austria)

13:00-14:00
Plenary Speech PS1
Chair: Takao Ito (Hiroshima University, Japan)

PS1 *Developing High-Speed Working Motion of the Multi Robot in DENSO*
Tomoaki Ozaki (DENSO CORPORATION, Japan)

14:20-15:20
Plenary Speech PS2
Chair: Marion Oswald (Vienna University of Technology, Austria)

PS2 *Experimenting with Variable Arm Quadrotors: Realizing Dynamic Configurations for Enhanced Flight Performance*
Hazry Desa (Universiti Malaysia Perlis, Malaysia)

Room 403
10:00-11:15 OS24-Robotics and Intelligent Casting (5)
Chair: Jiwu Wang (Beijing Jiaotong University, China)

OS24-1   *A high-performance motion planning method based on asymptotically optimal RRT*
Tianbin Meng, Jiwu Wang (Beijing Jiaotong University, China)

OS24-2   CSM-RRT*: an improved RRT* algorithm based on constrained sampling mechanism
Hang Yang, Jiwu Wang (Beijing Jiaotong University, China)
Xueqiang Shang (Aero Engine Corporation of China, China)

OS24-3   *Small Sample Object Detection Based on Improved YOLOv5*
Yuxuan Gao (Beijing Jiaotong University, China),
Jiwu Wang (Beijing Jiaotong University, China),
Zixin Li (Aero Engine Corporation of China)

OS24-4   *Research on Gas Pore Prediction Method Based on Sand Core Characteristic Time*
Xiaolong Wang (Beijing Jiaotong University, China), Qihua Wu (Weichai Power Co., Ltd. China), Jiwu Wang (Beijing Jiaotong University, China), Jinwu Kang (Tsinghua University,China),
Na Li (Weichai Power Co., Ltd. China), Yucheng Sun(Weichai Power Co., Ltd. China)

OS24-5   *Optimization Analysis of a Deep Learning-Based Model for Predicting Temperature Fields in the Solidification Process of Castings*
Yahui Yang (Beijing Jiaotong University, China), Jiwu Wang (Beijing Jiaotong University, China),
Jinwu Kang (Tsinghua University,China)

15:40-17:10 OS1 Intelligent Life and Cybersecurity (6)
Chair: I-Hsien Liu (National Cheng Kung University, Taiwan)
Co-Chair: Chu-Fen Li (National Formosa University, Taiwan)
Co-Chair: Pang-Wei Tsai (National Cheng Kung University, Taiwan)

OS1-1   *Detecting abnormal operations in ICS using finite-state machines*
Pei-Wen Chou, Nai-Yu Chen, Jung-Shian Li, I-Hsien Liu
(National Cheng Kung University, Taiwan)

OS1-2   *Industrial Control System State Monitor Using Blockchain Technology*
Yun-Hao Chang, Tzu-En Peng, Jung-Shian Li, I-Hsien Liu
(National Cheng Kung University, Taiwan)

OS1-3   *Enhancing Dam Security and Water Level Alerting with Blockchain Technology*
YingCheng Wu, Jung-Shian Li, Chu-Fen Li, I-Hsien Liu
(National Cheng Kung University, Taiwan)

OS1-4   *MiniDAM: A Dam Cybersecurity Toolkit*
Tzu-En Peng, Meng-Wei Chang, Yun-Hao Chang, Jung-Shian Li, I-Hsien Liu
(National Cheng Kung University, Taiwan)

OS1-5   *Case Study of Network-Based Intrusion Detection System Deployment in Industrial Control Systems with Network Isolation*
Nai-Yu Chen, Pei-Wen Chou, Jung-Shian Li, I-Hsien Liu
(National Cheng Kung University, Taiwan)

OS1-6   *The AI integration service innovation model of real estate industry in Taiwan*
Li-Min Chuang, Chih-Hung Chen (Chang Jung Christian University, Taiwan)

Room 404
10:00-11:00 OS9 Software Development Support Method (4)
Chair: Tetsuro Katayama (University of Miyazaki, Japan)
Co-Chair: Tomohiko Takagi (Kagawa University, Japan)

OS9-1   *An Improved Conversion Technique from EPNAT Models to VDM++ Specifications for Simulation of Abstract Software Behavior*
Sho Matsumoto[1], Ryoichi Ishigami[1], Tetsuro Katayama[2], Tomohiko Takagi[1]
[1] Kagawa University, Japan, [2] University of Miyazaki, Japan

OS9-2   *Prototype of RAGESS Which Is a Tool for Automatically Generating SwiftDiagrams to Support iOS App Development*
Haruki Onaga*, Tetsuro Katayama*, Yoshihiro Kita†, Hisaaki Yamaba*, Kentaro Aburada*, and Naonobu Okazaki*
*University of Miyazaki, Japan, †University of Nagasaki, Japan

OS9-3   *Extension to Support Types and Operation/Function Definitions in BWDM to Generate Test Case Tool from the VDM++ Specification*
Shota Takakura*, Tetsuro Katayama*, Yoshihiro Kita†, Hisaaki Yamaba*, Kentaro Aburada*, and Naonobu Okazaki*
*University of Miyazaki, Japan, †University of Nagasaki, Japan

OS9-4   *Proposal of ASLA Which Is a Segmentation and Labeling Tool for Document Images Based on Deep Learning*
Kanta Kakinoki*, Tetsuro Katayama*, Yoshihiro Kita†, Hisaaki Yamaba*, Kentaro Aburada*, and Naonobu Okazaki*
*University of Miyazaki, Japan, †University of Nagasaki, Japan

15:40-17:10 OS15 Artificial Intelligence for Embedded Systems and Robotics (6)
Chair: Hakaru Tamukoh (Kyushu Institute of Technology, Japan)
Co-Chair Yuma Yoshimoto (National Institute of Technology, Kitakyushu College, Japan)

OS15-1   *YOLO real-time object detection on EV3-Robot using FPGA hardware Accelerator*
Dinda Pramanta[1], Ninnart Fuengfusin[2], Arie Rachmad Syulistyo[2], Hakaru Tamukoh[2]
([1]Kyushu Institute of Information Sciences, Japan) ([2]Kyushu Institute of Technology, Japan)

OS15-2   *A Low Computational Cost Hand Waving Action Recognition System with Echo State Network for Home Service Robots*
Hiromasa Yamaguchi[1], Akinobu Mizutani[1], Arie Rachmad Syulistyo[1], Yuichiro Tanaka[1], Hakaru Tamukoh[1] ([1]Kyushu Institute of Technology, Japan)

OS15-3   *A Rapidly Adjustable Object Recognition System through Language Based Prompt Engineering*
Naoki Yamaguchi[1], Tomoya Shiba[1], Kosei Isomoto[1], Hakaru Tamukoh[1]
([1]Kyushu Institute of Technology, Japan)

OS15-4   *Development of a SayCan-based task planning system capable of handling abstract nouns*
Kosei Yamao[1], Daiju Kanaoka[1], Kosei Isomoto[1], Akinobu Mizutani[1], Yuichiro Tanaka[1], Hakaru Tamukoh[1], ([1]Kyushu Institute of Technology, Japan)

OS15-5   *RoboCup@Home 2023: Stickler for the Rules Task Solutions*
Tomoya Shiba[1], Hakaru Tamukoh[1], ([1]Kyushu Institute of Technology, Japan)

OS15-6   *Offloading Intellectual Processing from Home Service Robots to Edge Devices*
Yuma Yoshimoto[1], Mizuki Kawashima[1], Shun Yonehara[1]
([1]National Institute of Technology, Kitakyushu College, Japan)

Room 405
10:00-11:15 GS2 Image Processing I (5)
Chair: Yui Tanjo (Kyushu Institute of Technology, Japan)

GS2-1   *A Method for Embedding Multiple Photographic Images in a Photographic Image*
Naoki Kouno, Kanya Goto, Toru Hiraoka (University of Nagasaki, Japan)

GS2-2   *Generation of Flowing-Line Images Using Vertical and Horizontal Smoothing Filters*
Karin Kuroki, Toru Hiraoka (University of Nagasaki, Japan)

GS2-3   *Human Motion Recognition from Multiple Directions and Its Gait Cycles Analysis*
Miki Ooba, Yui Tanjo (Kyushu Institute of Technology, Japan)

GS2-4   *A Method of Improving the QOL of the People with Visual Impairment by MY VISION*
Shun Kitazumi, Yui Tanjo (Kyushu Institute of Technology, Japan)

GS2-5   *Human Behavior Segmentation and Recognition Using a Single-camera*
Jing Cao, Yui Tanjo (Kyushu Institute of Technology, Japan)

15:40-17:10 GS4 Robotics (6)
Chair: Jiwu Wang (Beijing Jiaotong University, China)

GS4-1   *Development and evaluation of a learning support robot for vector learning*
Kosei Machida, Shinichi Imai (Tokyo Gakugei University, Japan)

GS4-2   *Feasibility Study on Methods to Measure the Strain on Young Children's Bodies*
Sachiko Kido, Praveen Nuwantha Gunaratne, Hiroki Tamura
(University of Miyazaki, Japan)

GS4-3 *An Integration of Contact Force Models with Multibody Dynamics Analyses for Human Joint Mechanisms and Effects of Viscoelastic Ground Contact*
Shintaro Kasai[1], Dondogjamts Batbaatar[2], Hiroaki Wagatsuma[1]
([1]Kyushu Institute of Technology, Japan; [2]Mongolian University of Science and Technology, Mongolia)

GS4-4 *Haptic Sensation Enhancement via the Stochastic Resonance Effect and Its Application to Haptic Feedback for Myoelectric Prosthetic Hands*
Yoshitaka Mizumoto, Taro Shibanoki (Okayama University, Japan)

GS4-5 *PID Parameter Tuning of a Low-Cost DC Motor Speed Control for Mobile Robot Application*
Munkh-Erdene Ayurzana[1], Erkhembayar Gankhuyag1, Naranbaatar Erdenesuren[1], Dondogjamts Batbaatar[1]
([1]Mongolian University of Science and Technology, Mongolia)

GS4-6 *Reinforcement Learning DDPG Algorithm Based Wheeled Mobility Aid Robot Control Methods*
Junkai Li, Mohd Rizon Mohamad Juhari, Tiang Sew Sun (UCSI University, Malaysia)

Room 406
10:00-11:15 OS23 Industrial Revolution (5)
Chair: Hazry Desa (Universiti Malaysia Perlis (UniMAP), Malaysia)

OS23-1 *Investigating the Engineering Interventions in the Conservation of Malaysia Heritage Structures: A Review on Preserving Historical Edifices Through Advance Civil Engineering Techniques*
Muhammad Azizi Azizan, Nurfadzillah Ishak, Hazry Desa (UniMAP, Malaysia)

OS23-2 *Drones and Data: A Comprehensive Exploration of UAVs in Data Mining*
Muhammad Azizi Azizan, Nurfadzillah Ishak, Hazry Desa (UniMAP, Malaysia)

OS23-3 *Development of Variable Arm to Control the Manoeuvrability of Quadrotor*
L. Y. Hong, H. Desa and M. A. Azizan (UniMAP, Malaysia)
M. H. Tanveer (Kennesaw State University, USA)

OS23-4 *Development of IOT-Enabled Smart Water Metering System*
S. D. Wen, H. Desa and M. A. Azizan (UniMAP, Malaysia)
A. -S. T. Hussain (Al-Kitab University, Iraq)
M. H. Tanveer and R. Patan (Kennesaw State University, USA)

OS23-5 *Object Detection and Instance Segmentation with YOLOV8: Progress and Limitations*
L. J. Lee, H. Desa and M. A. Azizan (UniMAP, Malaysia)
A. -S. T. Hussain (Al-Kitab University, Iraq)
M. H. Tanveer (Kennesaw State University, USA)

11:30-12:45 OS2 Pattern Recognition and Robotics I (5)
Chair: Fengzhi Dai (Tianjin University of Science and Technology, China)
Co-Chair: Yunzhong Song (Henan Polytechnic University, China)

OS2-1 *A Study on Sales Patterns for Vegetable Products in Retail Stores*
Yuhao Zhang [1], Shuangshuang Ma [1], Jiashuai Wang [1], Fengzhi Dai [1], Lijiang Zhang [2]
([1] Tianjin University of Science and Technology, China, [2] Xinjiang Shenhua Biotechnology Co., Ltd, China)

OS2-2 *Research and Implementation of Cooperative Control for ROS Mobile Robot*
Saijie Zhang, Huailin Zhao
(Shanghai Institute of Technology, China)

OS2-3 *An OpenCV-based Method for Workpiece Residue Image Processing*
Jiaxin Wang, Hao He, Fangyv Liu, Fengzhi Dai
(Tianjin University of Science and Technology, China)

OS2-4 *On Nonblockingness Verification and Enforcement of Controlled Nondeterministic Discrete-Event Systems*
Xiang Ren, Zipei Wang (Tianjin University of Science and Technology, China)

OS2-5 *Modeling and Reachability Verification of Controlled Nondeterministic Finite-State Automata*
Zipei Wang, Xiang Ren (Tianjin University of Science and Technology, China)

15:40-16:25 GS3 Image Processing II (3)
Chair: Seiji Ishikawa (Kyushu Institute of Technology, Japan)

GS3-1 *Online Classroom Student Engagement Analysis using Enhanced YOLOv5*
Shuai Wang, Abdul Samad Shibghatullah (UCSI University, Malaysia)

GS3-2 *A Method of Recognizing Body Movements Based on a Self-viewpoint Video*
Iichirou Moribe, Yui Tanjo (Kyushu Institute of Technology, Japan)

GS3-3 *Supporting Safe Walk of a Visually Impaired Person at a Station Platform based on MY VISION*
Shintaro Yamada, Yui Tanjo, Seiji Ishikawa (Kyushu Institute of Technology, Japan)

Room 407
10:00-11:15 OS13-1 Robot Control (5)
Chair: Yizhun Peng (Tianjin University of Science and Technology, China)

OS13-1   *Intelligent Logistics Handling Robot: Design, Control, and Recognition*
Yanchao Bi, Jiale Cheng, Limei Wang, Yizhun Peng
(Tianjin University of Science and Technology, China)

OS13-2   *Greenhouse Design Using Visual Recognition and IoT Technology*
Yuntian Xia, Yizhun Peng
(Tianjin University of Science and Technology, China)

OS13-3   *Design of Modular Photovoltaic Environmentally Friendly Portable Stroller*
Suqing Duan, Yizhun Peng
(Tianjin University of Science and Technology, China)

OS13-4   *"Teenage Mutant Ninja Turtles" - Design of a Bionic Quadrupedal Rescue Robot*
Hongpi Zhao, Yingfan Zhu, Zhihan Zhao, Xin Liang, Lei Lv, Yizhun Peng (Tianjin University of Science and Technology, China)

OS13-5   *Design of Grass Lattice Planter for Complex Environment Based on Adaptive Suspension Technology*
Shaokai Tian, Wenqi Fu, Yizhun Peng
(Tianjin University of Science and Technology, China)

11:30-12:45 OS13-2 Robot Control (5)

Chair: Yizhun Peng (Tianjin University of Science and Technology, China)

OS13-6   *Design of Intelligent Ecological Multifunctional Plant and Animal Breeding System*
Suqing Duan, Yuntian Xia, Siyi Wang, Yizhun Peng (Tianjin University of Science and Technology, China)

OS13-7   *Design of a Fully Automated Logistics Handling Platform*
Hongpi Zhao, Jianfeng Qin, Yizhun Peng (Tianjin University of Science and Technology, China)

OS13-8   *Design and Application of AI-based Brush Calligraphy and Painting Robot*
Haibo Li, Yizhe Sun, Shuxin Wang, Yizhun Peng (Tianjin University of Science and Technology, China)

OS13-9   *Intelligent Wheelchair System: Non-contact Heart Rate and Body Temperature Measurement*
Dongpo Ma, Junsheng Zhang, Yizhun Peng (Tianjin University of Science and Technology, China)

OS13-10   *Recognition and Localization Method for Automotive Axle Holes in Assembly Robots*
Junsheng Zhang, Dongpo Ma, Yizhun Peng (Tianjin University of Science and Technology, China)

13:00-14:15 OS3-1 Pattern Recognition and Robotics II (5)

Chair: Fangyan Li (Tianjin University of Science and Technology, China)

Co-Chair: Haozhe Sun (Tianjin University of Science and Technology, China)

OS3-1   *Detection and Identification of Daylily Maturity Based on YOLOv8*
Fangyan Li (Tianjin University of Science and Technology, China)

OS3-2   *Chaos Synchronization and Circuit Design of Chen System and Lü System with Different Structures*
Haozhe Sun (Tianjin University of Science and Technology, China)

OS3-3   *A Parking Space Recognition Method Based on Digital Image Technology*
Hao He, Fangyv Liu, Jiaxin Wang (Tianjin University of Science and Technology, China)

OS3-4   *A Design of Intelligent Handling Robot Based on AT89C52*
Fangyv Liu, Jiaxin Wang, Hao He (Tianjin University of Science and Technology, China)

OS3-5   *Application and Differences of Robotic Arms, Traditional Machines and Manual Work in Production*
Xue Yang, Ying Su, Yuping Mei, Haiquan Wang (Tianjin University of Science and Technology, China)

14:20-15:20 OS3-2 Pattern Recognition and Robotics II (4)

Chair: Fangyan Li (Tianjin University of Science and Technology, China)

Co-Chair: Haozhe Sun (Tianjin University of Science and Technology, China)

OS3-6   *A Deep Exploration of the Mounting Issues Related to Six Rotor UVA*
Yuping Mei, Ying Su, Xue Yang (Tianjin University of Science and Technology, China)

OS3-7   *The Application of Hexacopter UAV in The Field of Climbing Evasion*
Ying Su, Yuping Mei, Xue Yang (Tianjin University of Science and Technology, China)

OS3-8   *Deep Learning and Embedded Based Operational Safety System for Special Vehicles*
Haoran Gong, Yumei Huang, Jiahao Xie (Tianjin University of Science and Technology, China)

OS3-9   *"Green Fruit" - Intelligent Traceable Agricultural Product Production and Marketing Platform Based on Blockchain Technology*
Yumei Huang, Jiahao Xie, Haoran Gong, Ziyue Xiao (Tianjin University of Science and Technology, China)

15:40-17:10 OS10 Intelligent Life and Robotics (6)
Chair: Evgeni Magid (Kazan Federal University, Russia)
Co-Chair: Kuo-Hsien Hsia (National Yunlin University of Science and Technology, Taiwan)

OS10-1   *A Design of a Modular Mobile Robot for Rescue Operations*
Baris Celiker, Shifa Sulaiman, Tatyana Tsoy (Kazan Federal University, Russia)

OS10-2   *Implementation of Bug1 and Bug2 Path Planning Algorithms for TurtleBot Using ROS Noetic*
Ilya Spektor[1], Aidar Zagirov[2], Ramil Safin[2], Evgeni Magid[1,2]
([1]HSE University, Russia) ([2]Kazan Federal University, Russia)

OS10-3   *Implementation of Alg1 and Alg2 Path Planning Algorithms for Mobile Robots Using ROS Noetic*
Anastasia Yankova1, Timur Gamberov[2], Tatyana Tsoy[2]
([1]HSE University, Russia) ([2]Kazan Federal University, Russia)

OS10-4   *Implementation of VisBug-21 and VisBug-22 Path Planning Algorithms Using ROS Noetic*
Viktoriia Mirzoian[1], Maxim Mustafin[2], Evgeni Magid[1,2]
([1]HSE University, Russia) ([2]Kazan Federal University, Russia)

OS10-5   *DistBug path planning algorithm package for ROS Noetic*
Alexander Pak[1], Alexander Eremin[2], Tatyana Tsoy[2]
([1]HSE University, Russia) ([2]Kazan Federal University, Russia)

OS10-6   *On sensor modeling in Gazebo simulator*
Niez Yuldashev, Alexandra Dobrokvashina, Roman Lavrenov
(Kazan Federal University, Russia)

February 24 (Saturday)
9:40-Registration

Room 302, 303
13:00-14:00
Plenary Speech PS3
Chair: Yingmin Jia

PS3 *Artificial Intelligence and Technologies of Arm-type and Mobile Robots in Industry*

Haruhisa Okuda

(Mitsubishi Electric Corporation, Japan)

Room 303
10:00-11:15 OS20-1 Advances in Field Robotics and Their Applications (5)
Chair: Kazuo Ishii (Kyushu Institute of Technology, Japan)
Co-Chair: Keisuke Watanabe (Tokai University, Japan)

OS20-1 *Image Collection Experiments of a Handy AUV for Offshore Structure Inspection*
Keisuke Watanabe, Koki Amano, Shingen Urano, Yasutaka Taniguchi, Konosuke Watanabe (Tokai University, Japan)

OS20-2 *Design of Disassembly-reassembly Type USV for Coral Reef Research*
Keisuke Watanabe, Koki Amano, Gaku Minato, Yasutaka Taniguchi, Konosuke Watanabe (Tokai University, Japan)

OS20-3 *Optimization method to improve visual SLAM in dynamic environment*
Yufei Liu, Kazuo Ishii (Kyushu Institute of Technology, Japan)

OS20-4 *Design of flexible mechanism for flexible manipulator*
Huang Jiawei, Kazuo Ishii (Kyushu Institute of Technology, Japan)

OS20-5 *Driver Drowsiness Detection Method based on Deep Learning*
Shi Puwei, Kazuo Ishii (Kyushu Institute of Technology, Japan)

11:30-12:00 OS20-2 Advances in Field Robotics and Their Applications (2)
Chair: Kazuo Ishii (Kyushu Institute of Technology, Japan)
Co-Chair: Keisuke Watanabe (Tokai University, Japan)

OS20-6 *An Analysis of Translational Motion for a Mobile Robot with Line-Symmetric Rollers Arrangement*
[1]Kenji Kimura, [2]Kazuo Ishii
([1]National Institute of Technology, Matsue College, Japan, [2]Kyushu Institute of Technology, Japan)

OS20-7 *Development of Teaching Materials for Robot Programming for Junior High School Students: Student-Based Educational Activities*
Kenji Kimura, Youta Takano (National Institute of Technology, Matsue College, Japan)

14:15-15:30 OS20-3 Advances in Field Robotics and Their Applications (5)
Chair: Kazuo Ishii (Kyushu Institute of Technology, Japan)
Co-Chair: Keisuke Watanabe (Tokai University, Japan)

OS20-8 *Design and Software Production of Robotics Educational Design for Elementary and Junior High School Student*
Youta Takano, Kenji Kimura (National Institute of Technology, Matsue College, Japan)

OS20-9 *Development of visual inspection system for low-reflective material utilizing of a string shadow*
Keiji Kamei[1], Tomorou Kawahara[1], Yoshiyuki Daimaru[2]
([1]Nishinippon Institute of Technology, Japan) ([2]Nissan Motor Kyushu, Japan)

OS20-10 *Development of IoT-Based Remote Monitoring Module for Greenhouse Environment to Facilitate Crop Growth Data Analysis*
Moeko Tominaga, Yasunori Takemura, Junya Era, Wataru Kaishita
(Nishinippon Institute of Technology, Japan)
Eiji Mizoe, Tomoyasu Furukawa
(Kumamoto Fruit and Vegetable Shippers Association Co. Ltd., Japan)

OS20-11 *Development greenhouse environment prediction system using IoT data*
Yasunori Takemura, Naoya Nishida, Moeko Tominaga
(Nishinippon Institute of Technology, Japan)
Eiji Mizoe, Tomoyasu Furukawa
(Kumamoto Fruit and Vegetable Shippers Association Co. Ltd., Japan)

OS20-12 *The Development of SaaS for Quantifying the Amount of Drifted Debris on the Coast*
Ayumu Tominaga, Ryohei Komori
(National Institute of Technology Kitakyushu College, Japan)
Eiji Hayashi (Kyushu Institute of Technology, Japan)

15:40-16:25 OS20-4 Advances in Field Robotics and Their Applications (3)
Chair: Kazuo Ishii (Kyushu Institute of Technology, Japan)
Co-Chair: Keisuke Watanabe (Tokai University, Japan)

OS20-13   *Development of an antagonistic wire-driven joint mechanism capable of rapid motion and variable stiffness*
Katsuaki Suzuki[1], Yuya Nishida[2], Kazuo Ishii[2]
([1]Kumamoto Industrial Research Institute, Japan, [2]Kyushu Institute of Technology, Japan)

OS20-14   *Gakken Hills Interdisciplinary Ekiden Competing with Humans, Animals, and Robots*
Takuya Fujinaga[1], Moeko Tominaga[2], Daigo Katayama[3], Kazuo Ishii[3]
([1]Fukuoka University, [2]Nishinippon Institute of Technology, [3]Kyushu Institute of Technology, Japan)

OS20-15   *Development of a Low-Cost Underwater Robot for Research and Education*
Takuya Fujinaga (Fukuoka University Japan)

16:40-17:40 OS7 Deep Learning and its Applications (4)

Chair: Mastaneh Mokayef (UCSI University, Malaysia)

Co-Chair: Takao Ito (Hiroshima University, Japan)

OS7-1   *Parallel Cross Window Attention Transformer and CNN Model for Segmentation of Instrument during Surgery*
Abdul Qayyum[1], M. K. A. Ahamed Khan[3], Moona Mazher[4], Imran Razzak[5], Steven Niederer[1,2], Mastaneh Mokayef[3], C.S. Hassan[3], Ridzuan, A.[3]
([1]Imperial College, London, United Kingdom) ([2]The Alan Turning Institute, London, UK) ([3]UCSI University, Malaysia) ([4]University College London, UK) ([5]University of New South Wales, Australia)

OS7-2   *Magnetic Resonance Spectroscopy (MRS) Reconstruction using Style Transfer Deep Depth wise Framework*
Abdul Qayyum[1], M. K. A. Ahamed Khan[3], Moona Mazher[4], Imran Razzak[5], Steven Niederer[1,2], Mastaneh Mokayef[6], C.S. Hassan[6], Ridzuan, A.[6]
([1]Imperial College, London, UK) ([2]The Alan Turning Institute, UK)
([3,6] UCSI University, Malaysia) ([4]University College London, UK)
([5]University of New South Wales, Australia)

OS7-3   *Federated Learning on Brain Disease Research: Segmentation of Cerebral Small Vessel Diseases (CSVD) using Multi-scale Hybrid Spatial Deep Learning Approach*
Moona Mazher[1], Abdul Qayyum[2], M. K. A. Ahamed Khan[3], Steven Niederer[2,4], Mastaneh Mokayef[3], Ridzuan, A[3], C. S. Hassan[3] ([1]University College London, UK) ([2]National Heart & Lung Institute, Imperial College, London, UK) ([3]UCSI University, Malaysia) ([4]Alan Turning Institute, London, UK)

OS7-4 *Hybrid Classical and Quantum Deep Learning Models for Medical Image Classification*
Moona Mazher[1], Abdul Qayyum[2], M. K. A. Ahamed Khan[3], Steven Niederer[2,4], Mastaneh Mokayef[3], Ridzuan, A[3], C. S. Hassan[3] ([1]University College London, UK,) ([2]National Heart & Lung Institute, Imperial College, UK) ([3]UCSI University, Malaysia) ([4]Alan Turning Institute, UK)

Room 403
10:00-11:15 OS17 Electronics and Kansei Engineering Based on ETT theory (5)
Chair: Tetsuo Hattori (Kagawa University, Hiroshima Institute of Technology, Japan)
Co-Chair: Yusuke Kawakami (NIT (Kagawa College), Japan)

OS17-1 *A High-Speed Estimation Method of Parameters in Impulse Response*
Toshiki Tanaka, Ivan Tanev (Doshisha University, Japan), Tetsuo Hattori (Kagawa University, Japan)

OS17-2 *A Consideration on Amplification Function in BJT Evers-Moll Model and PTT (I) ---- V-I Characteristics ----*
Shimon Hattori, Osamu Matoba (Kobe University, Japan), Tetsuo Hattori (Kagawa University, Japan), Toshiki Tanaka (Kinkei System, Japan), Yusuke Kawakami (NIT (Kagawa College), Japan)

OS17-3 *A Consideration on Amplification Function in BJT Evers-Moll Model and PTT (II) ---- H Parameters in the Small Signal Amplifier Circuit----*
Shimon Hattori, Osamu Matoba (Kobe University, Japan), Tetsuo Hattori (Kagawa University, Japan), Toshiki Tanaka (Kinkei System, Japan), Yusuke Kawakami (NIT (Kagawa College), Japan)

OS17-4 *Color Image Arrangement Using Histogram Matching*
Yusuke Kawakami (NIT (Kagawa College), Japan), Tetsuo Hattori (Kagawa University, Japan), R.P.C. Janaka Rajapakse (Tainan National University of the Arts, Taiwan (R.O.C.))

OS17-5 *Methodology for Creativity Oriented STEM Education Based on ETT Theory*
Tetsuo Hattori, Toshihiro Hayashi, Mai Hattori, Yoshiro Imai (Kagawa University, Japan), Asako Ohno (Osaka Sangyo University, Japan), Takeshi Tanaka (Hiroshima Institute of Technology, Japan)

14:15-15:30 OS27 Post-narratological Approaches to Cognition in Humans and Robots (5)
Chair: Jumpei Ono (Aomori University, Japan)
Co-Chair: Hiroki Fxyma (Kobe University, Japan)

OS27-1   *A Game Framework Based on the Disinformation Warfare in Russo-Ukrainian War*
Jumpei Ono (Aomori University, Japan), Takashi Ogata (Iwate Prefectural University, Japan)

OS27-2   *Comparative Analysis of Eye Tracking between Veteran and Novice during Radiological Interpretation*
Yuka Naito (Chuo University, Japan), Jun Nakamura (Chuo University, Japan), Yoshinobu Ishiwata (Yokohama City University, Japan)

OS27-3   *Development of Notification System to Prevent Working Productivity from Declining Caused by Increased Carbon Dioxide Concentration*
KAMODA Yohei, NAKAMURA Jun (CHUO University, Japan)

OS27-4   *Visualization of the Skilled Physician's Gaze Characteristic during Diagnosis*
Taiki Sugimoto (Chuo University, Japan), Jun Nakamura (Chuo University, Japan), and Yoshinobu Ishiwata (Yokohama City University, Japan)

OS27-5   *Comparative Analysis of Methods for Visualizing the Sensory Experience of Food and Beverages*
Hiroki Fxyma (Kobe University, Japan)

16:40-17:40 OS12 Machine Learning and its Applications (4)
Chair: Masato Nagayoshi (Niigata College of Nursing, Japan)
Co-Chair Takashi Kuremoto (Nippon Institute of Technology, Japan)

OS12-1   *Restoration of Guqin Music by Deep Learning Methods*
Takashi Kuremoto*, Kazuma Fujino*, Hirokazu Takahashi*, Shun Kuremoto**, ***,
Mamiko Koshiba**, Hiroo Hieda***, Shingo Mabu**
(*Nippon Institute of Technology, **Yamaguchi University, ***Institute for Future Engineering, Japan)

OS12-2   *Constructive Nurse Scheduling Using Reinforcement Learning Considering Variations in Nurse Work Patterns*
Masato Nagayoshi (Niigata College of Nursing, Japan),
Hisashi Tamaki (Kobe University, Japan)

OS12-3   *A Basic Study on Indicator of Transfer Learning for Reinforcement Learning*
Satoshi Sugikawa, Kenta Takeoka, Naoki Kotani
(Osaka Institute of Technology, Japan)

OS12-4   *Machine Learning Approach to Predict Cooling Load for Existing Buildings*
Makoto Ohara*, Hideo Isozaki**
(*International Professional University of Technology in Osaka, Japan,
**Kobe University, Japan)

Room 404
10:00-11:15 GS1-1 Machine learning& Neural Network & Artificial Life (5)
Chair: Hiroki Tamura (University of Miyazaki, Japan)

GS1-1   *Deep Learning Based Prediction of Heat Transfer Coefficient Using Spectrogram Images from Boiling Sound*
Fuga Mitsuyama, Ren Umeno, Tomohide Yabuki, Tohru Kamiya (Kyushu Institute of Technology, Japan)

GS1-2   *A Study on Classification of Faulty Motor Sound Using Convolutional Neural Networks*
Jamil Md Shafayet, Praveen Nuwantha Gunaratne, Hiroki Tamura
(University of Miyazaki, Japan)

GS1-3   *Design and Development of a Flexible Active Ankle Joint Orthosis for Locomotion Assistance*
Praveen Nuwantha Gunaratne, Hiroki Tamura (University of Miyazaki, Japan)

GS1-4 *A study on the Real-Time Biomechanical Analysis of Lamber Burden Utilizing Stereoscope Cameras*
Taufik Hidayat Soesilo, Praveen Nuwantha Gunaratne, Hiroki Tamura
(University of Miyazaki, Japan)

GS1-5 *Verification of Determination Possibility using Convolutional Autoencoder for Machine Tool Abnormality Detection*
Yuta Sumoto, Praveen Nuwantha, Hiroki Tamura (University of Miyazaki, Japan)

11:30-12:15 GS1-2 Machine learning& Neural Network & Artificial Life (3)
Chair: Hiroki Tamura (University of Miyazaki, Japan)

GS1-6 *Basic Research for High-speed Heart Sound Determination using AI*
Riku Nakashima, Praveen Nuwantha Gunaratne, Hiroki Tamura
(University of Miyazaki, Japan)

GS1-7 *Prediction of High-Energy Electron Flux at Geosynchronous Orbit using a neural network technique*
Ami Iwabu, Kentaro Kitamura (Kyusyu Institute of Technology, Japan)

GS1-8 *Image Gradient-based Monocular Visual-Inertial Odometry*
Tae Ihn Kim (Hyundai Motor Company, Republic of Korea), Jae Hyung Jung, and Chan Gook Park (Seoul National University, Republic of Korea)

14:15-15:30 OS8 Intelligent Control (5)
Chair: Yingmin Jia (Beihang University, China)
Co-Chair: Weicun Zhang (University of Science and Technology Beijing, China)

OS8-1 *Global Stabilization of A Class of Nonholonomic Integrators via Discontinuous Control*
Lixia Yan, Yingmin Jia (Beihang University (BUAA), China)

OS8-2 *Frequency Dependence Performance Limit of Vibration Absorbers*
Jiqiang Wang[1], Xiaoyu Yin[2], Weicun Zhang[3]
([1]Chinese Academy of Sciences, China)
([2]Science & Technology Bureau of Zhenhai District, China)
([3]University of Science and Technology Beijing, China)

OS8-3 *Adaptive Concurrent Learning Algorithm Based on Pontryagin's Maximum Principle for Nonlinear System Optimal Tracking Control with State Inequality Constraints*
Yuqi Zhang, Bin Zhang (Beijing University of Posts and Telecommunications, China)

OS8-4   *Privacy preserving Mean-square consensus for discrete-time heterogeneous multi-agent systems with Communication Noises*
Tongqing Yang[1], Lipo Mo[1], Yingmin Jia[2]
([1]Beijing Technology and Business University, China)
([2]Beihang University (BUAA), China)

OS8-5   *Event-Triggered Consensus Control for Nonlinear Singular Multi-Agent Systems under Directed Topology*
Lin Li, Tong Yuan, Mei Huang
(University of Shanghai for Science and Technology, China)

15:40-16:25 OS14 Robotic Manipulation (3)
Chair: Kensuke Harada (Osaka University, Japan)
Co-Chair: Akira Nakamura (Saitama Institute of Technology, Japan)
Co-Chair: Tokuo Tsuji (Kanazawa University, Japan)

OS14-1   *Evaluation Standard of Error Recovery Planning Focused on Revival Process from Failures in Robotic Manufacturing Plants*
Akira Nakamura[1], Kensuke Harada[2]
([1]Saitama Institute of Technology, [2]Osaka University, Japan)

OS14-2   *Robotic Food Handling Utilizing Temperature Dependent Variable-Stiffness Material*
Rozilyn Marco[1], Prashant Kumar[2], Xinyi Zhang[2], Weiwei Wan[2] , Kensuke Harada[2]
([1]University of Toronto, Canada, [2]Osaka University, Japan)

OS14-3   Vegetable maturity evaluation for harvest robots
Reno Muhammad Fadilla, Tokuo Tsuji, Tatsuhiro Hiramitsu, and Hiroaki Seki
(Kanazawa University, Japan)

16:40-17:40 GS5 Applications I (4)
Chair: Chair Jiwu Wang (Beijing Jiaotong University, China)

GS5-1   *Unsupervised image registration based on Residual-connected DRMINE for diagnostic metastatic bone tumors*
Shogo Baba[1], Tohru Kamiya[1], Takashi Terasawa[2], Takatoshi Aoki[2]
([1]Kyushu Institute of Technology, Japan), ([2]University of Occupational and Environmental Health, Japan)

GS5-2   *Developing a smart Belt for Monitoring Elderly Activities Based on Multi-Modal Sensors Integration*
and Internet of Things
Abdul Jalil, Pujianti Wahyuningsih, Najirah Umar, Muhammad Risal, Suwatri Jura, A. Edeth Fuari Anatasya (Universitas Handayani Makassar, Indonesia)

GS5-3   *A penalized motion detection model for extracting ionospheric echoes from low signal-to-noise ratio Ionogram video images*
Yuu Hiroshige[1], Akiko Fujimoto[1], Akihiro Ikeda[2], Shuji Abe[3], Akimasa Yoshikawa[3]
([1]Kyushu Institute of Technology, Japan) ([2]National Institute of Technology, Kagoshima College, Japan), (3Kyushu University, Japan)

GS5-4   *Verification Experiments on the Lower Back Burden caused by Posture and Environment during Lifting Operations*
Tomoka Kimura, Yutaro Fujino, Sachiko Kido, Praveen Nuwantha Gunaratne, Hiroki Tamura
(University of Miyazaki, Japan)

Room 405
10:00-10:45 OS19 Natural Computing (3)
Chair: Marion Oswald (Vienna University of Technology, Austria)
Co-Chair: Yasuhiro Suzuki (Nagoya University, Japan)

OS19-1   *A Model of Reaction-diffusion phenomena with Multiset Processing*
Yasuhiro Suzuki (Nagoya University Japan)

OS19-2   *Extract tactile qualities from time series data*
Yasuhiro Suzuki (Nagoya University, Japan)

OS19-3   *Healthcare applications of vibrotactile stimulation developed by Tactile Score*
Yasuhiro Suzuki, Rie Taniguchi (Nagoya University, Japan)

11:00-13:00 OS22 Mathematical Informatics (8)
Chair Takao Ito (Hiroshima University, Japan)
Co-Chair Makoto Sakamoto (University of Miyazaki, Japan)

OS22-1   *Adsorption Behavior of Arsenic and Selenium using NiZn Hydroxy Double Salts with Acetate, Chloride, Nitrate, and Sulfate Anions*
Kaoru Ohe*, Ryosuke Tabuchi, Tatsuya Oshima
(University of Miyazaki, Japan)

OS22-2   *Parallel Acoustic Analysis Based on the Domain Decomposition Method with Higher-Order Element*
Amane Takei[1], Makoto Sakamoto[1], Akihiro Kudo[2]
([1]University of Miyazaki, Japan)
([2]National Institute of Technology, Tomakomai Collage, Japan)

OS22-3   *Sound Field Evaluation on Acoustical Experiment using Non-Steady State Analysis*
Akihiro Kudo[1], Makoto Sakamoto[2], Amane Takei[2]
([1]National Institute of Technology, Tomakomai College, Japan)
([2] University of Miyazaki, Japan)

OS22-4   *A DeepInsight Method with Morphological Analysis*
Toyoaki Tomioka[1], Satoshi Ikeda[1], Makoto Sakamoto[1], Takao Ito[2]
([1]University of Miyazaki, Japan), ([2]Hiroshima University, Japan)

OS22-5   *Support for Museum Exhibition of Small Fungi using AR Technology*
Kakeru Takemura[1], Ota Hamasuna[1], Fumito Hamakawa[1], Satoshi Ikeda[1], Kaoru Ohe[1],
Amane Takei[1], Makoto Sakamoto[1], Shuichi Kurogi[2]
([1]University of Miyazaki, Japan)
([2]Miyazaki Prefectural Museum of Nature and History, Japan)

OS22-6   *Automatic Selection of High-Grade Dried Shiitake Mushrooms using Machine Learning*
Ota Hamasuna[1], Kakeru Takemura[1], Kodai Hasebe[1], Fumito Hamakawa[1], Bidesh
Biswas Biki[1], Satoshi Ikeda[1], Kaoru Ohe[1], Amane Takei[1], Makoto Sakamoto[1],
Kazuhide Sugimoto[2]
([1]University of Miyazaki, Japan), ([2] SUGIMOTO Co., Ltd. Japan)

OS22-7   *Predicting High Volatility Cryptocurrency Prices using Deep Learning*
Tsutomu Ito[1], Kodai Hasebe[2], Fumito Hamakawa[2], Bidesh Biswas Biki[2], Satoshi
Ikeda[2], Amane Takei[2], Makoto Sakamoto[2], Md Riajuliislam[3], Sabrina Bari Shital[4],
Takao Ito[5]
([1]National Institute of Technology, Ube College, Japan)
([2]University of Miyazaki, Japan), ([3] TU Dortmund, Germany)
([4]Daffodil International University, Bangladesh), ([5]Hiroshima University, Japan)

OS22-8   *Simulation of Weathering Representation using Vertex and UV Information*
Tsutomu Ito[1], Fumito Hamakawa[2], Kodai Hasebe[2], Satoshi Ikeda[2], Amane Takei[2],
Makoto Sakamoto[2], Takao Ito[3]
([1]National Institute of Technology, Ube College, Japan),
([2]University of Miyazaki, Japan), ([3]Hiroshima University, Japan)

14:15-15:15 OS5-1 Intelligent Life and Robotics (4)
Chair Kuo-Hsien Hsia (National Yunlin University of Science and Technology, Taiwan)
Co-Chair Evgeni Magid (Kazan Federal University, Russia)

OS5-1   *Research on dynamic obstacle avoidance and complex path planning strategies based on ROS robots*
Yi-Wei Chen, Jr-Hung Guo
(National Yunlin University of Science and Technology, Taiwan)

OS5-2   *Research on Multi-Robot Formation on Two-Dimensional Plane*
Kuo-Hsien Hsia, Chun-Chi Lai, Yi-Ting Liu, Yu-Le Chen
(National Yunlin University of Science and Technology, Taiwan)

OS5-3   *The Development of Utilization Rate and Energy Consumption Monitoring and Networking System*
Chung-Wen HUNG, Chun-Chieh WANG, Heng-En CHANG
(National Yunlin University of Science and Technology, Taiwan)

OS5-4   *MCU Based Edge Computing Platform for Liquid Level Measurement*
Chung-Wen HUNG, Chun-Liang LIU, Tai-Hsuan WU (National Yunlin University of Science and Technology, Taiwan)

15:40-16:10 OS5-2 Intelligent Life and Robotics (2)
Chair Kuo-Hsien Hsia (National Yunlin University of Science and Technology, Taiwan)
Co-Chair Evgeni Magid (Kazan Federal University, Russian Federation)

OS5-5   *Potential of genetic algorithms in multi-UAV coverage problem*
Ramil Faizullin[1], Tatyana Tsoy[1], Edgar A. Martínez-García[2], Evgeni Magid[1,3]
([1]Kazan Federal University, Russian Federation)
([2]The Autonomous University of Ciudad Juarez, Mexico)
([3]HSE University, Russia)

OS5-6   *Construction of Anthropomorphic Grippers with Adaptive Control*
Evgeny Dudorov[1], Julia Zhdanova[2], Ivan Zhidenko[1], Vladimir Moshkin[2], Alexander Eryomin[3], Evgeni Magid[3], Alexander Permyakov[1]
([1]JSC 'SPA 'Android technics', Russia) ([2]MIREA – Russian Technological University, Russia)
([3]Kazan Federal University, Russia)

16:40-17:40 OS16 Industrial Artificial Intelligence Robotics (4)
Chair: Eiji Hayashi (Kyushu Institute of Technology, Japan)

OS16-1   *A Research on Performance Information Editing Support System for Automatic Piano*
Yoshiki Hori, Eiji Hayashi (Kyushu Institute of Technology, Japan)

OS16-2   *Development of Autonomous Mobile Field Robots – Accuracy Verification of Self-Localization through Simulation -*
Takamasa Hayashi[1], Shintaro Ogawa[1], Yuto Okawachi[1], Tan Chi Jie[1], Janthori Titan[1], Ayumu Tominaga[2], Eiji Hayashi[1], Satoko Seino[3]
([1]Kyushu Institute of Technology, Japan), ([2]National Institute of Technology (Kitakyushu College), Japan), ([3]Kyushu University, Japan)

OS16-3   *Development of AR System for Grasping String Foods on Introduction of Industrial Robot*
Yoshihiro Koyama, Eiji Hayashi (Kyushu Institute of Technology, Japan), Akira Kawaguchi (The City College of New York of The City University of New York, United States of America)

OS16-4   *An Image Analysis of Coastal Debris Detection -Detection of microplastics using deep learning*
Yuto Okawachi[1], Ayumu Tominaga[2], Shintaro Ogawa[1], Takamasa Hayashi[1], Tan Chi Jie[1],
Janthori Titan[1], Eiji Hayashi[1], Satoko Seino[3]
([1]Kyushu Institute of Technology, Japan), ([2]National Institute of Technology (Kitakyushu College),
Japan), ([3]Kyushu University, Japan)

Room 406
10:00-11:15 OS18-1 Computer and Information Engineering (5)
Chair: Norrima Mokhtar (University of Malaya, Malaysia)
Co-Chair: Heshalini Rajagopal (UCSI University, Malaysia)

OS18-1   *Efficient Campus Shuttle Tracking and Management Mobile Application for College Campus*
Andrea Tantay Gonzales[1], Kavitha Thamadharan[1], Neesha Jothi[2]
([1]INTI International College Penang, Malaysia) ([2]UCSI University, Malaysia)

OS18-2   *GCN Analysis of Task-Based fMRI Data for Diagnosis of Schizophrenia*
Tejaswini Thota[1], Reuben Stephen John[1], Dr R Dhanush[1], Dr Amutha S[1]
([1] Vellore Institute of Technology, India)

OS18-3   *AR-Based Application for Campus Navigation*
Renuka Devi Rajagopal[1], Akshay S[1], Manoj Rathinam[1], Shakthi B[1], Heshalini Rajagopal[2]
([1] Vellore Institute of Technology, India), ([2]UCSI University, Malaysia)

OS18-4   *Development of Robotic Assistant for Health Care Sector with A Special Focus to Aid the Geriartric Patients*
Narayanan Ganesh (Vellore Institute of Technology, India)

OS18-5   *Development of a Desktop Application Restaurant Management System*
Gabriela Maria Ancilla, Heshalini Rajagopal, Ismail Ahmed Al-Qasem Al-Hadi
(UCSI University, Malaysia)

11:30-12:00 OS18-2 Computer and Information Engineering (2)
Chair: Norrima Mokhtar (University of Malaya, Malaysia)
Co-Chair: Heshalini Rajagopal (UCSI University, Malaysia)

OS18-6   *Face Recognition based on Attendance System*
Koh Pei Cong[1], Heshalini Rajagopal[1], Ghassan Saleh[1], Norrima Mokthar[2]
([1]UCSI University, Malaysia), ([2]University of Malaya, Malaysia)

OS18-7   *U-Reserve: Development of a Facility Reservation System for UCSI University*
Esther Chong Jun Lynn[1], Neesha Jothi[1], Ismail Ahmed[1]
([1]UCSI University, Malaysia)

14:15-15:30 OS18-3 Computer and Information Engineering (5)
Chair: Norrima Mokhtar (University of Malaya, Malaysia)
Co-Chair: Heshalini Rajagopal (UCSI University, Malaysia)

OS18-8   *The Smart Document Processing with Artificial Intelligence*
Raenu Kolandaisamy[1], Heshalini Rajagopal[1], Indraah K[2], Glaret Shirley Sinnappan[3]
([1] UCSI University, Malaysia), ([2]University Utara Malaysia, Kedah)
  ([3]Tunku Abdul Rahman University of Management and Technology, Malaysia)

OS18-9   *Digital Security Challenges Faced by Business Organizations*
Raenu Kolandaisamy[1], Heshalini Rajagopal[1], Indraah K[2], Glaret Shirley Sinnappan[3]
([1]UCSI University, Malaysia), ([2]University Utara Malaysia, Malaysia), ([3]Tunku Abdul Rahman University of Management and Technology, Malaysia)

OS18-10   *The Study on Perception on E-Waste Among the People*
Raenu Kolandaisamy[1], Heshalini Rajagopal[1], Indraah K[2]
([1]UCSI University, Malaysia), ([2]University Utara Malaysia, Malaysia)

OS18-11   *Emergence of Cybercrimes in Online Social Networks*
Raenu Kolandaisamy[1], Heshalini Rajagopal[1], Indraah K[2]
([1]UCSI University, Malaysia) ([2]University Utara Malaysia, Malaysia)

OS18-12   *Development of a music recommendation application by using facial emotion recognition*
Shengke Xie, Raenu Kolandaisamy, Ghassan Saleh, Heshalini Rajagopal
(UCSI University, Malaysia)

15:40-16:25 OS21 Human-Machine Interface (3)
Chair: Norrima Mokhtar (University of Malaya, Malaysia)
Co-Chair: Heshalini Rajagopal (UCSI University, Malaysia)

OS21-1   *Enhancing Reconnaissance Missions Through Multiple Unmanned Systems in ROS*
Anees ul Husnain[1,2], Norrima Mokhtar[1], Takao Ito[3], Siti Sendari[4], Muhammad Farris Kyasudeen[5], Muhammad Badri M Noor[1,6], Heshalini Rajagopal[7]
([1] Universiti Malaya, Malaysia), ([2] The Islamia University of Bahawalpur, Pakistan), ([3]Hiroshima University, Japan), ([4] Universitas Negeri Malang, Indonesia), ([5] University Technology MARA (UiTM), Malaysia)
([6] Ifcon Technology Sdn Bhd, Malaysia), ([7]UCSI University, Malaysia)

OS21-2   *Illumination Effects on Facial Expression Recognition using Empirical Mode Decomposition*
Hashimah Ali[1], Wan Khairunizam Wan Ahmad[1], Hariharan Muthusamy[2], Mohamed Elshaikh[1]
([1] Universiti Malaysia Perlis, Malaysia)
([2] National Institute of Technology Uttarakhand, India)

OS21-3   *Supercontinuum Generation Pump By a Molydenum Disulfide Based Soliton Mode-Locked Fiber Laser*
Aeryn D. Ahmad[1], Norrima Mokhtar[1], Hamzah Arof[1], Sulaiman Wadi harun[1], Ahmad Haziq Aiman Rosol[2]
([1]Universiti Malaya, Malaysia), ([2]MJIIT, UTM, Malaysia)

16:40-17:40 GS6 Applications II (4)
Chair: Masayuki Fujiwara (Kyushu Institute of Technology, Japan)

GS6-1   *Development of Smartphone Application for Calculating the Low Back Pain Risk*
Seigo Imura, Praveen Nuwantha, Hiroki Tamura (University of Miyazaki, Japan)

GS6-2   *A Computational Approach for Global Trade Analysis Sensitive to Free Trade Agreement Circumstances: A Case Study Focusing on the Great Mekong Subregion*
Ahmad Altaweel[1], Bo-Young Lee[2], Masayuki Fujiwara[1], Jang-Sok Yoon[2], Hiroaki Wagatsuma[1]
([1]Kyushu Institute of Technology, Japan; [2]Logistics Revolution Korea, Korea)

GS6-3   *Trigger circuit design and system integration for simultaneous measurement of human EEG, motion, and gaze*
Masayuki Fujiwara[1], Phan Hoang Huu Duc[1], Laurent Bougrain[2], Patrick Hénaff[2], Hiroaki Wagatsuma[1] ([1]Kyushu Institute of Technology, Japan; [2]Université de Lorraine, France)

GS6-4   *Terminal Synergetic Controller for Car's Active Suspension System Using Dragonfly Algorithm*
Tinnakorn Kumsaen[1], Sorn Simatrang[2], Arsit Boonyaprapasorn[3], Thunyaseth Sethaput[4]
([1]Khon Kaen University, Thailand; [2]Nacres Co., Ltd, Thailand; [3]Chulachomklao Royal Military Academy, Thailand; [4]Thammasat University, Thailand)

February 25 (Sunday)
9:40-Registration

10:00-11:00
Room 405
10:00-11:15 OS6 Intelligence and Optimization (5)
Chair: Mastaneh Mokayef (UCSI University, Malaysia)
Co-Chair Takao Ito (Hiroshima University, Japan)

OS6-1   *Simulation-Based Enhancement of SNR in Drone Communication through Uniform Linear Array Configurations*
Gershom Phiri[1], Mastaneh Mokayef[1], MHD Amen Summakieh[1], M.K.A Ahamed Khan[1], Sew Sun Tiang[1], Wei Hong Lim[1], Abdul Qayyum[2]
([1]UCSI University Malaysia)
([2]National Heart and Lung Institute, Imperial College London, UK)

OS6-2   *Empowering Elderly Individuals through the Intelligent Shopping Trolley*
Mastaneh Mokayef[1], Muzaiyanah Binti Hidayab[1], MHD Amen Summakieh[1], M.K.A Ahamed Khan[1], Kim Soon Chong[1], Chin Hong Wong[2], Chua Huang Shen[3], Abdul Qayyum[4]
([1]UCSI University Malaysia) ([2]Fuzhou University, China) ([3]UOW Malaysia University, Malaysia) ([4]National Heart and Lung Institute, Imperial College, UK)

OS6-3   *Optimized Microstrip Slot UWB Patch Antenna for Medical Imaging*
Maxime Duvacher[1], Mastaneh Mokayef[2], MHD Amen Summakieh[2], M.K.A Ahamed Khan[2], Sew Sun Tiang[2], Wei Hong Lim[2], Abdul Qayyum[3]
([1]Polytech Nantes, France) ([2]UCSI University Malaysia) ([3]National Heart and Lung Institute, Imperial College London, UK)

OS6-4   *Development of an Innovative Undergraduate Industrial Automation and Robotics Degree Program*
M.K.A. Ahamed Khan[1], Mastaneh Mokayef[1], Ridzuan, A.[1], Irraivan Elamvazuthi, Badli Shah Yusoff[2], Abu Hassan Darusman[3] ([1]UCSI University Malaysia) ([2]University Technology Petronos, Malaysia) ([3]UNIKL Malaysia France Institute, Malaysia)

OS6-5   *Smart Assistive Trolley for Elderly Care and Independence*
Dina Ashraf[1], Mastaneh Mokayef[1], MHD Amen Summakieh[1], M.K.A Ahamed Khan[1], Abdul Qayyum[2], Sivajothi A/L Paramasivam[3] ([1]UCSI University Malaysia) ([2]National Heart and Lung Institute, Imperial College London, UK) ([3]UOW Malaysia University College, Malaysia)

12:30-13:45 OS4-1 Pattern Recognition and Robotics III (5)
Chair: Huahao Li (Tianjin University of Science and Technology, China)
Co-Chair Hongshuo Zhai (Tianjin University of Science and Technology, China)

OS4-1   *A Digital Twin Design Based on Robot Workstation*
Huahao Li (Tianjin University of Science and Technology, China)

OS4-2   *A Study of Chemical Reactor Simulation System Based on PCS7*
Hongshuo Zhai (Tianjin University of Science and Technology, China)

OS4-3   *Analysis of Learning Quality Evaluation for University Student Courses with Process Assessment*
Yuhao Zhang, Ying Gong, Xuran Wang
(Tianjin University of Science and Technology, China)

OS4-4   *Motion Analysis and Transfer Applications Based on Posture Recognition*
Yuhao Zhang, Mingyue Li, Jianhao Jiao (Tianjin University of Science and Technology, China)

OS4-5   *Functional Safety Assessment of the Safety Protection System Based on Petri Net*
Peng Wang, Mengyuan Hu
(Tianjin University of Science and Technology, China)

14:00-15:15 OS4-2 Pattern Recognition and Robotics III (5)
Chair: Huahao Li (Tianjin University of Science and Technology, China)
Co-Chair: Hongshuo Zhai (Tianjin University of Science and Technology, China)

OS4-6   *Pedestrian Attribute Recognition Based on Deep Learning*
Peng Wang, Qikun Wang, Shengfeng Wang
(Tianjin University of Science and Technology, China)

OS4-7   *Simulation of office air conditioning air supply based on COMSOL*
Peng Wang, Mengda Liu, Qikun Wang
(Tianjin University of Science and Technology, China)

OS4-8   *Solo Wheel Technology-Self-balancing Wheelbarrow*
Ziyue Xiao, Yumei Huang, Zhencheng Chang, Mingxuan Li
(Tianjin University of Science and Technology, China)

OS4-9   *Second-order self-balancing inverted pendulum*
Ziyue Xiao, Zhencheng Chang, Mingxuan Li, Yumei Huang
(Tianjin University of Science and Technology, China)

OS4-10   *Design of Nanny's Abnormal Behavior Recognition Bracelet Based on Human Activity Recognition (HAR) Deep Learning Model*
Depeng Wang, Yingfan Zhu, Yande Xiang, Ziyue Xiao
(Tianjin University of Science and Technology, China)

Room 406
10:00-11:30 OS26-1 Navigating the Digital Frontier: Innovations in the Age of Industry Revolution 4.0(6)
Chair: Wei Hong Lim (UCSI University, Malaysia)
Co-Chair: Takao Ito (Hiroshima University, Japan)

OS26-1   *An Intelligent Cargo/Warehouse Management System*
Zhongheng Sun, Zhou Yue, Xun Sun, Wenzhuo Fan, Wenxuan Zhou
(Fuzou University, China)

OS26-2   *A Comprehensive Approach to Design and Implement an IoT-Enabled Intelligent Shopping Cart System with Obstacle-Aware Navigation and Enhanced Customer Engagement for Elevated Consumer Experiences*
Yao Chen[1], Jiacheng Du[1], Bo Peng[1], Ningfei Wang[1], Zehan Huang[1], Wei Hong Lim[2], Sew Sun Tiang[2], Mastaneh Mokayef[2], Chin Hong Wong[1]
([1]Fuzhou University, China) ([2]UCSI University, Malaysia)

OS26-3   *Design of a Four-Port Flexible UWB-MIMO Antenna for Wearable and IoT Applications*
Jia Wei Tan, Sew Sun Tiang, Kim Soon Chong, Mohammad Arif Ilyas, Mastaneh Mokayef, Zhi Ying Yeoh, Wei Kang Lai, Wei Hong Lim
(UCSI University, Malaysia)

OS26-4   *Investigate Power Efficiency in PLECS and MATLAB Software by Designing USB 5W Charger*
Zhi Ying Yeoh, Kim Soon Chong, Sew Sun Tiang, Mohammad Arif Ilyas, Jia Wei Tan, Wei Kang Lai, Wei Hong Lim (UCSI University, Malaysia)

OS26-5   *Design and Simulation and Performance of Grid Connected Photovoltaic System for Small, Tall Building in Malaysia*
Wei Kang Lai, Kim Soon Chong, Sew Sun Tiang, Mohammad Arif Ilyas, Jia Wei Tan, Zhi Ying Yeoh, Wei Hong Lim
(UCSI University, Malaysia)

OS26-6   *Design of Dual-Band Coplanar Waveguide (CPW) Printed Antenna for 1.9 - 3.6GHz Applications*
Jia Wei Tan, Sew Sun Tiang, Kim Soon Chong, Mohammad Arif Ilyas, Mastaneh Mokayef, Zhi Ying Yeoh, Wei Kang Lai, Wei Hong Lim
(UCSI University, Malaysia)

12:30-13:45 OS26-2 Navigating the Digital Frontier: Innovations in the Age of Industry Revolution 4.0 (5)
Chair: Wei Hong Lim (UCSI University, Malaysia)
Co-Chair: Takao Ito (Hiroshima University, Japan)

OS26-7 *Optimized Convolutional Neural Network Towards Effective Wafer Defects Classification*
Koon Hian Ang[1], Koon Meng Ang[1], Chin Hong Wong[2,3], Abhishek Sharma[4], Chun Kit Ang[1], Kim Soon Chong[1], Sew Sun Tiang[1], Wei Hong Lim[1]
([1]UCSI University, Malaysia) ([2]Fuzou University, China) ([3]Maynooth University, Ireland) ([4]Graphic Era Deemed to be University, India)

OS26-8 *Tackling Photovoltaic (PV) Estimation Challenges: An Innovative AOA Variant for Improved Accuracy and Robustness*
Rayan Mohammed Noor Mohammed Bakhit[1], Abhishek Sharma[2], Tiong Hoo Lim[3], Chin Hong Wong[4,5], Kim Soon Chong[1], Li Pan[1], Sew Sun Tiang[1], Wei Hong Lim[1]
([1]UCSI University, Malaysia) ([2]Graphic Era Deemed to be University, India) ([3]Universiti Teknologi Brunei, Brunei Darussalam) ([4]Fuzou University, China) ([5]Maynooth University, Ireland)

OS26-9 *Deep Learning in Manufacturing: A Focus on Welding Defect Classification with CNNs*
Tin Chang Ting[1], Hameedur Rahman[2], Tiong Hoo Lim[3], Chin Hong Wong[4,5], Chun Kit Ang[1], Mohamed Khan Afthab Ahamed Khan[1], Sew Sun Tiang[1], Wei Hong Lim[1]
([1]UCSI University, Malaysia) ([2]Air University, Pakistan) ([3]Universiti Teknologi Brunei, Brunei Darussalam) ([4]Fuzou University, China) ([5]Maynooth University, Ireland)

OS26-10 *Enhancing Global Optimization Performance of Arithmetic Optimization Algorithm with a Modified Population Initialization Scheme*
Tin Chang Ting[1], Hameedur Rahman[2], Meng Choung Chiong[1], Mohamed Khan Afthab Ahamed Khan[1], Cik Suhana Hassan[1], Farah Adilah Binti Jamaludin[1], Sew Sun Tiang[1], Wei Hong Lim[1]
([1]UCSI University, Malaysia) ([2]Air University, Pakistan)

OS26-11 *Enhancing Precision Object Detection and Identification for Autonomous Vehicles through YOLOv5 Refinement with YOLO-ALPHA*
Guandong Li[1], Yanzhe Xie[1], Yuhao Lu[1], Jingzhen Fan[1], Yuankui Huang[1], Zongyan Wen[1], Wei Hong Lim[2], Chin Hong Wong[1]
([1]Maynooth University, Ireland) ([2]UCSI University, Malaysia)

14:00-15:30 OS11 Intelligent Life and Robotics (6)
Chair: Evgeni Magid (Kazan Federal University, Russia)
Co-Chair: Kuo-Hsien Hsia (National Yunlin University of Science and Technology, Taiwan)

OS11-1   *An Overview of Kinect Based Gesture Recognition Methods*
Alexander Alexeev[1], Tatyana Tsoy[1], Edgar A. Martínez-García[2], Evgeni Magid[1,3]
([1]Kazan Federal University, Russia)
([2]The Autonomous University of Ciudad Juarez, Mexico)
([3]HSE University, Russia)

OS11-2   *An investigation on the impact of human-robot interactions during an autonomous obstacle avoidance task*
Riham Salman, Shifa Sulaiman, Renata Islamova, Tatyana Tsoy
(Kazan Federal University, Russia)

OS11-3   *A Comparative Analysis of Object Detection Methods for Robotic Grasping*
Nikita Kolin, Elvira Chebotareva (Kazan Federal University, Russia)

OS11-4   *Vision-based autonomous navigation for medical examination using a UR3e manipulator*
Bulat Abbyasov[1], Aidar Zagirov[1], Timur Gamberov[1], Hongbing Li[2], Evgeni Magid[1]
([1]Kazan Federal University, Russia)
([2]Shanghai Jiao Tong University, China)

OS11-5   *Robot-Assisted Language Learning: Scientific Data Analysis*
Karina Sadyikova, Valeriya Zhukova, Roman Lavrenov
(Kazan Federal University, Russia)

OS11-6   *Monitoring Beehive Sound Levels with Arduino-based System*
Kulmukhametov Ramis[1], Ramil Safin[1], Tatyana Tsoy[1], Kuo-Hsien Hsia[2], Evgeni Magid[1]
([1]Kazan Federal University, Russia)
([2]National Yunlin University of Science and Technology, Taiwan)

Room 407
10:00-11:15 OS25-1 Research Towards the Sustainable Development Goals (SDG's) (5)
Chair: Ammar A.M. Al Talib (UCSI University, Malaysia)
Co-Chair: Takao Ito (Hiroshima University, Japan)

OS25-1   *Portable Green Energy Mobile Laptop Charging Station*
Ammar A. M. Al-Talib[1], Rodney Tan[1], Ang Aun Jie[1], Idayu M. Tahir[1], Sarah Atifah Saruchi[2], Cik Suhana Bt. Hassan[1], Amar Rizwan[1]
([1]UCSI University) ([2]UMPSA, Malaysia)

OS25-2   *Auto Indoor Hydroponics Plant Growth Chamber*
Ammar A.M. Al-Talib[1], Tew Hwa Hui[1], Sarah Atifah Saruchi[2], Idayu M. Tahir[1], Nor Fazilah Binti Abdullah[1]
([1]UCSI University) ([2]UMPSA, Malaysia)

OS25-3   *A Design and Fabrication of a Solar Agriculture Water Pumping System*
Ammar A.M. Al-Talib[1], Idayu M. Tahir[1], Ain Atiqa[1], Amar Rizwan[1], Sarah Atifah Saruchi[2], Yazan Abu Al shaikh[1]
([1]UCSI University) ([2]UMPSA, Malaysia)

OS25-4   *Design and Performance of a Power Generating Manual Treadmill*
Ammar A. M. Al-Talib[1], Sarah Atifah Saruchi[2], Cik Suhana Bt. Hassan[1], Nor Fazilah Binti Abdullah[1], Ain Atiqa[1], Ahmad Jelban[1]
([1]UCSI University) ([2]UMPSA, Malaysia)

OS25-5   *Smart Car Jack Using Internet of Things*
Idayu M.T. Noor, Ammar A.M. Al- Talib, Mahmoud E.A. Zeiad, Suhana B.H. Cik,
(UCSI University, Malaysia)

12:30-13:45 OS25-2 Research Towards the Sustainable Development Goals (SDG's) (4)
Chair: Ammar A.M. Al Talib (UCSI University, Malaysia)
Co-Chair: Takao Ito (Hiroshima University, Japan)

OS25-6   *Gas Detection for Biogas System Using Internet Of Things (IoT)*
Ammar A.M. Al Talib, I.H.W. Yang, Idayu M.T. Noor, Haslija A.B. Ayu, Afifi. Z. Nur Muhammad
(UCSI University, Malaysia)

OS25-7   *IoT- Based Smart Mushroom Growing Kit*
Ammar A.M. Al- Talib, C.K.J. Ting, Noor Idayu M. Tahir, Ain Atiqa, T.Y. Hui
(UCSI University, Malaysia)

OS25-8   *Design and Analysis of Artificial Magnetic Conductor for Metal Shielding Applications in RFID Car Detection Applications*
Eryana Hussin[1,2], Azman Zakariya[2], Md. Ashraful Haque[2,3], Nur Izzati Ali[4]
([1]UCSI University) ([2]UTPSA, Malaysia) ([3]Daffodil International University, Bangladesh) ([4]UMP, Malaysia)

OS25-9   *Effect of Fibre Orientation on the Mechanical Performance of Natural Fibre Polymer Composite Bicycle Frame using Finite Element Analysis*
Kok Sem Too, Cik Suhana Hassan, Nor Fazilah Abdullah, Ammar Abdulaziz Majeed Al-Talib,
(UCSI University, Malaysia)

OS25-10   *Modelling of short-circuit protection for a residential grid-connected BESS*
Kong De Kang, Farah Adilah Jamaludin, Rodney H.G. Tan
(UCSI University, Malaysia)

14:00-15:30 GS7 Applications III (6)
Chair: Kasthuri Subaramaniam (UCSI University, Malaysia)

GS7-1   *Rehabilitating Flood-Damaged Cars for Sustainable Car Rental Services: A Web-Based System*
Pon Xiao Qi, Abdul Samad Shibghatullah, Kasthuri Subaramaniam
(UCSI University, Malaysia)

GS7-2   *Optimizing E-Invoicing Rollout: Adaptive E-Invoicing Rollout (AER) Framework for Navigating Malaysia's Digital Transformation*
Koh Chee Hong, Abdul Samad Shibghatullah
(UCSI University, Malaysia)

GS7-3   *App Alert System for Smart Phones*
Chee Kin Hoe, Kasthuri Subaramaniam, Abdul Samad Shibghatullah
(UCSI University, Malaysia)

GS7-4   *Developing Hand Gesture Recognition System in Interpreting American Sign Language*
Kong Seh Chong, Kasthuri Subaramaniam, Ismail Ahmed Al-Qasem Al-H
(UCSI University, Malaysia)

GS7-5   Miniature Enterprise Resource Planning
Adim Khalid Aldireejah, Kasthuri Subaramaniam, Ghassan Saleh
(UCSI University, Malaysia)

GS7-6   *Developing Cloud-based Sportswear Website*
Lim Wei Yee, Kasthuri Subaramaniam, Raenu Kolandaisamy
(UCSI University, Malaysia)

Farewell Party

Abstract
PS Abstract (3)
PS1 Developing High-Speed Working Motion of the Multi Robot in DENSO
Tomoaki Ozaki (DENSO CORPORATION, Japan)

The number of labor force is expected to decrease due to the declining birthrate and aging population. Although many efforts to solve this social issue using automation technology by robots around the world have been implemented, most of the current applications of robots are still repetitive works such as picking and placing in mass production lines in factories, and small progress of the application of robots for high-mix low-volume production lines where the operations are frequently changed. In this paper, we discuss the reasons and the potential solutions for autonomous control technologies including the AI and deep learning. Moreover, the high-speed working motion of the multi robot developed by DENSO will be presented in this paper.

High-Speed Working Motion of the Multi-Robot

PS2 Experimenting with Variable Arm Quadrotors: Realizing Dynamic Configurations for Enhanced Flight Performance
Hazry Desa, Muhammad Azizi bin Azizan (Universiti Malaysia Perlis, Malaysia)

This paper introduces two innovative variable arm concepts for quadrotors, enhancing precise movement control by manipulating bending moments through arm length variations. Its key goal is to identify the optimal arm configuration for smooth and stable quadrotor maneuvers. Exploring two concept designs tailored for quadrotors, the study focuses on regulating manoeuvrability using variable arms, enabling bending moment adjustments. Results validate that the electric actuator with linear guide-type 2 variable arm ensures smooth and stable quadrotor movement.
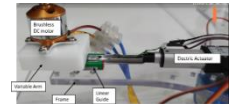
PS3 Artificial Intelligence and Technologies of Arm-type and Mobile Robots in Industry
Haruhisa Okuda (Mitsubishi Electric Corporation, Japan)

In recent years, labor shortage has become a serious issue in industrial fields. Various technologies including robot and information processing system to realize flexible work like humans are effective solutions to this issue. Artificial intelligence technology of arm-type robots equipped with 3D sensors and force sensors has been applied in the manufacturing field to cope with different intelligent and highly precise tasks in Mitsubishi Electric. In addition, various technologies to expand the scope of application to the service field, as well as to realize highly functional delivery with mobile robots is under development. Furthermore, IoT technology is also being used for easy and quick on-site implementation and efficient operation. This speech introduces these initiatives with actual examples.
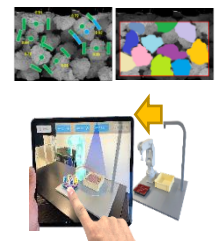
OS Abstract
OS1 Intelligent Life and Cybersecurity (6)
Chair I-Hsien Liu (National Cheng Kung University, Taiwan)
Co-Chair Chu-Fen Li (National Formosa University, Taiwan)
Co-Chair Pang-Wei Tsai (National Cheng Kung University, Taiwan)

OS1-1 Detecting abnormal operations in ICS using finite-state machines
Pei-Wen Chou, Nai-Yu Chen, Jung-Shian Li, I-Hsien Liu (National Cheng Kung University, Taiwan)

In 2021, a water treatment facility in Florida, USA, fell victim to an external malicious attack. In this incident, malicious actors attempted to manipulate the quantities of specific chemicals to impact water quality and safety. Given the intricacies of abnormal operation detection in Industrial Control Systems and the advantages of finite-state machine, we endeavored to apply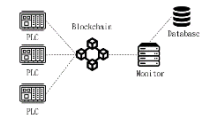 this approach for the detection of abnormal ICS operations. We conducted a series of tests using the dam control system cybersecurity testbed established by TWISC@NCKU, Taiwan. The results indicate that our approach effectively enhances the efficiency of identifying non-standard operational behaviors, enabling maintenance personnel to promptly identify anomalies.
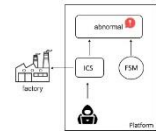
OS1-2 Industrial Control System State Monitor Using Blockchain Technology
Yun-Hao Chang, Tzu-En Peng, Jung-Shian Li, I-Hsien Liu (National Cheng Kung University, Taiwan)
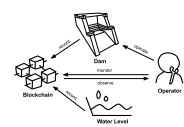
This paper introduces an innovative approach to enhance data verification and security in intelligent systems through the integration of blockchain technology. The proposed method amalgamates the transparency and decentralization inherent to blockchain with the command and oversight functionalities of PLC to ensure the utmost data integrity. The devised approach synergizes the decentralized attributes of blockchain with the control capabilities of PLCs, thus establishing robust safeguards for data integrity. Through the utilization of blockchain's tamper-resistant ledger, PLCs orchestrate data interactions and enforce real-time monitoring and control. The viability and efficacy of this innovative scheme are substantiated through empirical evaluations and simulations, conclusively affirming its practicality.

OS1-3 Enhancing Dam Security and Water Level Alerting with Blockchain Technology
YingCheng Wu, Jung-Shian Li, Chu-Fen Li, I-Hsien Liu (National Cheng Kung University, Taiwan)

Ensuring the security, monitoring, and timely alerting of water levels in dams is a major challenge. We use blockchain technology to enhance the security and monitoring of dam infrastructure, and also improving the alerting system for water level changes. The use of blockchain technology in dam infrastructure management provides a decentralized, transparent, and tamper-resistant platform for storing and managing data. This ensures the integrity and security of critical data related to dam operations and water levels. This research investigates the enhanced security, monitoring, and alerting capabilities that this integration offers, and aims to contribute to the improved security and efficiency of dam infrastructure, leading to more reliable operations and better protection against potential disasters.

OS1-4 MiniDAM: A Dam Cybersecurity Toolkit
Tzu-En Peng, Meng-Wei Chang, Yun-Hao Chang, Jung-Shian Li, I-Hsien Liu
(National Cheng Kung University, Taiwan)

Testbeds, serving as simulations of real-world scenarios, are of paramount importance for research in cybersecurity related to critical infrastructure. In this paper, we aim to offer a comprehensive exploration of the MiniDAM and our testbed, introducing its physical settings based on real dam operational standards. Furthermore, a comparative analysis between the Secure Water Treatment (SWaT) testbed, MiniCPS, our testbed, and MiniDAM is presented. This paper also includes insights into dataset generation and the integration of other functionalities. The exposition of MiniDAM's features and capabilities serves as a foundation for enhancing resilience and provides valuable support for advancing research within the broader field of dam-related studies.

OS1-5 Case Study of Network-Based Intrusion Detection System Deployment in Industrial Control Systems with Network Isolation

Nai-Yu Chen, Pei-Wen Chou, Jung-Shian Li, I-Hsien Liu (National Cheng Kung University, Taiwan)

Deploying intrusion detection systems is a common cybersecurity measure, and intrusion detection systems typically operate at the ports of gateways. In critical infrastructure, industrial control systems often employ network isolation strategies, lacking the role of gateways. This research primarily explores the deployment of the Snort intrusion detection system in such an environment, combined with specific OT rules. Validation is conducted using the cybersecurity testbed of the dam control system established by TWISC@NCKU in Taiwan. The results indicate that by employing our proposed approach, it is possible to effectively detect abnormal network traffic, addressing the common issue of inadequate monitoring in environments with network isolation.

OS1-6 The AI integration service innovation model of real estate industry in Taiwan

Li-Min Chuang, Chih-Hung Chen (Chang Jung Christian University, Taiwan)

In real estate transactions, the intermediary role is often played by real estate agents. In recent years, with the integration of information and AI, the real estate industry can now provide higher quality services. This study references relevant literature and collects the services currently offered by real estate agents. The main dimensions and sub-dimensions are extracted using the Likert scale. Then, the Fuzzy Analytic Hierarchy Process (FAHP) is employed in a questionnaire study to obtain the relative weights among four main dimensions and twelve sub-dimensions. This research develops propositions and conclusions, summarizing the key factors for real estate transactions in Taiwan's real estate industry. These findings serve as important reference points for the industry.
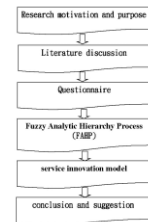
OS2 Pattern Recognition and Robotics I (5)
Chair Fengzhi Dai (Tianjin University of Science and Technology, China)
Co-Chair Yunzhong Song (Henan Polytechnic University, China)

OS2-1 A Study on Sales Patterns for Vegetable Products in Retail Stores

Yuhao Zhang [1], Shuangshuang Ma [1], Jiashuai Wang [1], Fengzhi Dai [1], Lijiang Zhang [2]
([1] Tianjin University of Science and Technology, China, [2] Xinjiang Shenhua Biotechnology Co., Ltd, China)

In fresh produce supermarkets, the shelf life of vegetable products is typically short, necessitating daily restocking based on historical sales data and the formulation of a rational pricing strategy to maximize the store's profits. This paper, based on sales data for vegetable products in a particular store from July 2020 to June 2023, employs various analytical methods, including multidimensional analysis, clustering, and regression, to explore the interrelationships among different types of vegetables. Furthermore, it combines cost-plus pricing and price elasticity models to establish a pricing framework that optimizes revenue for the supermarket.

OS2-2 Research and Implementation of Cooperative Control for ROS Mobile Robot

Saijie Zhang, Huailin Zhao
(Shanghai Institute of Technology, China)

This paper discusses on the distributed control for the multiple ROS-based robots. A communication platform is built to execute the data transmission. The SLAM mapping and autonomous navigation of the robots are completed. The simulation tools of both Gazebo and Rviz are applied to analyze the multi-point navigation. The proposed distributed control system is tried and the synchronous control of the robots is realized, which achieve more accurate and more synchronous robot motion control. At last, the multi-robot following and multi-robot formatting come true.

OS2-3 An OpenCV-based Method for Workpiece Residue Image Processing
Jiaxin Wang, Hao He, Fangyv Liu, Fengzhi Dai (Tianjin University of Science and Technology, China)

Workpiece residue refers to a thin film formed on the surface of the workpiece during the machining process, due to factors such as cutting fluid, chips, oil stains, etc., which affects the quality and performance of the workpiece. This paper proposes an OpenCV-based method for workpiece residue image processing, aiming to achieve automatic detection and analysis of workpiece residue. By building a workpiece image acquisition system, workpiece residue images of different types and degrees are collected, and the proposed method is verified and evaluated. The experimental results show that the proposed method can effectively detect and analyze workpiece residue, with high accuracy and robustness, providing an effective means for workpiece quality control.

OS2-4 On Nonblockingness Verification and Enforcement of Controlled Nondeterministic Discrete-Event Systems
Xiang Ren, Zipei Wang (Tianjin University of Science and Technology, China)

Discrete event systems, as an important kind of cyber-physical systems, have been widely used in engineering field. In this paper, we first express the dynamics of a controlled nondeterministic discrete-event system (acronym is DES) as an algebraic state-space representation using the semi-tensor product (STP) theory. And then, we discuss the problems of state-based nonblockingness verification and enforcement of nondeterministic DESs. Specifically, we obtain a criterion of verifying whether a given controlled nondeterministic DES is nonblocking. Further, we develop an efficient matrix-based approach to enforce state-based nonblockingness. We illustrate the applications of the proposed theoretical results using an example.

OS2-5 Modeling and Reachability Verification of Controlled Nondeterministic Finite-State Automata
Zipei Wang, Xiang Ren (Tianjin University of Science and Technology, China)

In this paper, we investigate the modeling and state reachability of controlled nondeterministic finite-state automata (NFA). The key feature of a controlled NFA is to admit a supervisor to intervene the behavior of original system. We first express the dynamics of a controlled NFA as an algebraic state-space representation in the framework of the semi-tensor product (STP) of matrices. Then, the necessary and sufficient condition for verifying state reachability of controlled NFA is presented. An explicit formula for calculating all paths of any two states is derived. Finally, we use an example to illustrate the application of the proposed theoretical results.
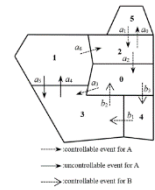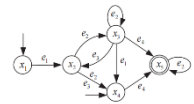
OS3 Pattern Recognition and Robotics II (9)
Chair Fangyan Li (Tianjin University of Science and Technology, China)
Co-Chair Haozhe Sun (Tianjin University of Science and Technology, China)

OS3-1 Detection and Identification of Daylily Maturity Based on YOLOv8
Fangyan Li (Tianjin University of Science and Technology, China)

To better apply object detection and identification techniques from deep learning to the field of agricultural automation, this paper focuses on the growth process of daylilies. It employs the state-of-the-art YOLOv8 model to achieve accurate assessment of daylily maturity. The backbone network of YOLOv8 draws inspiration from the CSPDarkNet network structure to extract image features. Experimental results demonstrate that the detection accuracy of daylilies based on YOLOv8 exceeds 95.6%, with a recall rate of 90.5% and a mean average precision (mAP) reaching 0.94. Moreover, the identification speed is significantly improved.

OS3-2 Chaos Synchronization and Circuit Design of Chen System and Lü System with Different Structures

Haozhe Sun (Tianjin University of Science and Technology, China)

In this paper, by using nonlinear feedback control, chaos synchronization is achieved between the Chen system and the Lü system with different initial values, and the error curves and state synchronization curves of the corresponding states in the response Lü system and the drive Chen system are plotted. Finally, the simulation circuit model of the synchronization system of the drive Chen system and the response Lü system is designed by Multisim circuit simulation software. Comparing the output curves with the curves obtained by MATLAB simulation software, it can be found that the two curves achieve a good qualitative agreement. The synchronization of the drive Chen system and the response Lü system is accomplished.
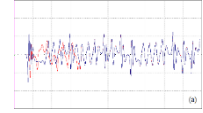
OS3-3 A Parking Space Recognition Method Based on Digital Image Technology

Hao He, Fangyv Liu, Jiaxin Wang (Tianjin University of Science and Technology, , China)

In recent years, the number of cars in the city has been increasing, leading to an increasingly prominent issue of urban parking space. Consequently, the automatic identification method for parking spaces has emerged as a crucial research direction. This paper presents a design and implementation scheme for recognizing the status of parking spaces in urban areas based on digital image processing and other technologies. Real-time images of multiple parking spaces are collected and transmitted for splicing and detection purposes to determine their availability. The experimental results demonstrate the feasibility and effectiveness of the proposed method, which holds practical significance in addressing parking space detection problems through digital image processing.
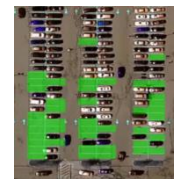
OS3-4 A Design of Intelligent Handling Robot Based on AT89C52

Fangyv Liu, Jiaxin Wang, Hao He (Tianjin University of Science and Technology, China)

Aimed at the shortcomings of the low efficiency and the limitation of the artificial logistics handling, this paper presents a logistics handling robot based on AT89C52 single chip microcomputer. The integration of the power module, sensor, and drive motor module enables automatic obstacle avoidance and information collection.The incorporation of an ultrasonic obstacle avoidance module and an infrared tracking module enhances the capability for obstacle avoidance and path searching, and can easily cope with different workplace. The Yaskawa MPL manipulator is highly suitable for high-speed and high-precision palletizing, picking, packaging, and other industries.

OS3-5 Application and Differences of Robotic Arms, Traditional Machines and Manual Work in Production

Xue Yang, Ying Su, Yuping Mei, Haiquan Wang
(Tianjin University of Science and Technology, China)

This article aims to explore the differences between robotic arms and traditional machines and humans. Firstly, robotic arms are a kind of automated equipment with high flexibility and accuracy, which can perform a variety of complex tasks. In contrast, traditional machines lack flexibility and accuracy, while humans have problems such as low work efficiency and high error rates. Secondly, the emergence of robotic arms can solve many problems in traditional machines and humans, and improve production efficiency and quality. Finally, there is no contradiction between robotic arms, traditional machines and humans, but they can complement each other. By using robotic arms and humans reasonably, we can give full play to their respective advantages and improve the overall production efficiency and economic benefits.

OS3-6 A Deep Exploration of the Mounting Issues Related to Six Rotor UVA
Yuping Mei, Ying Su, Xue Yang (Tianjin University of Science and Technology, China)

In recent years, the hexacopter UAV has developed rapidly, especially the technological breakthrough in the field of automatic driving, which is of great significance in both military and civilian fields. The six-rotor UAV uses six rotors as the power source and adjusts the attitude by changing the rotor speed to further achieve position control. It has excellent hovering ability and sensitivity, and is equipped with a precise positioning system and advanced sensors. However, there are still common problems such as weak mounting capacity and single mounting mode. In view of this problem, this paper will test the fuselage structure, avionics system and power system one by one from two aspects: changing the fuselage structure and installing the motor position.

OS3-7 The Application of Hexacopter UAV in The Field of Climbing Evasion
Ying Su, Yuping Mei, Xue Yang (Tianjin University of Science and Technology, China)

With the development of tourism to climb peaks, the safety of high mountain walls plays an increasingly important role. In the past, the terrain of the mountain wall was too steep and too fast, and people could not patrol the mountain wall climbers, resulting in the safety of climbers could not be guaranteed, and the needs of climbers were difficult to solve in time. In order to solve the above problems, the method of real-time monitoring of UAVs and delivery of materials by UAVs was proposed. Hexacopter UAV is an unmanned small aerial vehicle equipment that can carry out vertical lifting, which is expected to improve the safety factor of climbing mountain walls and solve the needs of climbers in a timely manner.

OS3-8 Deep Learning and Embedded Based Operational Safety System for Special Vehicles
Haoran Gong, Yumei Huang, Jiahao Xie (Tianjin University of Science and Technology, China)

Aiming at the characteristics of special operation vehicles, such as complex working environment, large vehicle weight, long braking distance, and many visual dead angles for drivers, this project designs and implements a special vehicle operation safety system based on deep learning and embedded system. First of all, deep learning model is applied to detect the humanoid target appearing around the vehicle, get the detection frame of the humanoid target and estimate the distance from the target to the camera. According to different distance distances, relevant voice announcements are made to remind the driver to make timely actions such as avoiding or braking the vehicle.

OS3-9 "Green Fruit" - Intelligent Traceable Agricultural Product Production and Marketing Platform Based on Blockchain Technology
Yumei Huang, Jiahao Xie, Haoran Gong, Ziyue Xiao (Tianjin University of Science and Technology, China)

The production process of agricultural products has been greatly improved and optimized through the "Qingguo" intelligent agricultural greenhouse system. Using corrosion-resistant, ageing-resistant, solar-spectrum-converting plastic film, combined with Internet of Things technology, a simple, highly automated and intelligent greenhouse intelligent control system has been created. This not only improves the yield and quality of agricultural products, but also effectively reduces production costs and risks.
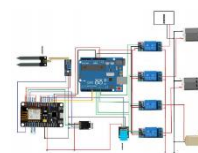
OS4 Pattern Recognition and Robotics III (10)
Chair Huahao Li (Tianjin University of Science and Technology, China)
Co-Chair Hongshuo Zhai (Tianjin University of Science and Technology, China)

OS4-1 A Digital Twin Design Based on Robot Workstation
Huahao Li (Tianjin University of Science and Technology, China)

With the introduction of Made in China 2025 and Industry 4.0, digital twins have rapidly become a trend. By interacting with physical objects and virtual models, mapping is completed in the virtual simulation space to reflect the actual operation process of the entire lifecycle of the corresponding device workstation. This article starts from the research background and significance of digital twin technology, the establishment of Solidworks physical model, the construction of PDPS workstation, and virtual simulation. We have completed the construction of a robot seven color panel assembly workstation model and combined it with PDPS to simulate the workstation.
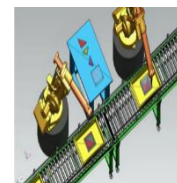
OS4-2 A Study of Chemical Reactor Simulation System Based on PCS7
Hongshuo Zhai (Tianjin University of Science and Technology, China)

This paper introduces a simulation method of plant control, and corrects the system through PID Tuner, and finally gets a reasonable and correct simulation curve. In this paper, PCS7 software is used to build the factory model, and the simulation of four circuits of feed, pressure, liquid level and temperature is built by CFC block in the software. The automatic operation of the system is completed by SFC block. The final simulation curve obtained by PID Tuner tool based on PCS7 software is optimized.

OS4-3 Analysis of Learning Quality Evaluation for University Student Courses with Process Assessment
Yuhao Zhang, Ying Gong, Xuran Wang (Tianjin University of Science and Technology, China)

With the progressive exploration and application of formative assessment in university pedagogy, this evaluative method has become widely adopted for appraising students' everyday learning attitudes and conditions. Drawing upon pertinent data regarding classroom learning experiences of students at a specific university, this paper employs machine learning, K-means clustering, the Topsis evaluation model, and the entropy weighting method to investigate the relationship between formative assessment and the quality of university student learning, culminating in the creation of an evaluation model. This model allows us to pinpoint the key factors influencing student learning attitudes and offers support for formative assessment in the university context.

OS4-4 Motion Analysis and Transfer Applications Based on Posture Recognition
Yuhao Zhang, Mingyue Li, Jianhao Jiao (Tianjin University of Science and Technology, China)

This paper investigates the posture trajectories in human motion using pose recognition technology based on the deep learning framework MediaPipe. By detecting key points on the human body and plotting and calculating these trajectories in the temporal dimension, we further conduct comparative analysis of these trajectories with professional sports coaches' motion guidance to assist athletes in correcting their posture. Additionally, this technology has been deployed on Jetson Nano, enabling its practical application in mobile scenarios and providing robust tools and methods for fields such as rehabilitation therapy, sports training, and animal behavior analysis. This study offers insights into the transfer applications of posture recognition.

OS4-5 Functional Safety Assessment of the Safety Protection System Based on Petri Net
Peng Wang, Mengyuan Hu (Tianjin University of Science and Technology, China)

In this paper, the functional safety evaluation of the safety protection system of gasoline hydrogenation unit was carried out using Petri net. Firstly, the principle and framework of the gasoline hydrogen refueling unit was described. Secondly, the safety integrity level was introduced, and the influencing factors of the safety integrity level were summarized. Thirdly, the Petri net model and the Markov model are compared and the Petri net model is used to verify its security integrity level. Finally,the calculation result demonstrated that the SIL did not reach the target level, and then reached the target level after improvement.This analysis method can provide reference for the safety integrity level evaluation of similar devices.

OS4-6 Pedestrian Attribute Recognition Based on Deep Learning
Peng Wang, Qikun Wang, Shengfeng Wang (Tianjin University of Science and Technology, China)

This paper studied pedestrian attribute recognition based on deep learning, for its importance in the fields of smart city construction. Firstly, the research status of pedestrian attribute recognition and common deep learning models was introduced. Secondly, considering the accuracy decline problem and gradient problem of the neural network, the residual network was used as the main body of the neural network model. Thirdly, the model was trained to classify multiple person attributes through two data sets, Market-1501 and DukeMTMC-reID. Finally, the pedestrian attribute recognition model was tested, and good results were obtained.

OS4-7 Simulation of office air conditioning air supply based on COMSOL
Peng Wang*, Mengda Liu,Qikun Wang (Tianjin University of Science and Technology, China)

This paper analyzed the summer thermal environment of an office in Tianjin University of Science and Technology based on COMSOL software.Firstly,the principle of indoor thermal environment distribution was introduced. Secondly, according to the actual position of the furniture in the office, the mathematical model is constructed based on the basic theory of computational fluid dynamics.Thirdly, the COMSOL software was used for simulation and calculation, combining boundary conditions.Finally, the simulation results are analyzed through the simulated indoor three-dimensional velocity field and temperature field. The simulation results show that the air conditioning supply can well achieve indoor occupants' comfort.

OS4-8 Solo Wheel Technology-Self-balancing Wheelbarrow
Ziyue Xiao, Yumei Huang, Zhencheng Chang, Mingxuan Li
(Tianjin University of Science and Technology, China)

As a new type of personal transportation, self-balancing unicycle has attracted wide attention with its unique design and advanced control system. We aim to study and optimize the self-balancing performance of the unicycle to improve its stability and maneuverability. Firstly, we gain an in-depth understanding of its operation principle. On this basis, an advanced control algorithm is proposed to realize real-time attitude adjustment. Second, for applications in complex environments, we propose an intelligent sensing system to enhance its environment sensing capability.

OS4-9 Second-order self-balancing inverted pendulum

Ziyue Xiao, Zhencheng Chang, Mingxuan Li, Yumei Huang
(Tianjin University of Science and Technology, China)

The team utilizes the second-order inverted pendulum cart based on LQR controller for steady pendulum control with light rods. First, the second-order inverted pendulum is used as a research object to obtain its set of dynamical equations, then, the set of dynamical equations is written in the form of state-space expressions, and finally, the second-order inverted pendulum system of a balanced trolley with a light rod is controlled by a stabilized pendulum using the LQR controller. The inverted pendulum is a typical nonlinear, underdriven and unstable system, which can realize the all-round wind resistance of the rod and occupy less space with high stability.

OS4-10 Design of Nanny's Abnormal Behavior Recognition Bracelet Based on Human Activity Recognition (HAR) Deep Learning Model

Depeng Wang, Yingfan Zhu, Yande Xiang, Ziyue Xiao (Tianjin University of Science and Technology, China)

This article designs a nanny abnormal behavior recognition bracelet. The bracelet is equipped with multiple sensors and a powerful control board, forming a complete nanny abnormal behavior recognition system, which realizes nanny abnormal behavior recognition and alarm in various environments. It uses a gyroscope to collect the three-axis acceleration information of the nanny, and uses the HAR model to infer the nanny's real-time behavior. When the abnormal behavior of the nanny is inferred, STM32 sends the GPS collected positioning information to the employer through ESP32 for timely alarm.

OS5 Intelligent Life and Robotics (6)
Chair Kuo-Hsien Hsia (National Yunlin University of Science and Technology, Taiwan)
Co-Chair Evgeni Magid (Kazan Federal University, Russia)

OS5-1 Research on dynamic obstacle avoidance and complex path planning strategies based on ROS robots

Yi-Wei Chen, Jr-Hung Guo (National Yunlin University of Science and Technology, Taiwan)

Robot Operating System (ROS) is a software system framework used by many robot systems. Although ROS provides a good development environment and related frameworks, ROS is not suitable for public places such as restaurants because of the coming and going of people. Dynamic obstacle avoidance is often handled by stopping the robot, or when there are frequent environmental map changes or when sensors such as optical radar fail, the stop action is also used. However, this often causes path obstructions or delays in completing tasks. Therefore, this study attempts to use images, auxiliary sensors, and various path avoidance strategies to solve the problem of the robot stopping and waiting for the obstacles to disappear. The problem of rapid changes in map paths.

OS5-2 Research on Multi-Robot Formation on Two-Dimensional Plane

Kuo-Hsien Hsia, Chun-Chi Lai, Yi-Ting Liu, Yu-Le Chen
(National Yunlin University of Science and Technology, Taiwan)

Mobile robots are playing an increasingly important role in both service and manufacturing industry. The management of multiple mobile robots is a very important issue on the research of mobile robotics. From a mathematical perspective, this paper discusses the problem of multiple robots on a two-dimensional plane reaching the designated positions in the shortest time to complete formation transformation. We improved the algorithm proposed by Hsia, Li and Su and proposed a new algorithm using a determinant and the Munkres allocation algorithm. Finally, the new algorithm is compared with the path distribution obtained by the Monte Carlo method under different numbers of robots and the excellence of the new algorithm has been verified.
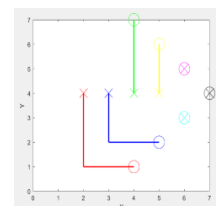
OS5-3 The Development of Utilization Rate and Energy Consumption Monitoring and Networking System

Chung-Wen Hung, Chun-Chieh Wang, Heng-En Chang

(National Yunlin University of Science and Technology, Taiwan)

In this paper, we present a system based on microcontroller unit for measuring the utilization rate of traditional, non-networked machinery. This equipment is designed for use with older machines equipped with andon lights. It employs optocoupler circuits to capture the status of these lights and current transformers to measure their operating and standby currents. Data is transmitted to the server using the Hypertext Transfer Protocol (HTTP) in JavaScript Object Notation (JSON) format. On the server side, a PHP interprets the data, connects to a Structured Query Language (SQL) database, and stores the data using SQL commands. Users can access graphical data through a web-based interface, using it to refine production processes, reduce production costs, and minimize carbon emissions.

OS5-4 MCU Based Edge Computing Platform for Liquid Level Measurement

Chung-Wen Hung, Chun-Liang Liu, Tai-Hsuan Wu

(National Yunlin University of Science and Technology, Taiwan)

An edge computing system based on micro control unit (MCU) for liquid level measurement is proposed in this paper. The system includes a solenoid electromagnet for bottle hit and a microphone to capture sound waves. The signals are converted from time domain to frequency domain by Fast Fourier Transform (FFT), employing an artificial intelligence (AI) model to predict the water level. Artificial Neural Network (ANN) model is applied for classification on the MCU. When optimizing hyperparameters, the accuracy of each parameter combinations should be considered. Ensure the model size suits the limited MCU memory and computing capabilities. Experimental results confirm the system's effectiveness with 99% accuracy.
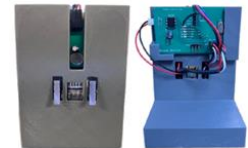
OS5-5 Potential of genetic algorithms in multi-UAV coverage problem

Ramil Faizullin[1], Tatyana Tsoy[1], Edgar A. Martínez-García[2], Evgeni Magid[1,3]

([1]Kazan Federal University, Russian Federation)

([2]The Autonomous University of Ciudad Juarez, Mexico)

([3]HSE University, Russian Federation)

For a rapid area coverage multiple UAVs are often used simultaneously. However, a path planning for a UAVs group during an area coverage task is computationally challenging. In practice, heuristic algorithms are applied to solve this problem. This paper overviews approaches to an area coverage problem with a group of UAVs using genetic algorithms. The article explores modifications that may be useful for a genetic algorithm for solving the coverage problem as well as representation methods for chromosomes that reflect a path of multi-UAV. Additionally, UAV group collision avoidance strategies during area coverage are considered.

OS5-6 Construction of Anthropomorphic Grippers with Adaptive Control

Evgeny Dudorov[1], Julia Zhdanova[2], Ivan Zhidenko[1], Vladimir Moshkin[2], Alexander Eryomin[3], Evgeni Magid[3], Alexander Permyakov[1]
([1]JSC 'SPA 'Android technics', Russian Federation)
([2]MIREA – Russian Technological University, Russian Federation)
([3]Kazan Federal University, Russian Federation)

A functionality of a service robot that physically interacts with a human is provided primarily by capabilities of its end effector. Rather limited capabilities of industrial robot grippers determined a gradual transition to anthropomorphic grippers. A number of degrees of freedom (DoF) of an end effector ensuring reliable grasping and holding of an arbitrary shaped object should be at least nine, preferably twelve. Such design implementation requires to switch toward underactuated grippers systems. This paper proposes a concept of constructing a group drive, which enables a motion of output links of two or more executive groups from a single motor. The presented technical solutions are based on methods of analyzing complex mechanical systems using functional circuits.

OS6 Intelligence and Optimization (5)
Chair Mastaneh Mokayef (UCSI University, Malaysia)
Co-Chair Takao Ito (Takao Ito, Hiroshima University, Japan)

OS6-1 Simulation-Based Enhancement of SNR in Drone Communication through Uniform Linear Array Configurations

Gershom Phiri[1], Mastaneh Mokayef[1], MHD Amen Summakieh[1], M.K.A Ahamed Khan[1], Sew Sun Tiang[1], Wei Hong Lim[1], Abdul Qayyum[2]
([1] UCSI University Malaysia, [2]National Heart and Lung Institute, Imperial College London, UK)

As drones navigate through shared airspace, they often encounter other drones, wireless devices, and communication systems. This coexistence creates potential sources of interference that can degrade the signal-to-noise ratio (SNR). To maintain reliable communication in drone systems, it is crucial to effectively manage and mitigate interference from other drones and wireless devices operating on the same frequency bands. By addressing these challenges, we can ensure a stable and dependable SNR for seamless communication among drones. This paper sheds light on the history of applications and challenges of utilizing flying base stations for wireless networks and analyzes different factors that affect signal-to-noise ratio (SNR) to enhance the performance of drone communication.

OS6-2 Empowering Elderly Individuals through the Intelligent Shopping Trolley

Mastaneh Mokayef[1], Muzaiyanah Binti Hidayab[1], MHD Amen Summakieh[1], M.K.A Ahamed Khan[1], Kim Soon Chong[1], Chin Hong Wong[2], Chua Huang Shen[3], Abdul Qayyum[4]
([1] UCSI University Malaysia, [2] Fuzhou University, China, [3] UOW Malaysia University, Malaysia, [4] National Heart and Lung Institute, Imperial College London, UK)

In this research, we have developed a prototype that aims to improve the weekly shopping experience for senior citizens. Our system tracks the movements of elderly individuals, eliminating the need for them to exert physical force in pushing or pulling the shopping trolley. To achieve this, we utilize a combination of sensors, including a gyroscope and magnetometer, to estimate the user's walking distance and direction. Additionally, we employ WiFi fingerprinting to accurately determine the user's position. Our experiments have yielded satisfactory results in terms of tracking accuracy and the overall functionality of the system. By addressing the specific challenges faced by senior citizens during the routine and essential process of grocery shopping, our smart shopping trolley concept seeks to enhance their experience. Through the integration of tracking technology and sensor-based solutions, we aim to make shopping more convenient and comfortable for elderly individuals. The positive outcomes observed in our experiments validate the effectiveness and feasibility of this approach.

OS6-3 Optimized Microstrip Slot UWB Patch Antenna for Medical Imaging

Maxime Duvacher[1], Mastaneh Mokayef[2], MHD Amen Summakieh[2], M.K.A Ahamed Khan[2], Sew Sun Tiang[2], Wei Hong Lim[2], Abdul Qayyum[3] ([1]Polytech Nantes, France, [2]UCSI University Malaysia, [3]National Heart and Lung Institute, Imperial College London, UK)

This research work presents the development of an Ultra-Wideband (UWB) microstrip patch antenna (MPA) with the specific purpose of tissue characterization. The antenna was carefully designed and simulated to operate within a frequency range of 4.8 to 6.9 GHz, optimized for its intended application. To ensure the best performance, a series of simulations and comparisons were conducted using CST Microwave Studio. Various antenna shapes were tested and evaluated to determine the most effective design. The results of these simulations were highly promising, as they revealed a simulated return loss (S_11) of -33dB. This indicates excellent performance and demonstrates the suitability and acceptability of the proposed antenna for medical imaging such as breast imaging, tumor detection, or monitoring physiological changes.

OS6-4 Development of an Innovative Undergraduate Industrial Automation and Robotics Degree Program

M.K.A. Ahamed Khan[1], Mastaneh Mokayef[1], Ridzuan, A.[1], Irraivan Elamvazuthi, Badli Shah Yusoff [2], Abu Hassan Darusman[3] ([1]UCSI University Malaysia, [2]University Technology Petronos, Malaysia, [3]UNIKL Malaysia France Institute, Malaysia)

In recent years, the need for integrated engineering courses has increased. Due to its multidisciplinary nature, Industrial Automation and Robotics degree course is an ideal example of curriculum integration. This paper discusses several issues such as course offerings, topical content, student profile, student performance and other pertinent matters related to the recent development of an Industrial Automation and Robotics undergraduate degree program at the University of Kuala Lumpur, Malaysia.

OS6-5 Smart Assistive Trolley for Elderly Care and Independence

Dina Ashraf[1], Mastaneh Mokayef[1], MHD Amen Summakieh[1], M.K.A Ahamed Khan[1], Abdul Qayyum[2], Sivajothi A/L Paramasivam ([1]UCSI University Malaysia, [2]National Heart and Lung Institute, Imperial College London, UK, [3] UOW Malaysia University College, Malaysia)

As people age, shopping can become increasingly challenging, especially when it involves pushing heavy trolleys and managing items throughout the entire trip. To address these difficulties, a prototype project was developed with the aim of introducing a robotic trolley. This innovative trolley is designed to autonomously follow senior citizens during their shopping journeys, thanks to face detection technology. The system eliminates the need for physical assistance and provides real-time feedback on the distance covered during the trip. To estimate the walking distance, the trolley incorporates an ultrasonic sensor that activates when the camera detects and tracks the user. The implementation of this project utilizes the OpenCV library, specifically tailored for Python programming. The results of this project have shown significant improvements in the lives of elderly individuals, offering them enhanced comfort and increased independence during their shopping experiences.

OS7 Deep Learning and its Applications (4)
Chair Mastaneh Mokayef (UCSI University, Malaysia)
Co-Chair Takao Ito (Hiroshima University, Japan)

OS7-1 Parallel Cross Window Attention Transformer and CNN Model for Segmentation of Instrument during Surgery

Abdul Qayyum[1], M. K. A. Ahamed Khan[3], Moona Mazher[4], Imran Razzak[5], Steven Niederer[1,2], Mastaneh Mokayef[3], C.S. Hassan[3], Ridzuan, A[3]
([1]Imperial College, London, UK) ([2]The Alan Turning Institute, UK) ([3]UCSI University, Malaysia)
([4]University College London, UK) ([5]University of New South Wales, Australia)

In this work, we present encoder and decoder-based hybrid parallel cross window attention-based transformer during the feature extraction, which consists of the multi-scale channel attention, convolutional layers, and Transformer layers, forming a unified block. Syn-ISS challenge dataset comprised of two tasks. In first task 1, they need to develop deep learning-based method for binary instrument segmentation and in second task multiclass instrument segmentation is required. Experiments conducted on Syn-ISS dataset achieved 0.993 F-score for task 1 and 0.993, 0.975, and 0.951 F-score for shaft, wrist, and jaw segmentation respectively for Task 2.

OS7-2 Magnetic Resonance Spectroscopy (MRS) Reconstruction using Style Transfer Deep Depth wise Framework

Abdul Qayyum[1], M. K. A. Ahamed Khan[3], Moona Mazher[4], Imran Razzak[5], Steven Niederer[1,2], Mastaneh Mokayef [6], C.S. Hassan[6], Ridzuan, A[6]
([1]Imperial College, London, UK) ([2]The Alan Turning Institute, UK)
([3,6]UCSI University, Malaysia) ([4]University College London, UK)
([5]University of New South Wales, Australia)

To analyze the chemical composition of tissues in brain, in vivo magnetic resonance spectroscopy allows non-invasive measurements of neurochemicals in either single voxel or multiple voxels. In this work, we present a deep depth-wise channel attention module (DCAM) based fine-tuned network for magnetic resonance spectroscopy image reconstruction. Besides, we have used channel-wise convolutions and average pooling without dimensionality reduction. We have trained the initial network from scratch on track-1 simulated dataset, however due to the limited dataset, we finetune the network on track-2 and track-3. Experiments are conducted on Edited-MRS-Rec-Challenge dataset1 that showed significantly better performance.

OS7-3 Federated Learning on Brain Disease Research: Segmentation of Cerebral Small Vessel Diseases (CSVD) using Multi-scale Hybrid Spatial Deep Learning Approach

Moona Mazher[1], Abdul Qayyum[2], M. K. A. Ahamed Khan[3], Steven Niederer[2,4], Mastaneh Mokayef[3], Ridzuan, A[3], C. S. Hassan[3]
([1]University College London, UK) ([2]National Heart & Lung Institute, Imperial College, London, UK)
([3]UCSI University, Malaysia) ([4]Alan Turning Institute, London, UK)

In this paper, we propose a hybrid architecture for medical image segmentation to produce efficient representations from global and local features and adaptively aggregate them, aiming to fully exploit their strengths to obtain better segmentation performance in federated learning. Furthermore, we propose a multi-scale feature extraction module embedded at the bottom of the proposed model, which can efficiently extract hidden multi-scale contextual information and aggregate multi-scale features. Experiments on segmentation over three-dimensional rotational angiography of internal Carotid Artery with aneurysm (SHINY-ICARUS) challenge dataset show the effectiveness of the proposed multiscale framework.

OS7-4 Hybrid Classical and Quantum Deep Learning Models for Medical Image Classification

Moona Mazher[1], Abdul Qayyum[2], M. K. A. Ahamed Khan[3], Steven Niederer[2,4], Mastaneh Mokayef[3],
Ridzuan, A[3], C. S. Hassan[3]
([1]University College London, UK) ([2]National Heart & Lung Institute, Imperial College, London, UK)
([3]UCSI University, Malaysia) ([4]Alan Turning Institute, UK)

In this paper, we proposed a hybrid classical and quantum convolutional neural network for Alzheimer's disease (AD) classification. The proposed model was further validated on the brain tumor classification task. The fundamental concept involves encoding data into quantum states, facilitating quicker information extraction, and subsequently utilizing this information to discern the data class. The proposed model results underscore the reliability and robustness and demonstrated by optimal performance accuracies across various datasets, the proposed model substantiates its efficacy in detecting and classifying AD disease and brain tumors.

OS8 Intelligent Control (5)
Chair Yingmin Jia (Beihang University, China)
Co-Chair Weicun Zhang (University of Science and Technology Beijing, China)

OS8-1 Global Stabilization of A Class of Nonholonomic Integrators via Discontinuous Control

Lixia Yan, Yingmin Jia (Beihang University (BUAA), China)

This paper investigates the discontinuous state feedback control for stabilizing a class of nonholonomic integrators with drift terms. The control design relies on constraining state trajectory in an invariant set. To this end, we apply constant controls to drive the states moving into the invariant set and then switch to a continuous control law with suitable gain selections. It is proven in the Lyapunov sense that the proposed control scheme achieves global exponential stabilization of the states, and the control switch would only occur at most once. Numerical simulations are carried out to validate the proposed control law.

$$\mathbb{R}^3 - C \qquad C$$
$$u = u_1 \qquad u = u_2(t)$$
$$C \triangleq \left\{ Z = [z_1, z_2, z_3]' \in \mathbb{R}^3 \mid z_1^2 + z_2^2 > |z_3| \right\}$$

OS8-2 Frequency Dependence Performance Limit of Vibration Absorbers

Jiqiang Wang[1], Xiaoyu Yin[2], Weicun Zhang[3]
([1]Chinese Academy of Sciences, China) ([2]Science & Technology Bureau of Zhenhai District, China)
([3]University of Science and Technology Beijing, China)

Optimal design of vibration absorbers has been extensively investigated. Most of the design methods are approached by optimizing certain performance indices, resulting in a set of optimal parameters that are independent of exogenous forcing frequencies. In practical designs, however, it is often desirable to know the performance limits over a frequency band of interest. This problem is tackled in the present paper where both lower and upper bounds are obtained. A refined upper bound is also derived that can further provide a systematic design methodology. Extensive remarks are also given exploring different avenues useful for design. Numerical examples are given to validate the corresponding designs.
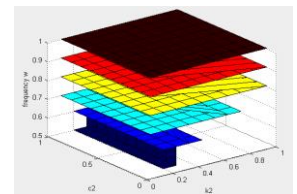
OS8-3 Adaptive Concurrent Learning Algorithm Based on Pontryagin's Maximum Principle for Nonlinear System Optimal Tracking Control with State Inequality Constraints

Yuqi Zhang, Bin Zhang (Beijing University of Posts and Telecommunications, China)

In this article, an adaptive learning algorithm is derived for finite-horizon optimal tracking problems of nonlinear systems. Concurrent learning is implemented to identify the unknown parameters of the system dynamics. Tracking of the desired trajectory and convergence of the developed policy are guaranteed via stability analysis. Different from the adaptive dynamic programming (ADP) infinite-horizon method, the proposed method solves the finite-horizon optimal tracking problems. Simulation results show the effectiveness of the proposed algorithm.
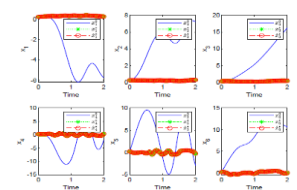
OS8-4 Privacy preserving Mean-square consensus for discrete-time heterogeneous multi-agent systems with Communication Noises

Tongqing Yang[1], Lipo Mo[1], Yingmin Jia[2]

([1]Beijing Technology and Business University, China) ([2]Beihang University (BUAA), China)

This paper studies the privacy preserving consensus of distributed heterogeneous multi-agent systems, which consists of first-order agents and second-order agents. A new protocol is designed for each agent. Then, by applying algebraic graph theory and matrix theory, it is proved that the closed-loop system could achieve consensus. After that, the designed protocol is encrypted by a cryptographic cryptosystem to prevent eavesdropping in the communication link and loss of information privacy between agents while achieving consensus. Finally, the effectiveness of the proposed consensus protocol and privacy protection algorithm are verified by numerical simulations.
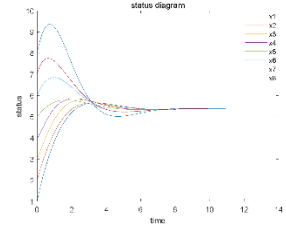
OS8-5 Event-Triggered Consensus Control for Nonlinear Singular Multi-Agent Systems under Directed Topology

Lin Li, Tong Yuan, Mei Huang (University of Shanghai for Science and Technology, China)

This paper is devoted to the problem of event-triggered consensus for a class of nonlinear singular multi-agent systems under directed topology. An event-triggered sampling mechanism is constructed, which naturally avoids Zeno behavior. And then, a distributed event-triggered consensus protocol is designed. By employing the Lyapunov-Krasovskii functional method and model transformation approach, sufficient conditions that can guarantee the consensus of the considered singular multi-agent systems are obtained, while the consensus control gain matrix and the event-triggered parameter are also given. Finally, a numerical example is included to illustrate the effectiveness of the proposed method.
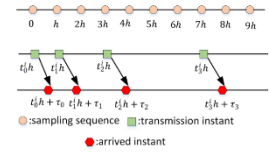
OS9 Software Development Support Method (4)
Chair Tetsuro Katayama (University of Miyazaki, Japan)
Co-Chair Tomohiko Takagi (Kagawa University, Japan)

OS9-1 An Improved Conversion Technique from EPNAT Models to VDM++ Specifications for Simulation of Abstract Software Behavior

Sho Matsumoto[1], Ryoichi Ishigami[1], Tetsuro Katayama[2], Tomohiko Takagi[1]

([1]Kagawa University, Japan) ([2]University of Miyazaki, Japan)

Formal software models based on EPNAT (Extended Place/transition Net with Attributed Tokens) can be converted to VDM++ specifications that enable simulation of abstract software behavior before implementation processes. However, the conversion technique has two problems, that is, (1) extracting all properties to be checked from the VDM++ specifications requires time and effort, and (2) the structure of the VDM++ specifications has less readability and maintainability. In this study, we improve the conversion technique by (1) adding a function to extract an abstract current state of software, and (2) dividing into classes that correspond to subnets of EPNAT models. This paper shows a new conversion rule, a new structure of VDM++ specifications, a simple example, and the discussion about their effectiveness.
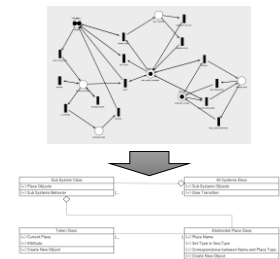
OS9-2 Prototype of RAGESS Which Is a Tool for Automatically Generating SwiftDiagrams to Support iOS App Development

Haruki Onaga[1], Tetsuro Katayama[1], Yoshihiro Kita[2], Hisaaki Yamaba[1], Kentaro Aburada[1], Naonobu Okazaki[1]
([1]University of Miyazaki, Japan) ([2]University of Nagasaki, Japan)

In the development of large-scale and complex mobile applications, it is difficult for developers to continually grasp the overall structure of the app. To support iOS app development, we proposed SwiftDiagram, a visualization of the static structure of Swift source code and confirmed its high usefulness. However, manually drawing SwiftDiagrams is labor-intensive. This paper implements a prototype of RAGESS (Real-time Automatic Generation of SwiftDiagram System), a tool that automatically generates SwiftDiagrams by performing static analysis on Swift source code every time an iOS app build is successful. Compared with other tools, RAGESS is confirmed to enable developers to visualize the static structure of source code effortlessly.
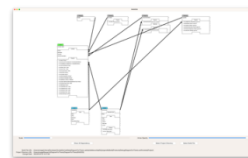
OS9-3 Extension to Support Types and Operation/Function Definitions in BWDM to Generate Test Case Tool from the VDM++ Specification

Shota Takakura[1], Tetsuro Katayama[1], Yoshihiro Kita[2], Hisaaki Yamaba[1], Kentaro Aburada[1], Naonobu Okazaki[1]
([1]University of Miyazaki, Japan) ([2]University of Nagasaki, Japan)

Generating test cases from the formal specification description VDM++, which is a method for disambiguating specifications, is time-consuming and labor intensive. Therefore, our laboratory has developed BWDM, a tool that automatically generates test cases from VDM++ specifications. However, existing BWDM has problems that it only supports integer types and cannot generate test cases for operation and function definitions including recursive structure. Therefore, in order to improve the usefulness of BWDM, this study extends BWDM to solve the above problems. Consequently, it has confirmed that the use of extended BWDM can reduce the test case generation time compared to manual test case generation.

OS9-4 Proposal of ASLA Which Is a Segmentation and Labeling Tool for Document Images Based on Deep Learning

Kanta Kakinoki[1], Tetsuro Katayama[1], Yoshihiro Kita[2], Hisaaki Yamaba[1], Kentaro Aburada[1], Naonobu Okazaki[1]
([1]University of Miyazaki, Japan) ([2]University of Nagasaki, Japan)

The current situation of the electronic documents is only a substitute for paper. As a new way to utilize electronic documents, we focus on dividing electronic documents into regions by their elements and generating keywords and sentences as labels from the contents of the elements. However, these tasks, when performed manually, are time-consuming and labor-intensive. This study proposes a prototype of ASLA (Automatic Segmentation and Labeling tool using AI), segmentation and labeling tool for document images based on deep learning, with the aim of reducing the time required for region segmentation and label generation. To evaluate the usefulness of ASLA, we have compared the time required for region segmentation and label generation using ASLA and by hand, and then confirmed the reduction in time.

OS10 Intelligent Life and Robotics (6)
Chair Evgeni Magid (Kazan Federal University, Russia)
Co-Chair Kuo-Hsien Hsia (National Yunlin University of Science and Technology, Taiwan)

OS10-1 A Design of a Modular Mobile Robot for Rescue Operations

Baris Celiker, Shifa Sulaiman, Tatyana Tsoy (Kazan Federal University, Russia)

Modular robotics is one of the subfields of mobile robotics, which is emerging as a new trend in various sectors. Modular mobile robots can be reconfigured to perform a wide variety of tasks. In this paper, applications of modular mobile robots in various sectors such as industry, space, surgery, rescue and entertainment tasks are discussed. Based on the study, an improved design of a modular mobile robot for navigating through different terrains during a rescue operation is presented. Simulation study of the robot is included to demonstrate a motion capability of the modular mobile robot.

OS10-2 Implementation of Bug1 and Bug2 Path Planning Algorithms for TurtleBot Using ROS Noetic

Ilya Spektor[1], Aidar Zagirov[2], Ramil Safin[2], Evgeni Magid[1,2]
([1]HSE University, Russia) ([2]Kazan Federal University, Russia)

Mobile robots typically operate in a dynamically changing unknown environments. Bug family algorithms were proposed to deal with a path planning of a ground robot in a 2D configuration space of an unknown environment with local data about obstacles that could be collected using only a touch sensor. This paper presents an implementation of Bug1 and Bug2 local path planning algorithms. The implementation brings classical 2D algorithms into a 3D environment of the Gazebo simulation using the Noetic version of the Robot Operation System (ROS). A Turtlebot3 Burger model was used as a target robot and its performance was evaluated in simple convex and maze environments.
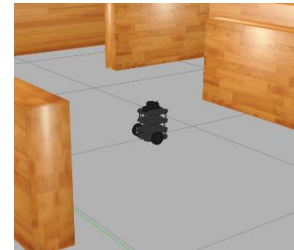
OS10-3 Implementation of Alg1 and Alg2 Path Planning Algorithms for Mobile Robots Using ROS Noetic

Anastasia Yankova[1], Timur Gamberov[2], Tatyana Tsoy[2]
([1]HSE University, Russia) ([2]Kazan Federal University, Russia)

Two standard approaches for a robot path planning include a global and a local navigation. The later does not require to store an environment model in a robot memory. This paper presents implementations of two local navigation algorithms, Alg1 and Alg2, with a robot having no prior information about an environment and obstacles. It calculates a path in a real time, continuously changing its states depending on correspondent conditions. The algorithms were implemented for an existing differential drive robot Turtlebot3 Burger using Robot Operation System (ROS). Virtual experiments were performed in the Gazebo simulator employing a simple 3D environment with only convex obstacles and a small 3D maze.

OS10-4 Implementation of VisBug-21 and VisBug-22 Path Planning Algorithms Using ROS Noetic

Viktoriia Mirzoian[1], Maxim Mustafin[2], Evgeni Magid[1,2]
([1]HSE University, Russia) ([2]Kazan Federal University, Russia)

Local navigation algorithms are crucial for autonomous robots operating in unknown environments where a presence of obstacles and dynamic changes pose significant challenges. A focus of these algorithms is to enable a real-time path calculation, allowing a robot to adapt its states dynamically based on corresponding environmental conditions, despite an absence of prior knowledge about surroundings. This paper presents an implementation of the VisBug-21 and VisBug-22 algorithms, designed to address challenges of a local navigation. The algorithms were implemented for a differential drive robot Turtlebot3 Burger using Robot Operation System (ROS). Virtual experiments were performed in the Gazebo simulator employing a simple 3D environment that contained only convex obstacles and a small 3D maze.

OS10-5 DistBug path planning algorithm package for ROS Noetic

Alexander Pak[1], Alexander Eremin[2], Tatyana Tsoy[2]
([1]HSE University, Russia) ([2]Kazan Federal University, Russia)

Algorithms of path-planning in an unknown environment play an important role in robotics. They do not require a prior information about obstacles' locations around a robot and allow calculating a path in a real time. This article presents an implementation of a sensory-based DistBug algorithm, which operates reactively using range data for immediate decision-making without constructing a world model. The algorithm was programmed in Python using robot operating system (ROS) and validated in the Gazebo simulator. For virtual experiments Turtlebot 3 Burger mobile robot was employed. The experiments were conducted in two types of environment: an environment with convex obstacles and a maze. The paper demonstrates analysis of experiments using several standard criteria of a path quality estimation.

OS10-6 On sensor modeling in Gazebo simulator

Niez Yuldashev, Alexandra Dobrokvashina, Roman Lavrenov (Kazan Federal University, Russia)

Sensor modeling in the Gazebo simulator is fundamental to robotics advancement. This review explores sensor modeling intricacies, methodologies, and applications, while emphasizing a critical role of a precise sensor modeling. Application scenarios demonstrate a sensor modeling's broad utility across fields including medical diagnostics, autonomous navigation, and industrial automation. Differences in research focus, methodology, and implementation underline a varied nature of sensor modeling studies. Key challenges include a need for more detailed world models. The paper guides research in sensor modeling and identifies crucial questions.
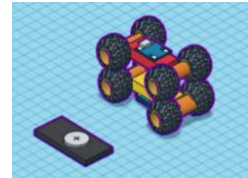
OS11 Intelligent Life and Robotics (6)
Chair Evgeni Magid (Kazan Federal University, Russia)
Co-Chair Kuo-Hsien Hsia (National Yunlin University of Science and Technology, Taiwan)

OS11-1 An Overview of Kinect Based Gesture Recognition Methods

Alexander Alexeev[1], Tatyana Tsoy[1], Edgar A. Martínez-García[2], Evgeni Magid[1,3]
([1]Kazan Federal University, Russia) ([2]The Autonomous University of Ciudad Juarez, Mexico)
([3]HSE University, Russia)

Visual sensors play an important role in a broad variety of robotic systems applications. Even though Kinect technology appeared over 10 years ago, Kinect sensors are still actively employed by researchers around the world. This paper presents an overview of Kinect and Kinect 2 sensors' applications in a human gesture based control. We analyzed existing research papers to estimate a popularity of particular feature extraction and gesture recognition methods, recommendations on a distance between an object of interest and a sensor, reported accuracy and latency of the sensor. Our analysis is supposed to facilitate a selection of a suitable combination of methods for a particular application of Kinect sensor in gesture recognition while considering its performance.

OS11-2 An investigation on the impact of human-robot interactions during an autonomous obstacle avoidance task

Riham Salman, Shifa Sulaiman, Renata Islamova, Tatyana Tsoy (Kazan Federal University, Russia)

The purpose of this research is to investigate how an interaction between humans and robots influences a safety of an autonomous obstacle avoidance task. The research collected and analyzed data from surveys and interviews using a combination of quantitative and qualitative methodologies. The findings contributed to our understanding of a complex interplay between a human-robot interaction, a perception, and a robot navigation safety. Based on these findings, the study proposed a number of recommendations for improving both physical and psychological safety aspects of an autonomous robot navigation.

OS11-3 A Comparative Analysis of Object Detection Methods for Robotic Grasping

Nikita Kolin, Elvira Chebotareva (Kazan Federal University, Russia)

The objects grasping is one of the fundamental robotic problems. Accurate and efficient real-time object detection is crucial for successful grasping in robots equipped with monocular vision. Deep machine learning has made significant progress in solving problems of object detection and image segmentation. At the same time, classical computer vision methods do not lose their relevance and can also be used for these tasks. In this research, we conduct a comparative analysis of the effectiveness the YOLOv8-seg neural network model versions for solving the image segmentation problem with classical segmentation methods. The obtained results allowed us to formulate some recommendations on the choice of a particular method for object detection depending on the surrounding environment conditions.

OS11-4 Vision-based autonomous navigation for medical examination using a UR3e manipulator

Bulat Abbyasov[1], Aidar Zagirov[1], Timur Gamberov[1], Hongbing Li[2], Evgeni Magid[1]
([1]Kazan Federal University, Russia) ([2]Shanghai Jiao Tong University, China)

Medical robotics is an emerging field of robotics within a healthcare sector. This interdisciplinary field focuses on prototyping, building and developing advanced robots for numerous clinical application. Modern robotic technologies have a tremendous potential for performing medical examination procedures such as ultrasonography using robotic manipulators that act in an autonomous manner. A manipulator navigation plays a key role in a safe and efficient exploration of a human body. This paper presents a development of a vision-based autonomous navigation system for a 6-DOF UR3e robotic arm. The developed system is based on a 3D point cloud and uses MoveIt for path planning. Gazebo was used as a simulation framework to validate the navigation system.

OS11-5 Robot-Assisted Language Learning: Scientific Data Analysis

Karina Sadyikova, Valeriya Zhukova, Roman Lavrenov (Kazan Federal University, Russia)

Robot-assisted language learning (RALL) is a direction in education that uses robots in a foreign languages studying process. In this paper, we present results of our research work devoted to studying trends in a quantity of publications in the field of using social robots for learning foreign languages. Additionally, we analyzed some published paper in the field of RALL, which allowed us to identify several popular robot models used in practical research in recent years. Among these models were the NAO, Pepper and DARWIN OP-2 robots. We have found that these robots are actively used in an educational process for learning foreign languages. In particular, these models are used to improve language and communication skills during implicit learning and interactive games.
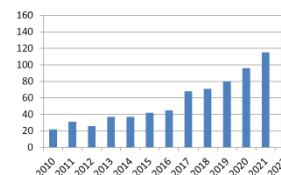
OS11-6 Monitoring Beehive Sound Levels with Arduino-based System

Kulmukhametov Ramis[1], Ramil Safin[1], Tatyana Tsoy[1], Kuo-Hsien Hsia[2], Evgeni Magid[1]
([1]Kazan Federal University, Russia) ([2]National Yunlin University of Science and Technology, Taiwan)

An automated beekeeping is a promising approach to addressing various issues associated with a beekeeping. Among primary problems, a swarming procedure stands out as a major concern. An uncontrolled swarming can lead to significant financial losses. During a swarming period a potential for losing a bee swarm is high, therefore a noise monitoring at this period gains a significant importance. Our long term research aims at a development of an intelligent monitoring system for beehive conditions, based on a hive-generated noises analysis. This paper presents an experiment that collected data about acoustic characteristics of bees' state using an Arduino microcontroller and a MAX9814 microphone module. The obtained data analysis is discussed.

OS12 Machine Learning and its Applications (4)
Chair Masato Nagayoshi (Niigata College of Nursing, Japan)
Co-Chair Takashi Kuremoto (Nippon Institute of Technology, Japan)

OS12-1 Restoration of Guqin Music by Deep Learning Methods

Takashi Kuremoto[1], Kazuma Fujino[1], Hirokazu Takahashi[1], Shun Kuremoto[2, 3], Mamiko Koshiba[*2],
Hiroo Hieda[3], Shingo Mabu[2]
([1]Nippon Institute of Technology, Japan) ([2]Yamaguchi University, Japan)
([3]Institute for Future Engineering, Japan)

Guqin music played an important role in the history of Asia cultures. The notation of Guqin are remained more than 600, however, only 100 music are played in nowadays. The handwriting Guqin notation named Jianzi Pu is hard to be understood, however, we challenge to restore the Guqin music by deep learning methods and few Jianzi Pu images. VGG16 and YOLOv5 were adopted in the recognition experiment for Guqin music restoration. As a result, YOLOv5 realized an online output of Guqin music as its output of audio or video forms.
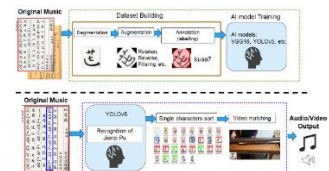
OS12-2 Constructive Nurse Scheduling Using Reinforcement Learning Considering Variations in Nurse Work Patterns

Masato Nagayoshi (Niigata College of Nursing, Japan), Hisashi Tamaki (Kobe University, Japan)

It is very difficult to create a work schedule that satisfies all the different requirements in nurse scheduling. For this reason, numerous studies have been conducted on the nurse scheduling problem. However, the created shift schedule is often not practical as it is, because adjustments including various constraints and evaluation criteria are required. Therefore, we have proposed a work revision method using reinforcement learning in a constructive nurse scheduling system. In this paper, we investigate the feasibility of creating a practical work schedule that considers different evaluations of nurses' work patterns, i.e., nurses' life stages.
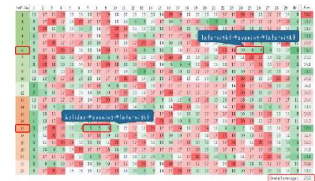
OS12-3 A Basic Study on Indicator of Transfer Learning for Reinforcement Learning

Satoshi Sugikawa, Kenta Takeoka, Naoki Kotani (Osaka Institute of Technology, Japan)

Reinforcement learning requires a lot of learning time for the agent to learn. Transition learning is a method to reduce this learning time, but it has the problem that the user does not know which knowledge is effective in which environment until it is learned. Therefore, it is necessary for the user to consider the relationship between the source and destination when transferring knowledge. Therefore, this study proposes Indicator of adaptation criteria that can determine this relationship in advance. In simulations, we demonstrate the usefulness of the proposed method by using some example problems.

OS12-4 Machine Learning Approach to Predict Cooling Load for Existing Buildings

Makoto Ohara[1], Hideo Isozaki[2]
([1]International Professional University of Technology in Osaka, Japan) ([2]Kobe University, Japan)

The objective of this study is to predict air conditioning loads for existing buildings using operational data, weather forecasts and visitor forecasts. The proposed prediction method is based on a neural network approach. However, it is important to note that the proposed method does not learn the entire loads. Loads are divided into factors which can be predicted by traditional thermodynamics and factors which are subject to machine learning. The proposed method has been applied to an example instance using operational data from an underground mall in Kobe, and its validity has been confirmed. This result could potentially lead to more efficient use of energy in buildings.
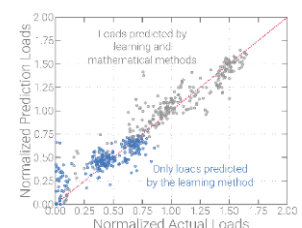
OS13 Robot Control (10)
Chair Yizhun Peng (Tianjin University of Science and Technology, China)

OS13-1 Intelligent Logistics Handling Robot: Design, Control, and Recognition
Yanchao Bi, Jiale Cheng, Limei Wang, Yizhun Peng (Tianjin University of Science and Technology, China)

This study aims to investigate various aspects of intelligent logistics handling robots, including mechanical design, automatic control, and image recognition. With the continuous development of the logistics industry and advancements in automation technology, intelligent logistics handling robots play a crucial role in improving logistics efficiency and reducing costs. Leveraging existing technologies, we have designed and developed an omnidirectional mobile intelligent logistics handling robot equipped with a SCARA-type robotic arm. The robot integrates functions such as task acquisition, global positioning, material detection, warehouse identification, material handling, and stacking, achieving a fully automated and streamlined logistics handling process.

OS13-2 Greenhouse Design Using Visual Recognition and IoT Technology
Yuntian Xia, Yizhun Peng (Tianjin University of Science and Technology, China)

This device is to solve the traditional pesticide spraying method on the human body has a greater impact and other issues, through the STM32-based visualisation of the intelligent greenhouse to achieve automatic spraying of pesticides and remote monitoring and other functions, the establishment of a visualisation of the intelligent greenhouse monitoring platform. This equipment through the MQTT protocol, not only through the Internet of Things platform real-time monitoring of crop growth status in the greenhouse, but also through the platform to determine whether to spray pesticides, data transmission, so as to use the cross slide to control the position of the nozzle, and then through the visual recognition algorithms to improve the accuracy of the visual recognition part of the accuracy of the spraying of plants affected by insect pests, the realization of the digital intelligent greenhouse.

OS13-3 Design of Modular Photovoltaic Environmentally Friendly Portable Stroller
Suqing Duan, Yizhun Peng (Tianjin University of Science and Technology, China)

The work is centered on ROS and integrates speech recognition and natural language processing modules to enhance the environmental awareness of the stroller and monitor the baby's condition to ensure a timely and appropriate response. The Raspberry Pi is the main control unit of the stroller and connects to a cloud-based IoT platform via the MQTT protocol. The platform facilitates seamless communication between the Raspberry Pi and cloud data for efficient data visualization on mobile devices. This innovative solution solves the challenge of balancing parental responsibilities with career advancement, promoting healthier and happier babies while allowing parents to maintain a harmonious work-life balance.

OS13-4 "Teenage Mutant Ninja Turtles" - Design of a Bionic Quadrupedal Rescue Robot
Hongpi Zhao, Yingfan Zhu, Zhihan Zhao, Xin Liang, Lei Lv, Yizhun Peng
(Tianjin University of Science and Technology, China)

"Teenage Mutant Ninja Turtles - Bionic Quadrupedal Rescue Robot is a quadrupedal robot based on the principle of bionics, inspired by the quadrupedal animals in nature. The robot has rescue and life detection capabilities and can perform rescue operations at disaster sites. The design of the robot enables it to operate efficiently and stably in complex environments, and at the same time it has the qualities of bravery and toughness, which are in line with the image of the Teenage Mutant Ninja Turtles. The robot can be used in a wide range of application scenarios, such as earthquake, fire and other disasters, to provide more efficient and safer support for rescue work.

OS13-5 Design of Grass Lattice Planter for Complex Environment Based on Adaptive Suspension Technology

Shaokai Tian, Wenqi Fu, Yizhun Peng (Tianjin University of Science and Technology, China)

The Complex environmental grass grid growers based on adaptive suspension adopted the high-performance stm32 as the master chip, and use the Bluetooth communication module, so that users can easily control the vehicle driving, feeding and the posture of the pressing wheel through the mobile phone application. In addition, we have introduced new mechanical structures such as adaptive suspension to ensure effective planting of grass squares in complex terrain such as sloping land. Our vehicle uses a Mecanum wheel motion system to squeeze seeds by rotating the friction wheels at high speed.

OS13-6 Design of Intelligent Ecological Multifunctional Plant and Animal Breeding System

Suqing Duan, Yuntian Xia, Siyi Wang, Yizhun Peng (Tianjin University of Science and Technology, China)

The intelligent ecological multi-functional system for animal and plant cultivation integrates PLC and MCU automatic control technology. This innovative approach amalgamates aquaculture with horticulture, fostering a symbiotic environment for the breeding of animals and plants. The system's operations are categorized into four key components: mechanical structure, environmental perception, automatic control, and intelligent networking. Its comprehensive functionality encompasses plant monitoring, self-checking of temperature and humidity, group control for fertilizer replenishment, consistent temperature regulation, full-spectrum illumination, advanced filtration, fogging supplementation, versatile water management, fertilizer blending, and water purification through sterilization.

OS13-7 Design of a Fully Automated Logistics Handling Platform

Hongpi Zhao, Jianfeng Qin, Yizhun Peng (Tianjin University of Science and Technology, China)

With the progress of science and technology and the development of society, the world's countries on industrial production efficiency, intelligent manufacturing transformation continues to grow. Therefore, we designed a fully automatic logistics handling platform, which is a multi-modular device based on mechanical design, microcontroller control, visual positioning, the device can realize autonomous identification, autonomous transportation, improve production efficiency, the device has a wide range of applications, can be applied to food packaging, parts processing, intelligent manufacturing and other automation scenarios.

OS13-8 Design and Application of AI-based Brush Calligraphy and Painting Robot

Haibo Li, Yizhe Sun, Shuxin Wang, Yizhun Peng (Tianjin University of Science and Technology, China)

With the rapid development of digital technology and artificial intelligence, the innovation and exploration of traditional painting forms in the field of digital art have become increasingly captivating. The main focus of this research is to create a brush calligraphy and painting robot system based on PLT files. It digitizes user creative instructions to autonomously generate brush calligraphy and landscape paintings. This technology not only advances the deep integration of digital art and traditional culture but also opens up new perspectives and vast possibilities for artistic creation.

OS13-9 Intelligent Wheelchair System: Non-contact Heart Rate and Body Temperature Measurement

Dongpo Ma, Junsheng Zhang, Yizhun Peng (Tianjin University of Science and Technology, China)

With the continuous development of society and human civilization, people, especially the disabled and the elderly, need more and more to use modern high technology to improve their lives and improve the comfort of life. Therefore, this paper introduces a single-chip microcomputer-based non-contact body temperature and facial recognition-based heart rate measurement method and applies it to the intelligent wheelchair system. Monitor the vital signs of the elderly, and give early warnings in time when abnormal values occur, so that family members or caregivers can rescue them in time.

OS13-10 Recognition and Localization Method for Automotive Axle Holes in Assembly Robots

Junsheng Zhang, Dongpo Ma, Yizhun Peng (Tianjin University of Science and Technology, China)

During the production of automotive axle holes, the roundness error at the pipe opening leads to low detection efficiency due to manual measurements, rendering real-time inspection unfeasible. This paper proposes a method for detecting and sorting the roundness of automotive axle holes based on visual inspection. Ensuring accuracy in grasping, it establishes the kinematic model of the robot Aubo_I5. Targeting automotive axle holes for grasping, it employs adaptive threshold segmentation to highlight the section features of the axle hole. The Canny algorithm is then used to extract edge information, and finally, the least squares method is utilized to detect roundness errors for sorting the axle holes based on this error.

OS14 Robotic Manipulation (3)
Chair Kensuke Harada (Osaka University, Japan)
Co-Chair Akira Nakamura (Saitama Institute of Technology, Japan)
Co-Chair Tokuo Tsuji (Kanazawa University, Japan)

OS14-1 Evaluation Standard of Error Recovery Planning Focused on Revival Process from Failures in Robotic Manufacturing Plants

Akira Nakamura[1], Kensuke Harada[2] ([1]Saitama Institute of Technology, Japan) ([2]Osaka University, Japan)

In recent years, many intelligent robots have been used in various fields. In many cases, these robots need to be able to perform not only repetitive tasks but also non-routine tasks. As work errors are more likely to occur in such situations, we have proposed several methods for error recovery. Our method uses both forward recovery and backward recovery to restore work. The former is a recovery method that is often used for minor modifications, and the latter is the recovery from relatively large failures that we mainly use. This paper proposes a method for selecting the best path from several possible recovery paths using a new evaluation method.
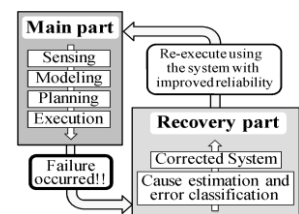
OS14-2 Robotic Food Handling Utilizing Temperature Dependent Variable-Stiffness Material

Rozilyn Marco[1], Prashant Kumar[2], Xinyi Zhang[2], Weiwei Wan[2], Kensuke Harada[2]
([1] University of Toronto, Canada) ([2] Osaka University, Japan)

This paper presents a robotic gripper that addresses challenges of automated robotic food handling. Depending on the food material, the robotic gripper should adjust its finger compliance. To cope with this problem, we apply the approach on using a variable stiffness fabric on the finger surfaces of the gripper, where the finger stiffness changes depending on its temperature. This gripper design was validated empirically through force and object-grasping experiments. By using the gripper, hard objects can be grasped with the hard mode of the finger while a fragile object can be safely grasped with the soft mode.

OS14-3 Vegetable maturity evaluation for harvest robots
Reno Muhammad Fadilla, Tokuo Tsuji, Tatsuhiro Hiramitsu, Hiroaki Seki (Kanazawa University, Japan)

We propose a method for detecting vegetables in an image and classifing their maturity by color. Vegetable images are classified according to their maturity and used as training data. We represent the images in different color systems and investigate the accuracy of the classification of maturity. The proposed method can be applied to harvesting robots.

OS15 Artificial Intelligence for Embedded Systems and Robotics (6)
Chair Hakaru Tamukoh (Kyushu Institute of Technology, Japan)
Co-Chair Yuma Yoshimoto (National Institute of Technology, Kitakyushu College, Japan)

OS15-1 YOLO real-time object detection on EV3-Robot using FPGA hardware Accelerator
Dinda Pramanta[1], Ninnart Fuengfusin[2], Arie Rachmad Syulistyo[2], Hakaru Tamukoh[2]
([1]Kyushu Institute of Information Sciences, Japan) ([2]Kyushu Institute of Technology, Japan)

The growing demand for robots necessitates faster and more precise processing. However, running large Artificial Intelligence (AI) models from cloud data centers to mobile robots via inference models uses considerable computation resources, which leads to power limitations, particularly for mobile robots. The use of reconfigurable semiconductor devices at the hardware level is a promising solution to this problem. We introduce the educational kit EV3-Robot with a co-design methodology utilizing Field-programmable Gate Arrays (FPGA) Kria KV260 as a hardware accelerator specifically for object detection. We apply the You Only Look Once (YOLO) model for object detection, which provides real-time results for practical applications. Additionally, we analyze the processing times of the local PC and EV3-Robot.

OS15-2 A Low Computational Cost Hand Waving Action Recognition System with Echo State Network for Home Service Robots
Hiromasa Yamaguchi, Akinobu Mizutani, Arie Rachmad Syulistyo, Yuichiro Tanaka, Hakaru Tamukoh
(Kyushu Institute of Technology, Japan)

This study proposes a low computational cost hand-waving action recognition system for non-verbal communication in home service robots. The system is based on an echo state network, which requires lower computational costs than that of deep neural networks (DNNs), and processes time-series data of skeletal coordinates of humans to recognize hand-waving actions. Additionally, this study proposes and compares two types of preprocessing methods of the skeletal coordinates to ensure the robustness of the human positions on the screen: one method extracts elbow and arm angles, which are invariable regardless of the humans' positions and the other normalizes the skeletal coordinates. The experimental result shows that the proposed system has competitive accuracy and is faster than DNN-based methods and robust to varying human positions.
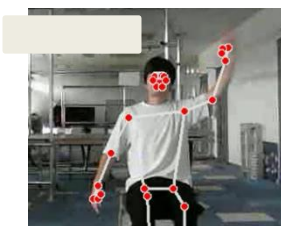
OS15-3 A Rapidly Adjustable Object Recognition System through Language Based Prompt Engineering
Naoki Yamaguchi, Tomoya Shiba, Kosei Isomoto, Hakaru Tamukoh (Kyushu Institute of Technology, Japan)

We propose the use of language-based prompt engineering to achieve rapidly adjustable object recognition in RoboCup@Home. The proposed prompt engineering involves humans adding features, such as the color and material of an object, into the text prompts inputted into Language Segment Anything. In this research, we evaluated the effectiveness of our proposed method in three benchmark tests for object recognition at RoboCup@Home held in France in 2023. The results showed that the highest scores were obtained in certain tasks, indicating that it could be applied to a range of recognition tasks.
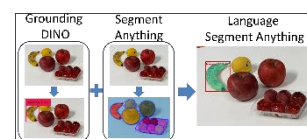
OS15-4 Development of a SayCan-based task planning system capable of handling abstract nouns

Kosei Yamao[1], Daiju Kanaoka, Kosei Isomoto, Akinobu Mizutani, Yuichiro Tanaka, Hakaru Tamukoh
(Kyushu Institute of Technology, Japan)

The task planning system is required to accomplish various requests from a human in real-world environments. SayCan, one of the task planning systems, has high accuracy. However, its accuracy decreases for requests that include abstract nouns of the ambiguous word/phrase, and the inference time increases as the number of skills increases. We propose a novel task planning system based on SayCan that introduces a function for listening back to abstract nouns and a rule-based skill extraction, enhancing accuracy and reducing inference time. The proposed system facilitates the interpretation of requests and enables appropriate task planning with low inference time. The effectiveness of the proposed system was demonstrated at RoboCup@Home, where it achieved high performance.

OS15-5 RoboCup@Home 2023: Stickler for the Rules Task Solutions

Tomoya Shiba, Hakaru Tamukoh (Kyushu Institute of Technology, Japan)

This paper proposes an approach where home service robots use only two recognition models that can be capable of prompt tuning to detect rule violations in the home. The robots are increasingly required to perform more advanced and complex perceptions beyond their traditional roles such as cleaning. Advanced tasks such as serving and security, for example, usually require multiple AI systems, including person detection, object detection and more. Our approach simplifies this by leveraging the combined capabilities of Grounding DINO and SAM to detect rule violations effectively. The success of our method was proven at RoboCup@Home 2023, where it secured the highest score among all participating teams.
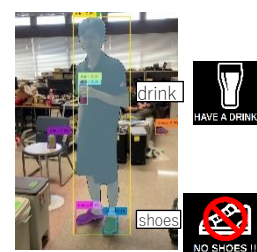
OS15-6 Offloading Intellectual Processing from Home Service Robots to Edge Devices

Yuma Yoshimoto, Mizuki Kawashima, Shun Yonehara
(National Institute of Technology, Kitakyushu College, Japan)

In this study, we focus on extending the operational time of home service robots by offloading intellectual processing to circuit devices such as Field Programmable Gate Arrays (FPGAs), which in turn reduces power consumption. The core of our approach involves developing a method for implementing intellectual processing on FPGAs, coupled with a dynamic circuit reconfiguration technique. This enables the FPGA to adaptively respond to frequent task changes. We present: (a) methods for transitioning circuits from a robot's control computer to FPGA in response to varying tasks, and (b) an evaluation of the effectiveness of using FPGA to extend operational time under these rapidly changing task conditions.

OS16 Industrial Artificial Intelligence Robotics (4)
Chair Eiji Hayashi (Kyushu Institute of Technology, Japan)

OS16-1 A Research on Performance Information Editing Support System for Automatic Piano

Yoshiki Hori, Eiji Hayashi (Kyushu Institute of Technology, Japan)

In 1996, Hayashi et al. from our laboratory developed an automatic piano playing device that requires data with nuances for each note to perform in a human-like piano playing. However, this device lacks the function to infer such nuanced data. Therefore, prior research focused on developing a system to infer data with nuances for each note. Yet, this system required manual inference, consuming a significant amount of time. In this study, we have constructed a system capable of automatically performing inference using deep learning. This system not only improves the accuracy of piano playing inference but also contributes to the efficiency of the inference process.

OS16-2 Development of Autonomous Mobile Field Robots – Accuracy Verification of Self-Localization through Simulation -

Takamasa Hayashi[1], Shintaro Ogawa[1], Yuto Okawachi[1], Tan Chi Jie[1], Janthori Titan[1], Ayumu Tominaga[2], Eiji Hayashi[1], Satoko Seino[3]
([1]Kyushu Institute of Technology, Japan) ([2]National Institute of Technology (Kitakyushu College), Japan), ([3]Kyushu University, Japan)

In recent years, the increase in marine debris has become a significant challenge in terms of its collection. Costal debris, a type of marine debris, can be collected by human hands, but the variety in shapes, types, and sizes presents limitations to human-only collection efforts. To address this, I focused on developing an autonomous mobile robot, establishing a simulation environment was considered crucial for facilitating smooth progress. This paper focuses on self-localization, an essential aspect for autonomous movement. We replicated an actual coastal cleaning site within the simulation environment and evaluated the accuracy of self-localization using an EKF (Extended Kalman Filter) with multiple sensors.

OS16-3 Development of AR System for Grasping String Foods on Introduction of Industrial Robot

Yoshihiro Koyama, Eiji Hayashi (Kyushu Institute of Technology, Japan), Akira Kawaguchi (The City College of New York of The City University of New York, United States of America)

In recent years, the food service industry has been facing a labor shortage. However, the introduction of industrial robots is not easy due to the high cost of equipment and system integration. Therefore, we are developing an Augmented Reality (AR) application for the purpose of introducing robots to small and medium-sized companies. By using this application to perform tasks necessary for introducing robots, such as teaching, cost reductions can be expected when introducing industrial robots. In a previous study, an AR-based grasping and serving simulation system was developed for solidified foods such as fried chicken and rice balls. In this study, we focused on string-shaped food items such as spaghetti, and attempted to develop an AR system for grasping and serving a string-shaped object by controlling a gripper.

OS16-4 An Image Analysis of Coastal Debris Detection -Detection of microplastics using deep learning

Yuto Okawachi[1], Ayumu Tominaga[2], Shintaro Ogawa[1], Takamasa Hayashi[1], Tan Chi Jie[1], Janthori Titan[1], Eiji Hayashi[1], Satoko Seino[3]
([1]Kyushu Institute of Technology, Japan) ([2]National Institute of Technology (Kitakyushu College), Japan) ([3]Kyushu University, Japan)

To address the issue of litter drifting ashore, this study developed a deep learning-based microplastic detection system. The system employed yolov7 as its deep learning network, complemented by SAHI (Slicing Aided Hyper Inference) as an additional vision library. yolov7 is renowned for its efficacy in real-time object detection. Our experimental framework involved four tests, utilizing two variations of yolov7 - the standard model and yolov7-e6e - in conjunction with SAHI. The effectiveness of each test was quantified using metrics such as Intersection over Union (IoU), Precision, Recall, F-measure, and Detection Time in seconds. For our dataset, we gathered images from actual cleanup locations, such as Hokuto Mizukumi Park. The model's discriminator underwent 700 training iterations, with a learning rate set at 0.001.

OS17 Electronics and Kansei Engineering Based on ETT theory (5)
Chair: Tetsuo Hattori (Kagawa University, Hiroshima Institute of Technology, Japan)
Co-Chair: Yusuke Kawakami (NIT (Kagawa College), Japan)

OS17-1 A High-Speed Estimation Method of Parameters in Impulse Response
Toshiki Tanaka, Ivan Tanev (Doshisha University, Japan), Tetsuo Hattori (Kagawa University, Japan)

This paper proposes a high-speed parameters estimation method for compartment model where output function is described by the convolution between input function and impulse response, which is like a time-invariant linear system. The proposed method uses linear regression analysis based on equations that can be obtained from the differentiation of convolution (DOC). In this paper, taking the parameters estimation problem of PET (Positron Emission Tomography) inspection system and RLC series electrical circuit for examples, we show that the method can estimate parameters in those impulse responses.
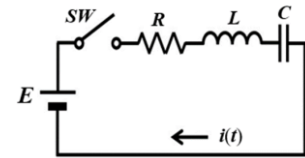
OS17-2 A Consideration on Amplification Function in BJT Evers-Moll Model and PTT (I)
---- V-I Characteristics ----
Shimon Hattori, Osamu Matoba (Kobe University, Japan), Tetsuo Hattori (Kagawa University, Japan), Toshiki Tanaka (Kinkei System, Japan), Yusuke Kawakami (NIT (Kagawa College), Japan)

In 1989, the notion of PTT (Photon Transport Transistor) has been proposed by B.J. Van Zeghbroeck et al., at IBM Research Center at that time. PTT is an optical coupling device of light emitting diode (LED) or laser diode and light receiving diode (Photo Diode, PD) where the carrier of the base layer is light (Photon) only. Later in 1996, W. N. Cheung and Paul J. Edwards have shown that PTT can be a very low noise amplifier in a positive feedback circuit, based on theoretical calculations of the noise figure. In this paper we consider the amplification principle showing the VI characteristics of the PTT and bipolar junction transistor (BJT).
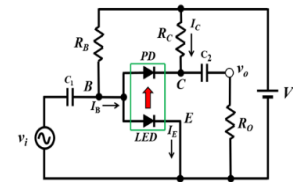
OS17-3 A Consideration on Amplification Function in BJT Evers-Moll Model and PTT (II)
---- H Parameters in the Small Signal Amplifier Circuit----
Shimon Hattori, Osamu Matoba (Kobe University, Japan), Tetsuo Hattori (Kagawa University, Japan), Toshiki Tanaka (Kinkei System, Japan), Yusuke Kawakami (NIT (Kagawa College), Japan)

The notion of PTT (Photon Transport Transistor) has been proposed in 1989 as an optical coupling device of light emitting diode (LED) or laser diode and light receiving diode (Photo Diode, PD), where the carrier of the base layer is light (Photon) only. In this paper we consider the amplification principle of the PTT and BJT (bipolar junction transistor), by illustrating those h parameters of equivalent small signal amplifier circuit. Moreover, this paper also discusses the similarities and differences between PTT and BJT, based on the derived h parameters.
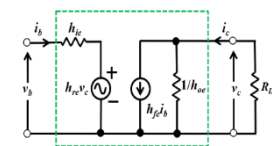
OS17-4 Color Image Arrangement Using Histogram Matching
Yusuke Kawakami (NIT (Kagawa College), Japan), Tetsuo Hattori (Kagawa University, Japan), R.P.C. Janaka Rajapakse (Tainan National University of the Arts, Taiwan (R.O.C.))

For the arrangement processing of image using its histogram, we previously have presented a Histogram Matching method based on Gaussian Distribution (HMGD). However, in the case where the brightness histogram of input image has multiple peaks, the HMGD processing does not always bring good results. In this paper, we present an improved histogram matching method using the reference histogram that is made by appropriate moving average processing over the histogram of input image. Also in this paper, we show the experimental results.
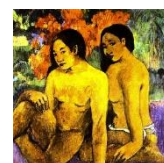
OS17-5 Methodology for Creativity Oriented STEM Education Based on ETT Theory

Tetsuo Hattori, Toshihiro Hayashi, Mai Hattori, Yoshiro Imai (Kagawa University, Japan),
Asako Ohno (Osaka Sangyo University, Japan), Takeshi Tanaka (Hiroshima Institute of Technology, Japan)

Recently, the necessity of integrated and comprehensive methodology for STEM (Science, Technology, Engineering and Mathematics) education is growing. In this paper, we propose an educational methodology utilizing the viewpoint of Equivalent Transformation Thinking (ETT) theory which has been proposed by Dr. Kikuya Ichikawa in 1955 as a principle of creativity. Especially, we show that the viewpoint is very useful not only for new technology invention but also for STEM Education in the sense that it deepens insights of the contents to be learned, motivates students to study further, and inspires their creativity.
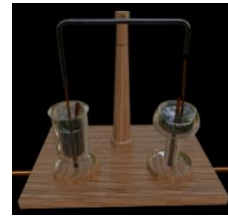
OS18 Computer and Information Engineering (12)
Chair Norrima Mokhtar (University of Malaya, Malaysia)
Co-Chair Heshalini Rajagopal (UCSI University, Malaysia)

OS18-1 Efficient Campus Shuttle Tracking and Management Mobile Application for College Campus

Andrea Tantay Gonzales[1], Kavitha Thamadharan[1], Neesha Jothi[2]
([1] INTI International College Penang, Malaysia) ([2]UCSI University, Malaysia)

Shuttle Stalk, a Real-Time Campus Shuttle Booking and Tracking System was a proposed solution which enables the students to register for the shuttle service through the mobile application, reducing the administrative burden and to track its current location in real-time, allowing them to plan their journeys more effectively. The application developed serves as a comprehensive solution for both students and administrators, providing real-time shuttle tracking and streamlined management capabilities.

OS18-2 GCN Analysis of Task-Based fMRI Data for Diagnosis of Schizophrenia

Tejaswini Thota[1], Reuben Stephen John[1], Dr R Dhanush[1], Dr Amutha S[1]
([1]Vellore Institute of Technology, India)

This study focuses on schizophrenia, characterized by distorted reality and delusions. We propose utilizing a Graph Convolutional Network (GCN) model on a task-based fMRI data to differentiate schizophrenia instances, with and without auditory hallucinations, utilizing healthy participants for comparison. Various node embedding algorithms are compared to extract structural properties, and statistical features for differentiation. The study focuses on identifying specific auditory stimuli that significantly differentiate individuals. Model validation deploys k-fold stratified crossvalidation, with evaluation metrics, including accuracy, precision, recall, F1 macro, and MCC guide the assessment of the GCN model.
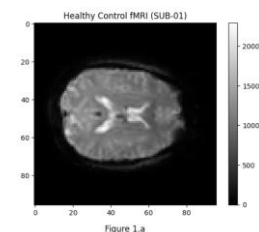
OS18-3 AR-Based Application for Campus Navigation

Renuka Devi Rajagopal[1], Akshay S[1], Manoj Rathinam[1], Shakthi B[1], Heshalini Rajagopal[2]
([1]Vellore Institute of Technology, India) ([2]UCSI University, Malaysia)

The "Campus Navigator" is an innovative augmented reality (AR) application designed for seamless campus navigation. It combines geolocation technology, QR code recognition, and predefined geospatial anchors to transition users from outdoor to indoor spaces. Augmented reality overlays provide real-time wayfinding guidance, eliminating the need for traditional maps. The user-centric design and user feedback integration enhance accessibility and user satisfaction. Real-world testing demonstrates promising performance, accuracy, and user engagement. The Campus Navigator transforms campus navigation, with potential for broader applications.

OS18-4 Development of Robotic Assistant for Health Care Sector with A Special Focus to Aid the Geriartric Patients

Narayanan Ganesh (Vellore Institute of Technology, India)

Human manual efforts in their day-to-day tasks are reduced by robotic helpers. This paper analyzes the robot created to assist the aged people. The created robot can be controlled by voice commands with its own inbuilt microphone to pick up human speech orders. This robot can do a variety of movements, turns, grab operations, move an object from one location to another. Personal assistant robot is built using Microcontrollers. The results reveal that the developed robot will be one of the best companions for the geriatric patients and will be an alternate solution will make a new mark in the health care sector.

OS18-5 Development of a Desktop Application Restaurant Management System

Gabriela Maria Ancilla, Heshalini Rajagopa, Ismail Ahmed Al-Qasem Al-Hadi (UCSI University, Malaysia)

The number of restaurants has grown rapidly around the world, the awareness of managing it efficiently has increased. This necessity yields the idea to invent a restaurant management system (RMS). This RMS will bring several benefits such as greater management of a restaurant, reduced resources cost, raise profit, and time-saving by allowing the admin to manage the restaurant easily through the functionality provided in the system. The provided features include adding, updating, and deleting information, stock calculation and auto re-order stock items, ingredients management, and finance calculation.

OS18-6 Face Recognition based on Attendance System

Koh Pei Cong [1], Heshalini Rajagopal[1], Ghassan Saleh[1], Norrima Mokthar[2]
([1]UCSI University, Malaysia) ([2]University of Malaya, Malaysia)

The use of face recognition technology for attendance tracking has grown popular in recent years. Hence, the main goal of this project is to produce an accurate, fast, and robust face recognition based on an attendance system. The system detects the user's unique features, understand the identity of the user through face recognition technology, and thus records the attendance from the face recognition dataset to the user that matches the user, in an attempt to help the user automatically check in with real-time attendance date. The system includes features such as face detection, face recognition, distance estimation, and attendance recording.

OS18-7 U-Reserve: Development of a Facility Reservation System for UCSI University

Esther Chong Jun Lynn[1], Neesha Jothi[1], Ismail Ahmed[1] ([1]UCSI University, Malaysia)

In today's technology-driven era, there is a growing emphasis on efficient and convenient methods to enhance user satisfaction. UCSI University, Kuala Lumpur, currently employs a manual facility reservation system that requires in-person or email submissions. This outdated approach often results in double booking and underutilization of resources. Therefore, this project aims to develop a web-based facility reservation system named U-Reserve. This system is developed by HTML, CSS, JS, JSP, Servlet, and MySQL, focusing on colour tone and font type. These considerations are vital in providing an improved user experience. With U-Reserve, not only the reservation requests and cancellations being streamlined, but informed decisions regarding facility management can also be made by a data-driven dashboard.

OS18-8 The Smart Document Processing with Artificial Intelligence

Raenu Kolandaisamy[1], Heshalini Rajagopal[1], Indraah K[2], Glaret Shirley Sinnappan[3]
([1]UCSI University, Malaysia) ([2]University Utara Malaysia, Malaysia)
([3]Tunku Abdul Rahman University, Malaysia)

This study focuses on the challenges and potential for Intelligent Document Processing (IDP) with Artificial Intelligence (AI) to manage unstructured data. A large amount of data in many different forms, such as information from the IoT, cybersecurity and more are produced during this modern Digital Age which has been widely distributed through a wide range of unformatted formats. IDP utilizes AI technologies like Machine Learning (ML), Natural Language Processing (NLP), and Computer Vision (CV), with the aim of converting unstructured data into structured, usable information. The IDP, are explored in the findings and discussion section that emphasizes its capability to automatically perform redundant tasks, reduce operating costs as well as improve employee productivity.

OS18-9 Digital Security Challenges Faced by Business Organizations

Raenu Kolandaisamy[1], Heshalini Rajagopal[1], Indraah K[2], Glaret Shirley Sinnappan[3]
([1]UCSI University, Malaysia) ([2]University Utara Malaysia, Malaysia)
([3]Tunku Abdul Rahman University, Malaysia)

Over the decades, the forms of cyber-attacks have evolved from disruption level, cybercriminal, followed by cyber espionage and lastly threatening level. Digital security has played a significant role in protecting enterprises from any form of cyber-attacks, especially in today's era of digitalization. Aligning with the global effort in emerging the concept of Industrial Revolution 4.0 (IR4.0) in organizations, where sensitive data and confidential information can be accessed at a fingertip of an employee. This paper discusses the difficulties of implementing digital security solutions in an enterprise in terms of external potential cyber threats, internal cyber security roadblocks within the organization and how Covid-19 pandemic imposed challenges towards cyber security in the organizations.

OS18-10 The Study on Perception on E-Waste Among the People

Raenu Kolandaisamy[1], Heshalini Rajagopal[1], Indraah K[2]
([1]UCSI University, Malaysia) ([2]Utara Malaysia, Malaysia)

E-waste is piling up at an alarming rate, which results in our environment's pollution, while only a few are being disposed of correctly. This research aims to discover people's awareness of e-waste and how to handle it. An online survey was done among 75 respondents, who were asked about their general knowledge of e-waste. Most respondents are aware of the effects of e-waste but lack knowledge on how to tackle them, including questions concerning e-waste management and its causes. In conclusion, the general public must be educated on e-waste to reduce and dispose of e-waste efficiently and effectively.

OS18-11 Emergence of Cybercrimes in Online Social Networks

Raenu Kolandaisamy[1], Heshalini Rajagopal[1], Indraah K[2]
([1]UCSI University, Malaysia) ([2]University Utara Malaysia, Malaysia)

The rise of social networking websites has been seen in recent years. Everyone will be spending most of their time on social networking websites such as Facebook, Instagram and Whatsapp. The great advantage that this social networking website offers benefited many of the users. It can help people to promote themselves or their business to gain more popularity and also customers through these social networking websites. There are many cybercrimes that can be identified which are identity theft, hacking, fraud and so on. The emergence of cybercrimes has created an awareness so that the users will know what the common attacks are and how they can be prevented from being lure and being a victim of this attack. This research will discuss about the attacks and how these attacks can be prevented by the users.

OS18-12 Development of a music recommendation application by using facial emotion recognition

Shengke Xie, Raenu Kolandaisamy, Ghassan Saleh, Heshalini Rajagopal (UCSI University, Malaysia)

Music is an important part of human life and culture, and it can affect people's emotions and moods. However, choosing music from a large library can be a challenging and time-consuming task. In this paper, we propose a facial expression recognition-based music recommendation system that can recommend suitable music which matches the user's current mood. The system uses a camera to capture the user's face and a convolutional neural network model which trained facial emotion recognition database to recognize seven basic emotions: anger, disgust, fear, happiness, sadness, surprise and neutral. The paper contributes to the research of facial emotion recognition and music recommendation and provides a convenient way for people to enjoy music.

OS19 Natural Computing (3)
Chair Marion Oswald (Vienna University of Technology, Austria)
Co-Chair Yasuhiro Suzuki (Nagoya University, Japan)

OS19 -1 A Model of Reaction-diffusion phenomena with Multiset Processing
Yasuhiro Suzuki (Nagoya University Japan)

We propose a model of reaction-diffusion phenomena using Abstract Rewriting System on Multisets ARMS, which is a model of Multiset Processing. Although proposed model is simple, computer simulations confirm that the Turing pattern is generated.

OS19-2 Extract tactile qualities from time series data
Yasuhiro Suzuki (Nagoya University, Japan)

We proposed the Tactile Score to describe time-varying tactile sensation by the time variation of vertical force. Tactile quality is essential in hardness/softness, roughness, and temperature. Hardness and softness can be extracted from the shape of the Tactile Score. Roughness can be extracted from the pattern of hardness and softness. The arbitrary time series data can be interpreted as a tactile score by considering the time variation of the vertical force, and the hardness and softness are extracted from the time series data interpreted as the tactile score. This method can extract different features from conventional data science methods.

OS19-3 Healthcare applications of vibrotactile stimulation developed by Tactile Score

Yasuhiro Suzuki, Rie Taniguchi (Nagoya University, Japan)

We constructed a system to convert the tactile score into vibratory tactile sensation. We conduct basic biological experiments on electronic tactile stimuli and apply the results to medical and healthcare applications. We use two types of electronic tactile stimuli. One is direct vibrating tactile stimulation, and the other is airborne vibrating tactile stimulation. Verification experiments of this system have confirmed cosmetic effects such as skin collagen aggregation, improvement of hypertension and diabetes, and reduction of pain and nerve paralysis. In addition, a clinical study was conducted in which patients with dementia were exposed to vibrotactile sensation (low-frequency sound) via air vibration and an improvement in cognitive function was confirmed.

OS20 Advances in Field Robotics and Their Applications (15)
Chair Kazuo Ishii (Kyushu Institute of Technology)
Co-Chair Keisuke Watanabe (Tokai University)

OS20-1 Image Collection Experiments of a Handy AUV for Offshore Structure Inspection

Keisuke Watanabe, Koki Amano, Shingen Urano, Yasutaka Taniguchi, Konosuke Watanabe
(Tokai University, Japan)

In recent years, in Japan, with the decline in the working population, there has been a noticeable shortage of labor at construction sites and offshore operations. On the other hand, with the aim of realizing a GX society, it is expected that offshore wind power generation platforms will rapidly increase and be deployed offshore. Since it is not easy to access offshore platforms, it is necessary to automate equipment inspection work, and it is expected that the use of AUVs will be particularly desirable for underwater inspection work. The most basic inspection work is assumed to be to collect images of underwater structures. Therefore, in this study, we created an AUV and conducted an image collection experiment using a pier owned by our university as an example to confirm its functionality.

OS20-2 Design of Disassembly-reassembly Type USV for Coral Reef Research

Keisuke Watanabe, Koki Amano, Gaku Minato, Yasutaka Taniguchi, Konosuke Watanabe
(Tokai University, Japan)

Researchers are conducting physical surveys by diving and swimming to study the effects of microplastics and global warming on coral reefs. The area, time, and water depth that can be investigated by diving are extremely limited. In addition, to collect microplastics, it is necessary to tow the nets, but chartering a ship is expensive. Therefore, the authors are developing a system that simultaneously operates a USV and a UUV to simultaneously observe the sea surface and underwater. In this paper, we conducted a conceptual design, carried out fluid force measurements, trajectory tracking experiments, and image recognition using AI, with the aim of realizing a lightweight USV that can be divided during transportation to coral reef areas and reassembled on site for operation.

OS20-3 Optimization method to improve visual SLAM in dynamic environment

Yufei Liu, Kazuo Ishii (Kyushu Institute of Technology, Japan)

Aiming at the problem that the robustness of the classic visual SLAM system is greatly affected by dynamic feature points in the environment, a method of eliminating dynamic feature points using a target detection algorithm is proposed. Use the target detection algorithm YOLOv5 to identify and classify the collected feature points. Then remove the extracted dynamic feature points, use the remaining static feature points for map construction and positioning, and finally test on the TUM data set. The results show that this algorithm effectively improves the positioning accuracy and robustness of the visual SLAM system.
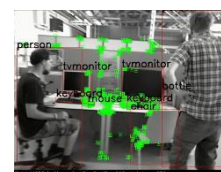
OS20-4 Design of flexible mechanism for flexible manipulator

Huang Jiawei, Kazuo Ishii (Kyushu Institute of Technology, Japan)

Based on the study of the kinematic limitations of the rigid manipulator structure and the characteristics of the existing flexible manipulator in different categories, this paper proposes a structure assumption of flexible manipulator based on the advantages of high load of rigid joint and high flexibility of flexible manipulator. This paper designed a flexible manipulator structure driven by wires. Using the forward kinematics analysis the sport model of the structure. The motion range of the end of the manipulator arm and the length changes of the wires were simulated with the movement of the model.

OS20-5 Driver Drowsiness Detection Method based on Deep Learning

Shi Puwei, Kazuo Ishii (Kyushu Institute of Technology, Japan)

Drowsiness driving will pose a serious threat to the lives of drivers and others. Determining the state of the driver through face recognition has the advantages of low cost and convenience. Therefore, this study deploys the face recognition model to the mobile phone, and finally realizes the recognition of the driver's Drowsiness by the front camera of the mobile phone. The whole research is divided into three parts. The first part is to train the face 68-keypoints recognition model based on the YOLO face algorithm. The second part is to deploy the trained model to the mobile phone using ONNX and NCNN. The third part is to calculate EAR and MAR using several facial key points, and finally complete the recognition of the driver's drowsiness state using EAR and MAR.

OS20-6 An Analysis of Translational Motion for a Mobile Robot with Line-Symmetric Rollers Arrangement

Kenji Kimura[1], Kazuo Ishii[2]

([1]National Institute of Technology, Matsue College, Japan) ([2]Kyushu Institute of Technology, Japan)

In fields such as logistics, robots are required to have efficient mobility. There are various types of rollers used in mobile robots. Among them, omni-rollers have excellent omnidirectional mobility and are easy to control. In this study, mechanism kinematics has been proposed that assumes arbitrary changes in the roller arrangement position on a circular mechanism and the roller arrangement has been evaluated from the viewpoint of speed efficiency. Furthermore, we aim to evaluate the mobility of mobile robots by focusing on their translational components. Moreover, we have examined the behavior of the area of the area generated by the end point of the robot velocity vector and evaluated the velocity efficiency.

OS20-7 Development of Teaching Materials for Robot Programming for Junior High School Students:Student-Based Educational Activities

Kenji Kimura, Youta Takano (National Institute of Technology, Matsue College, Japan)

With the extensive use of robots in recent years, an age of co-existence between humans and robots is expected to arrive in the future. Therefore, providing robot education in the early stages of elementary and junior high school is necessary to stimulate students' interest in robots. Furthermore, educational institutions are becoming more active in robot education as a part of their contribution to the local community. For elementary school students, a Beauto Racer (Vstone) has been adopted as a teaching material for robot education, and teaching materials have been developed for teaching line tracing. In this study, we will develop educational materials for junior high school students on maze exploration using a Beauto Rover (Vstone). The students of NIT, Matsue College took the initiative in this activity. The purpose is to improve their teaching skills.

OS20-8 Design and Software Production of Robotics Educational Design for Elementary and Junior High School Student

Youta Takano, Kenji Kimura (National Institute of Technology, Matsue College, Japan)

The Ministry of Education, Culture, Sports, Science and Technology (MEXT) has been promoting cross-curricular learning including STEAM education in its educational policy for 2021. The purpose of this study is to have students experience not only control engineering but also basic programming techniques, and to have them become interested in mathematical subjects in general, which are the basis for control engineering. The educational design and accompanying software were designed using Beauto Balancer2 (Viston) educational robot and Scratch, a visual programming editor developed by the Scratch Foundation. In addition, workshops were conducted and the lesson design was evaluated by a questionnaire.
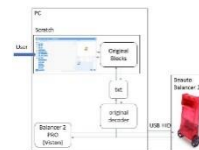
OS20-9 Development of visual inspection system for low-reflective material utilizing of a string shadow

Keiji Kamei[1], Tomorou Kawahara[1], Yoshiyuki Daimaru[2]
([1]Nishinippon Institute of Technology, Japan) ([2]Nissan Motor Kyushu, Japan)

Visual inspection of products in manufacturing plant is usually conducted by workers. However, that inspection process by workers occasionally happen oversight due to careless check, so that automatic visual inspection system is needed from manufacturer. Products made of low-reflective material are difficult to inspect by light reflection because contours of defect areas are got blurred. To solve this problem, we propose to inspect products utilizing of a string shadow. Form our research, a string shadow was different between contours of defect area and its of non-defect. On the other hand, former research did not success to detect defect area due to determination algorithm for defect. In this paper, we apply linear approximation of a string shadow to defect detection. From the experimental results, we succeed in detecting defects.

OS20-10 Development of IoT-Based Remote Monitoring Module for Greenhouse Environment to Facilitate Crop Growth Data Analysis

Moeko Tominaga, Yasunori Takemura, Junya Era, Wataru Kaishita
(Nishinippon Institute of Technology, Japan)
Eiji Mizoe, Tomoyasu Furukawa (Kumamoto Fruit and Vegetable Shippers Association Co. Ltd., Japan)

As the shortage of agricultural workforce becomes a concern, IoT proves effective in alleviating labor burdens. Monitoring, storing, and analyzing the growing environment, especially for new crops, is crucial for subsequent agricultural cycles. However, small farms often face challenges in data collection due to cost constraints, particularly with new crops. An agricultural IoT system was developed for small farmers, featuring a module for collecting, accumulating, and visualizing time-series data from sensors. To demonstrate feasibility, the IoT system was implemented and validated by intermittently monitoring banana and coffee growth parameters in a Kumamoto greenhouse over two years. Results confirmed that the proposed IoT remote monitoring module empowers farmers to monitor crop growth environments.

OS20-11 Development greenhouse environment prediction system using IoT data

Yasunori Takemura, Naoya Nishida, Moeko Tominga (Nishinippon Institute of Technology, Japan)
Eiji Mizoe, Tomoyasu Furukawa (Kumamoto Fruit and Vegetable Shippers Association Co. Ltd., Japan)

Currently, the decrease in the working population is a significant problem in agriculture in Japan. In addition, the prices of imported vegetables and fruits have been rising due to international trade problems such as the weak yen. This study has developed an environmental data storage system using IoT to assist the labor force and to understand the cultivation data of crops that have not been cultivated in Japan. Using this big data, we have constructed and evaluated a forecasting system. In this study, we focus on the domestic cultivation of bananas grown in the tropics as the subject to obtain data and make predictions. For learning, a system was constructed to infer the next day's temperature by time-series analysis, using as input data obtained from the IoT. In this paper, we report a comparison of two time series analysis algorithms.

OS20-12 The Development of SaaS for Quantifying the Amount of Drifted Debris on the Coast

Ayumu Tominaga, Ryohei Komori (National Institute of Technology Kitakyushu College, Japan)
Eiji Hayashi (Kyushu Institute of Technology, Japan)

In the process of collecting drifted debris along a coastal area, it has been required to quantify the collection results. This study introduced a Software as a Service (SaaS) equipped with image processing capabilities for the detection and identification of debris in photographs. The SaaS was developed to facilitate the automated weighing of the collected debris. In the beach cleanup event in June 2023, 317 images capturing debris found during the coastal cleanup were gathered on a cloud server from 90 participants via this SaaS. These images were subsequently analyzed by object recognition AI executed on a cloud server, leading to the detection and identification of 954 instances of debris. This result represents the actual amount of debris collected during this cleanup event, and indicates that the achievements of the event were quantified via the SaaS.

OS20-13 Development of an antagonistic wire-driven joint mechanism capable of rapid motion and variable stiffness

Katsuaki Suzuki[1], Yuya Nishida[2], Kazuo Ishii[2]
([1]Kumamoto Industrial Research Institute, Japan, [2]Kyushu Institute of Technology, Japan)

The advancement of digital transformation in manufacturing is expected to demand novel tasks for industrial machinery unlike before. When adding mechanisms or actuators to enhance machine functionality, concerns arise about potential system enlargement and increased complexity. The design concept of integrating multiple functions into mechanisms using a few actuators and components serves as one means to address these issues. We aim to elucidate the structural arrangement capable of achieving three functions, including normal motion, rapid motion, and variable stiffness, using two electric actuators. In this paper, we introduce a newly devised antagonistic wire-driven joint mechanism aimed at improving the variable stiffness function within the articulated mechanism we have developed, enabling three functionalities.

OS20-14 Gakken Hills Interdisciplinary Ekiden Competing with Humans, Animals, and Robots

Takuya Fujinaga[1], Moeko Tominaga[2], Daigo Katayama[3], Kazuo Ishii[3]
([1]Fukuoka University, [2]Nishinippon Institute of Technology, [3]Kyushu Institute of Technology, Japan)

With the advancement of science and technology, the societal implementation of robots has been expected. In recent years, the number of robots exiting in human living spaces has increased, and interaction between humans and robots has become an important issue. The purpose of this study is to propose a future society where humans, animals, and robots coexist. We organize relay races, or ekiden in Japanese, where they compete together as runners to discuss the future society. The competition rules have been established for this ekiden to ensure that each runner can compete fairly. For example, an advantage is given to the time of humans, animals, and robots, so that the results are the same. This paper explains the competition rules and reports the competition results of the 8th race held in 2023.

OS20-15 Development of a Low-Cost Underwater Robot for Research and Education

Takuya Fujinaga (Fukuoka University Japan)

Underwater robots are frequently utilized for underwater observation. When used as tools, commercially available underwater robots are sufficient. However, when utilized for research, development, or education, the required specifications vary based on the purpose, necessitating custom development. Generally, the development of underwater robots demands a significant amount of funds, posing a high barrier for research, development, and education. The aim of this study is to reduce these barriers, facilitating broader implementation of research, development, and education in the field of underwater robotics. This paper explains the development requirements, process and cost of a low-cost underwater robot designed for a robotics competition and discusses key considerations during development.

OS21 Human-Machine Interface (3)
Chair Norrima Mokhtar (University of Malaya, Malaysia)
Co-Chair Heshalini Rajagopal (UCSI University, Malaysia)

OS21-1 Enhancing Reconnaissance Missions Through Multiple Unmanned Systems in ROS

Anees ul Husnain[1,2], Norrima Mokhtar[1], Takao Ito[3], Siti Sendari[4], Muhammad Farris Kyasudeen[5], Muhammad Badri M Noor[1,6], Heshalini Rajagopal[7]
([1] Universiti Malaya, Malaysia), ([2] The Islamia University of Bahawalpur, Pakistan)
([3]Hiroshima University, Japan), ([4] Universitas Negeri Malang, Indonesia)
([5] University Technology MARA (UiTM), Malaysia)
([6] Ifcon Technology Sdn Bhd, Malaysia), ([7]UCSI University, Malaysia)

The synergistic collaboration between UAVs and UGVs addresses the limitations of individual platforms, offering a versatile solution for reconnaissance tasks in diverse environments. The proposed system employs ROS as the underlying architecture to facilitate seamless communication and coordination among multiple UAVs and UGVs. Optimal coverage is ensured through efficient exploration and coverage of the reconnaissance area, under by comparing raster-scan, expanding spiral and zig-zag area exploration approaches. The article concludes by discussing potential extensions, such as the integration of machine learning techniques for enhanced autonomy and the scalability of the system for larger-scale missions by presenting a ROS-based framework that maximizes the synergy between UAVs and UGVs.

OS21-2 Illumination Effects on Facial Expression Recognition using Empirical Mode Decomposition

Hashimah Ali[1], Wan Khairunizam Wan Ahmad[1], Hariharan Muthusamy[2], Mohamed Elshaikh[1]
([1]Universiti Malaysia Perlis, Malaysia)
([2]National Institute of Technology Uttarakhand, India)

Thus, this paper aims to investigate the illumination effects (brightness variations) on facial expression recognition using empirical mode decomposition reconstruction techniques. In this framework, firstly, the noisy facial expression images were simulated with the illumination effects using different brightness levels of 30%, 40%, 50%, 60%, and 70%. Then, the EMD will decompose the noisy facial expression images into a small set of intrinsic mode functions (IMF), namely IMF1, IMF2, IMF3, and residue. Based on property held by EMD, the signals are decomposed into several IMF components, each with a different time scale. Because the last several IMFs represent the majority of illumination effects, various reconstruction techniques for IMFs have been investigated at various brightness levels. Feature reduction techniques Principal component analysis (PCA) and linear discriminant analysis (LDA) have been employed to reduce the high-dimensional space of IMF features into low-dimensional IMF features. The reduced IMF reconstructions were then used as input to the k-nearest neighbour classifier to recognise the seven facial expressions. A series of experiments have been conducted on the JAFEE database using various reconstruction IMFs together with PCA plus LDA. Based on the results obtained, the reconstruction of IMF1 + IMF2 + IMF3 shows the highest accuracy in high illumination conditions, which is 99.06%.

OS21-3 Supercontinuum Generation Pump By a Molydenum Disulfide Based Soliton Mode-Locked Fiber Laser

Aeryn D. Ahmad[1], Norrima Mokhtar[1], Hamzah Arof[1], Sulaiman Wadi harun[1], Ahmad Haziq Aiman Rosol[2]
([1]Universiti Malaya, Malaysia), ([2]MJIIT, UTM, Malaysia)

In this work, a highly stable soliton mode-locked Erbium-doped fiber laser (EDFL) is passively obtained using a molybdenum disulfide (MoS2) thin film as a saturable absorber (SA). The MoS2 thin film obtained via electrochemical deposition technique is integrated into an EDFL cavity to generate mode-locked pulses operating at 1.88 MHz with a pulse duration of 3.03 ps. Supercontinuum (SC) light is generated using the proposed soliton mode pulses operating at 1560.4 nm as they are injected into a 100 m long highly nonlinear photonic crystal fiber (HN-PCF) after it is amplified to the output power of 17.8 dBm. The SC light operates in a wavelength range starting from 1360 nm to more than 1750 nm with the intensity above -35 dBm. The proposed supercontinuum laser can be seen as a promising light source for metrology and sensing applications.

87

OS22 Mathematical Informatics (8)
Chair Takao Ito (Hiroshima University, Japan)
Co-Chair Makoto Sakamoto (University of Miyazaki, Japan)

OS22-1   Adsorption Behavior of Arsenic and Selenium using NiZn Hydroxy Double Salts with Acetate, Chloride, Nitrate, and Sulfate Anions

Kaoru Ohe, Ryosuke Tabuchi, Tatsuya Oshima (University of Miyazaki, Japan)

Arsenic(As) and selenium(Se) has become an increasingly serious water contamination worldwide, so the development of adsorbents to improve the adsorption performance of As and Se is desired. In this study, hydroxy double salt NiZn-AcO, -Cl, -NO$_3$, and -SO$_4$, with acetate, chloride, nitrate, and sulfate anions between layers and the adsorption behavior of As and Se was investigated. The Langmuir isotherm model, characteristic of monolayer adsorption, fit the data best with $R^2 > 0.94$. The adsorption capacity(q_{max}) of As(V) by NiZn-Cl was about twice that of the other adsorbents. The q_{max} of Se(IV) by NiZn-AcO, NiZn-Cl, and NiZn-NO$_3$ was almost the same value as that of Se(VI) and 1.4-1.7 times larger than that of NiZn-SO$_4$. The prepared adsorbent was confirmed to be useful for the removal of As and Se.
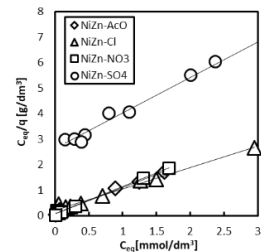
Fig.1 Langmuir plots of Se(VI) on NiZn-AcO, -Cl, NO$_3$ and SO$_4$

OS22-2   Parallel acoustic analysis based on the domain decomposition method with higher-order element

Amane Takei[1], Makoto Sakamoto[1], Akihiro Kudo[2]
([1]University of Miyazaki, Japan) ([2]National Institute of Technology, Tomakomai Collage, Japan)

Numerical analysis approaches, such as the finite element method, are widely used for the estimation of the sound field in architectural spaces. Large-scale analyses, using numerical models with over 10 trillion elements, are required for the analysis of a large space such as a concert hall with higher-frequency bands. Large spaces are often limited to low-frequency analysis. In this study, the number of elements is reduced by wave acoustic analysis using higher-order elements. Based on the results using higher-order elements, it is shown that it is possible to analyze a real environment model such as a live music club and a concert hall.

(a)   Visualization results (X-Y)
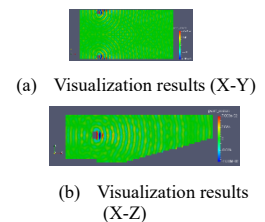
(b)   Visualization results (X-Z)
Fig.1 Numerical result using concert hall model

OS22-3   Sound field evaluation on the acoustical experiment using non-steady state analysis

Akihiro Kudo[1], Makoto Sakamoto[2], Amane Takei[2]
([1]National Institute of Technology, Tomakomai College, Japan) ([2] University of Miyazaki, Japan)

Identifying acoustic effects of sound waves emitted by a sound source on the listener is important for understanding the localization of sound as perceived by the listener. In this presentation, we will analyze the effect of the presence of a chair on the psychoacoustic effect using a practical model simulating a psychoacoustic experiment consisting of a loudspeaker, a listener, and a chair. Non-steady state analysis is required to analyze changes in acoustic effects over time. To analyze the interference of sound waves in the vicinity of the listener, a short Gaussian pulse-like solitary wave is set up on the loudspeaker's vibrating surface. Since sound waves in the audible band are used for these simulations, a mesh size of less than 25 mm is used. The laboratory is 3m of depth and width, and 2.5m of height, which means that the computational whole target consists of over 15 million elements. This simulation requires a large number of calculations, so a parallel computing method with ADVENTURE_Sound that is an open-source software for sound simulation, is used to run the simulation. From the simulation results of the non-stationary analysis and consider the effect of the presence of the chair on the sound image localization.

Fig.1 Example of results of non-steady state analysis with gaussian pulse

OS22-4 A DeepInsight method with Morphological Analysis

Toyoaki Tomioka[1], Satoshi Ikeda[1], Makoto Sakamoto[1], Takao Ito[2]
([1]Faculty of Engineering, University of Miyazaki, Japan) ([2]Hiroshima University, Japan)

The DeepInsight method is a comprehensive approach aimed at applying convolutional neural networks (CNN) to non-image data. This study specifically focuses on using the DeepInsight method to forecast stock prices, which involve time series data. To achieve this, we employ morphological analysis to extract the necessary word feature for predicting stock prices from news articles. Additionally, t-SNE is utilized to reduce the dimensionality of the high-dimensional data, transforming it into a more manageable, low-dimensional format suitable for stock price prediction. After reducing dimensionality in both news text data and stock price time series data using t-SNE, and subsequently applying CNN for learning from the transformed representations, predictions were made with an accuracy of approximately 60%.

OS22-5 Support for Museum Exhibition of Small Fungi using AR Technology

Kakeru Takemura[1], Ota Hamasuna[1], Fumito Hamakawa[1], Satoshi Ikeda[1], Kaoru Ohe[1], Amane Takei[1],
Makoto Sakamoto[1], Shuichi Kurogi[2]
([1]University of Miyazaki, Japan) ([2]Miyazaki Prefectural Museum of Nature and History, Japan)

There are 25 species of bioluminescent fungi that have been confirmed in Japan, and 12 species have been confirmed in Miyazaki Prefecture. Those fungi are very small. Therefore, it is difficult to observe the structure of fungi exhibited in museums with the naked eye. The purpose of this research is to display 3DCG models of mushrooms using AR (Augmented Reality) technology in order to facilitate observation of these small mushrooms exhibited in museums. Two hundred visitors to the museum were asked to use the application and complete a survey. In order to measure the ease of observation of different app functions and mobile devices used in the survey, we divided the respondents into four groups.

OS22-6 Automatic Selection of High-Grade Dried Shiitake Mushrooms using Machine Learning

Ota Hamasuna[1], Kakeru Takemura[1], Kodai Hasebe[1], Fumito Hamakawa[1], Bidesh Biswas Biki[1], Satoshi
Ikeda[1], Kaoru Ohe[1], Amane Takei[1], Makoto Sakamoto[1], Kazuhide Sugimoto[2]
([1]University of Miyazaki, Japan) ([2]SUGIMOTO Co., Ltd. Japan)

SUGIMOTO Co., Ltd. collects dried shiitake mushrooms directly from approximately 600 producers in Takachiho. Shiitake mushrooms are at their peak in spring and fall, and at peak times, more than 1 ton of mushrooms can be brought in each day. However, shiitake mushrooms are still sorted manually, and sorting this much requires a lot of effort. In this paper, we aim to efficiently determine the quality of shiitake mushrooms using image processing and deep learning.

OS22-7 Predicting High Volatility Cryptocurrency Prices using Deep Learning

Tsutomu Ito[1], Kodai Hasebe[2], Fumito Hamakawa[2], Bidesh Biswas Biki[2], Satoshi Ikeda[2], Amane Takei[2],
Makoto Sakamoto[2], Md Riajuliislam[3], Sabrina Bari Shital[4], Takao Ito[5]
([1]National Institute of Technology, Ube College, Japan) ([2]University of Miyazaki, Japan)
([3]TU Dortmund, Germany) ([4]Daffodil International University, Bangladesh) ([5]Hiroshima University, Japan)

Even if you want to make a profit from cryptocurrency, you are worried that you will lose money, and it is difficult to afford it. There are a vast number of papers that study such unpredictable price fluctuations of cryptocurrency. Currently, it is mainstream to use deep learning to predict the price of cryptocurrency. The goal of this research is to predict the price of cryptocurrency over the long-term using deep learning. The algorithms used are LSTM, GRU, and Bi-LSTM. The cryptocurrency targeted are Bitcoin, Ethereum, Litecoin, and Cardano. Finally, we will compare it with previous research and verify the performance of our model.
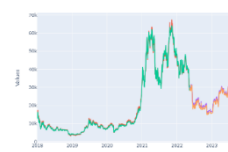
OS22-8 Simulation of Weathering Representation using Vertex and UV Information

Tsutomu Ito[1], Fumito Hamakawa[2], Kodai Hasebe[2], Satoshi Ikeda[2], Amane Takei[2], Makoto Sakamoto[2], Takao Ito[3] ([1]National Institute of Technology, Ube College, Japan) ([2]University of Miyazaki, Japan), ([4]Hiroshima University, Japan)

In recent years, three-dimensional computer graphics (3DCG) technology has been developed. In addition, a plethora of research have been published on weathering representations such as rust and moss for realistic representation. However, when outputting simulation results on an image, an enormous number of images are required to display different simulation results in a three-dimensional space. In this paper, a simulation method using vertex information of a 3D model and simple images is proposed. In this method, when the number of vertices is sufficient, the simulation result output does not use images, thus reducing the data increase. This figure is one of the experimental results.

OS23 Industrial Revolution (5)
Chair Hazry Desa (Universiti Malaysia Perlis (UniMAP), Malaysia)

OS23-1 Investigating the Engineering Interventions in the Conservation of Malaysia Heritage Structures: A Review on Preserving Historical Edifices Through Advance Civil Engineering Techniques

Muhammad Azizi Azizan, Nurfadzillah Ishak, Hazry Desa (COE-UAS, UniMAP, Malaysia)

This review delves into the realm of Malaysia's heritage conservation, spotlighting the transformative impact of advanced artificial robotic aid in civil engineering techniques. Through the integration of Fiber-Reinforced Polymers and nanotechnology, historical edifices are fortified, seamlessly blending modern engineering with architectural elegance. Non-destructive testing (NDT) methods, including ground-penetrating radar and Finite Element Analysis (FEA), empower conservationists with deep insights into structural intricacies, guiding targeted interventions. In the digital sphere, 3D laser scanning captures intricate details, while Virtual Reality (VR) simulations facilitate immersive exploration and decision-making. This harmonious fusion of cutting-edge engineering ensures the enduring legacy of Malaysia's architectural treasures.
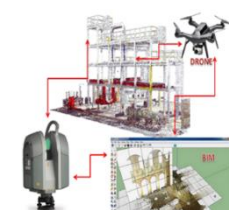
OS23-2 Drones and Data: A Comprehensive Exploration of UAVs in Data Mining

Muhammad Azizi Azizan, Nurfadzillah Ishak, Hazry Desa (UniMAP, Malaysia)

Encapsulates a comprehensive investigation into the symbiotic relationship between Unmanned Aerial Vehicles (UAVs) and data mining, as encapsulated in the discourse titled "Drones and Data." The technological transitions facilitated by UAVs, emphasizing the integration of machine learning algorithms, cloud-based data processing, and the collaborative synergy between stakeholders. Therefore, this study a glimpse into the intricate web of applications and technological advancements at the intersection of UAVs and data mining. It serves as a scholarly guide through the evolving landscape of "Drones and Data," UAVs play a central role in unlocking unprecedented insights and efficiencies, reshaping the future of data mining.

OS23-3 Development of Variable Arm to Control the Manoeuvrability of Quadrotor

L. Y. Hong, H. Desa and M. A. Azizan (UniMAP, Malaysia)

M. H. Tanveer (Kennesaw State University, USA)

This paper introduces a variable arm concept for quadrotors, enabling manoeuvrability by adjusting arm length. Variations in arm's length impact thrust-generated bending moments, tilting and directing quadrotor movement. The main aim is to create a quadrotor with adaptable arm length to efficiently control maneuvers, reducing the need for extra thrust during flight. The focus is on designing a quadrotor capable of extending or retracting its arms. The proposed idea centers on using the variable arm to alter bending moments and control manoeuvrability. Testing a quadrotor with this adjustable arm showed successful performance in agile manoeuvres. Results confirm the variable arm's ability to induce body rotation, effectively regulating quadrotor manoeuvrability.
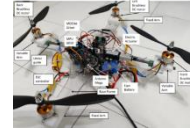
OS23-4 Development of IOT-Enabled Smart Water Metering System

S. D. Wen, H. Desa and M. A. Azizan (UniMAP, Malaysia)

A. -S. T. Hussain (Al-Kitab University, Iraq) M. H. Tanveer, R. Patan (Kennesaw State University, USA)

This paper presents a smart IoT water meter for automated readings, aiming to create both a device and a compatible mobile app. Instead of manual reading, the IoT-enabled meter, using a camera and CNN, accurately captures data. It utilizes ESP32 CAM for collection, a laptop as a gateway, and MQTT for transfer. Data is stored in Firebase and analyzed via the app. A functional prototype tested in a housing area collects data, monitored through the app. The analysis evaluates method suitability, offering future improvement recommendations.
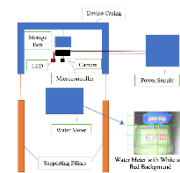
OS23-5 Object Detection and Instance Segmentation with YOLOV8: Progress and Limitations

L. J. Lee, H. Desa and M. A. Azizan (UniMAP, Malaysia)

A. -S. T. Hussain (Al-Kitab University, Iraq) M. H. Tanveer (Kennesaw State University, USA)

This research employs object detection and instance segmentation algorithms to distinguish between objects & backgrounds and to interpret the detected objects. The YOLOV8 (You Only Look Once) framework and COCO dataset are utilized for detecting and interpreting the objects. Additionally, the accuracy of detection, segmentation, and interpretation is tested by placing objects at various distances from the camera. The algorithm's performance was evaluated, and the results were documented. In the experiments, a sample of 11 objects was tested, and 8 of them were successfully detected at distances of 45cm, 75cm, 105cm, and 135cm. For instance, segmentation, segmentation maps appeared clean when detecting a single object but faced challenges when multiple objects overlapped.

OS24-Robotics and Intelligent Casting (5)
Jiwu Wang (Beijing Jiaotong University, China)

OS24-1 A high-performance motion planning method based on asymptotically optimal RRT
Tianbin Meng (Beijing Jiaotong University, China), Jiwu Wang (Beijing Jiaotong University, China)

The motion planning algorithm of the robot arm plays an important role in the working process of the robot arm, especially in the complex environment, an efficient algorithm is more conducive to the robot arm to complete the corresponding planning task. Aiming at the problems such as low exploration efficiency and poor planning path in the current motion planning task of robotic arm, we propose a distance constraint mechanism. Based on RRT^* and $Informed-RRT^*$, the algorithm uses halton sequence to generate random points and introduces the current lowest cost path length. To avoid useless nodes extension. The simulation results show that the algorithm with distance constraint mechanism can improve the exploration efficiency and planning quality to some extent.

OS24-2 CSM-RRT*: an improved RRT* algorithm based on constrained sampling mechanism
Hang Yang (Beijing Jiaotong University, China), Jiwu Wang (Beijing Jiaotong University, China),
Xueqiang Shang (Aero Engine Corporation of China, China)

In this paper,an improved RRT* algorithm(CSM-RRT*) based on constrained sampling mechanism is proposed .The entire planning process is divided into two steps: fast exploration and optimization of the initial path.Firstly,with the removing of redundant nodes and saving of collision nodes by constrained sampling strategy,the number of redundant nodes and collision times is reduced which avoids excessive exploration of repetitive and collision regions.The target bias strategy is used to reduce unnecessary expansion while preserving the node tree's ability to explore unknown regions and quickly finding a path connecting the starting and ending points.Subsequently, a dynamic sampling region consists of removed redundant nodes and collision nodes is formed around initial path .By prioritizing exploration within this dynamic region,computational resources can be saved and the asymptotic optimal path can be quickly converged from the initial path.Eventually,simulation results presented in various obstacle cluttered environments confirm the efficiency of CSM-RRT*.
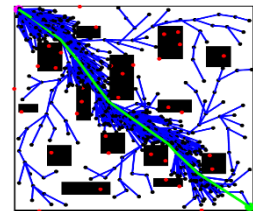
OS24-3 Small Sample Object Detection Based on Improved YOLOv5
Yuxuan Gao(Beijing Jiaotong University, China),Jiwu Wang(Beijing Jiaotong University, China), Zixin Li(Aero Engine Corporation of China)

Object detection is widely used in various production and life, such as mask detection and recognition during the epidemic, face recognition with masks. Object detection algorithm based on deep learning has always been an important research content and implementation method in the field of object detection. Due to the large number of lead seals and fuses, their locations are not fixed, the lead seals and fuses have difficulties such as few sample datasets, complex target background and easy to be blocked, and strong reflective interference, and the conventional image processing methods are difficult to solve the problem of effective object recognition. In this study, by expanding the datasets, using different data enhancement methods, and training in the improved algorithm, the detection accuracy, detection speed, and adaptability were effectively improved.

OS24-4 Research on Gas Pore Prediction Method Based on Sand Core Characteristic Time

Xiaolong Wang[1], Qihua Wu[2], Jiwu Wang[1], Jinwu Kang[3], Na Li[2], Yucheng Sun[2]
([1]Beijing Jiaotong University, China) ([2]Weichai Power Co., Ltd. China) ([3]Tsinghua University, China)

In the production of castings, gas pores are a prevalent defect, particularly in components where air-tightness is crucial, such as cylinder blocks and heads. These defects primarily arise from the intrusion of gases into the casting during the combustion of resin in the sand core. Due to the complexity of the sand core's gas evolution and venting process, predicting gas pores using numerical simulation poses significant computational and time challenges. This paper introduces a rapid prediction method for gas pores based on the characteristic time of the sand core. By setting the heat transfer boundaries, initial conditions, and termination criteria for computation, the thermal conductivity of the sand core is adjusted. The termination computation time is used as the characteristic time of the sand core. This time is then compared with a critical characteristic time to predict the distribution of gas pores. As the analysis of more sand cores is incorporated, the precision of the critical characteristic time improves, leading to more accurate predictions of gas pores in castings.

OS24-5 Optimization Analysis of a Deep Learning-Based Model for Predicting Temperature Fields in the Solidification Process of Castings

Yahui Yang, Jiwu Wang (Beijing Jiaotong University, China) Jinwu Kang (Tsinghua University, China)

In the foundry industry, accurate prediction of the temperature field of castings during solidification is essential to ensure high product quality and improve productivity. This study provides an in-depth investigation of the Unet deep learning model developed specifically for this purpose, focusing on the impact of key training parameters such as optimiser selection, batch size, number of iterations and loss function selection on the prediction performance of the Unet model. The findings provide practical insights on how these parameters can be selected and tuned to improve the accuracy and reliability of model predictions. The results of this study not only provide new ideas for the practice of deep learning applications in the foundry industry, but also help to improve the accuracy and efficiency of the production process.
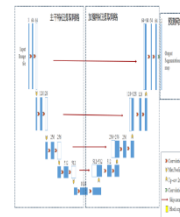
OS25 Research Towards the Sustainable Development Goals (SDG's) (10)
Chair Ammar A.M. Al Talib (UCSI University, Malaysia)
Co-Chair Takao Ito (Hiroshima University, Japan)

OS25-1 Portable Green Energy Mobile Laptop Charging Station

Ammar A. M. Al-Talib[1], Rodney Tan[1], Ang Aun Jie[1], Idayu M. Tahir[1], Sarah Atifah Saruchi[2], Cik Suhana Bt. Hasan[1], Amar Rizwan[1]
([1]UCSI University) ([2]UMPSA, Malaysia)

Mobile phones and laptop computers require electrical power to recharge when the battery is down.. As a result, it would be very useful if a portable charging station derived from renewable energy harvesting could be built, so that individuals could recharge their phones and laptops whenever needed. The objective of this project is to design and develop a green energy mobile and laptop charging station that uses wind and solar energy and evaluate the performance of the designed station under different working conditions. The efficiency of the power generated to charge the station is achieving 95.6% for solar charging, which is considered a high efficiency for a renewable energy charging station. Based on the analysis of the charging station results, it has been proven that it can provide sufficient power and is safe for use as a portable mobile laptop charging station.

OS25-2 Auto Indoor Hydroponics Plant Growth Chamber

Ammar A.M. Al-Talib[1], Tew Hwa Hui[1], Sarah Atifah Saruchi[2], Idayu M. Tahir[1], Nor Fazilah Binti Abdullah[1]
([1]UCSI University) ([2]UMPSA, Malaysia)

The objective of this project is to build an auto indoor hydroponics plant growing chamber that has an auto monitoring and controlling system. A ESP32 based hydroponics electrical system is built with the attachment of hardware components such as temperature and humidity sensor, light intensity sensor, water level sensor, and water flow rate sensor. The software development of the system is through Arduino IoT Cloud platform, which has an overall suitability in terms of features, cost, and user intuitiveness for starters. Results have shown that ESP32 can ensure stable power supply.. After testing and validation, all of the electrical components are stored in a power enclosure box to prevent contact with liquid. In short, the developed auto indoor hydroponics plant growth chamber has effectively demonstrated the ability in easing the plant cultivation procedure for agricultural community.

OS25-3 A Design and Fabrication of a Solar Agriculture Water Pumping System

Ammar A.M. Al-Talib[1], Idayu M.Tahir[1], Ain Atiqa[1], Amar Rizwan[1], Sarah Atifah Saruchi[2], Yazan Abu Al shaikh[1]
([1]UCSI University) ([2]UMPSA, Malaysia)

This study explores the use of a solar driven water pump. PV technology replaces conventional electricity and diesel pumps by using solar energy to power DC or AC water pumps. The main objective of this study is to design and construct a solar-powered agriculture water pumping system and to evaluate its performance. The solar agriculture water pumping system used in this project consists of a 40-watt monocrystalline solar cell with an efficiency conversion of between 23% – 24%. This can supply power to a 16.8 W DC   Flow   Submersible Pump. It could lift the water up to 5m and a flow range of 700 Liter/Hour. The system also includes a PWM 30A Solar Charge Controller to regulate the input power to a 12V, 7.2A Sealed Rechargeable Battery. Finally, a 20m long watering kit with nozzles irrigation system is connected to a 12V DC Submersible Pump to water the plants.

OS25-4 Design and Performance of a Power Generating Manual Treadmill

Ammar A. M. Al-Talib[1], Sarah Atifah Saruchi[2], Cik Suhana Bt. Hasan[1], Nor Fazilah Binti Abdullah[1], Ain Atiqa[1], Ahmad Jelban[1]
([1]UCSI University) ([2]UMPSA, Malaysia)

Treadmills are one of the most popular training equipment in the gym and at home. The working principle of treadmills is by moving the belt with the human knee bending, which creates mechanical energy to turn the belt. A gear or pulleyand belt system connects to the generator along the axel line of the rolling bars. The power generated by the DC generator is stored in a battery pack and could be used to charge phones or other equipment. It has been found that treadmills can provide an efficiency of 95% when the DC motor is used and 92% when the AC motor is used. The main objective of this study is to design and fabricate a powder-generating manual treadmill and to analyze the performance of the system under differentoperation conditions.

OS25-5 Smart Car Jack Using Internet of Things

Idayu M.T. Noor, Ammar A.M. Al- Talib, Mahmoud E.A. Zeiad, Suhana B.H. Cik,
(UCSI University, Malaysia)

This paper presents a modified car jack design that is safe, easy to operate, and reduces physical effort for lifting and lowering vehicles during automobile repair, especially for roadside situations. The study explores the integration of Internet of Things (IoT) technology in automotive tools, particularly in the development of smart car jacks. The proposed smart car jack offers convenience and enhanced control over vehicle maintenance tasks. Comparative analysis between smart car jacks and traditional jacks reveals several advantages of the former. Smart jacks demonstrate higher lifting capacities and efficiency due to their use of hydraulic systems and advanced motorized mechanisms. The most significant differentiator is the integration of technology in smart car jacks. Equipped with microcontrollers, sensors, and mobile applications, they offer remote control and real-time monitoring features. These technological advancements provide users with enhanced convenience and safety, setting smart jacks apart from their traditional counterparts.
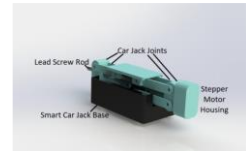
OS25-6 Gas Detection for Biogas System Using Internet-of-Things (IoT)

Ammar A.M. Al Talib, I.H.W. Yang, Idayu M.T. Noor, Haslija A.B. Ayu, Afifi. Z. Nur Muhammad
(UCSI University, Malaysia)

The melting of the polar icecaps, rising sea levels, increase in allergies and infectious disease outbreaks, those are only a few effects of global warming. Global warming remains one of the most detrimental by-products of industrialization. Fossil fuels contribute to the majority of greenhouse gases emitted but remain a popular option for the generation of energy. An easy fix for this conundrum is to utilize other forms of fuel for energy generation which burns cleaner and renewable as opposed to fossil fuels. The solution would be to use biomethane generated from waste products which burn cleaner and comes from a renewable source. In this paper, an IoT based biogas monitoring system for biogas reactors is proposed. An ESP-32 microcontroller system is deployed and tested to detect the presence of gas production. A dashboard plotting the data obtained from sensors is designed to help user monitor parameters.

OS25-7 IoT Based Smart Mushroom Growing Kit

Ammar A.M. Al- Talib, C.K.J. Ting, Noor Idayu M. Tahir, Ain Atiqa, T.Y. Hui
(UCSI University, Malaysia)

This project introduces an IoT-based smart mushroom growing kit to meet the rising global demand for high-quality mushrooms. Various species of mushrooms can be efficiently grown in different environmental conditions with the help of IoT devices that enable farmers to regulate the climate condition according to the specific needs of each type of mushroom. The kit employs sensors and actuators, including temperature, humidity, MQ-135, and ultrasonic sensors, along with an ESP32 camera, controlled by a microcontroller. The collected data is transmitted to an IoT platform via Wi-Fi, facilitating real-time monitoring and control through a user-friendly dashboard on Blynk website and Blynk app. This innovative system optimizes mushroom cultivation by adjusting environmental conditions, offering efficiency and profitability. Users can remotely monitor and regulate the growth environment through their smartphones, enhancing the overall mushroom cultivation experience.

OS25-8 Design and Analysis of Artificial Magnetic Conductor for Metal Shielding Applications in RFID Car Detection Applications

Eryana Hussin[1,2], Azman Zakariya[2], Md. Ashraful Haque[2,3], Nur Izzati Ali[4]

([1]UCSI University) ([2]UTPSA, Malaysia) ([3]Daffodil International University, Bangladesh) ([4]UMP, Malaysia)

This paper presents the design of symmetrical Artificial Magnetic Conductor (AMC) for metal shielding in RFID applications. It integrates a simple straight dipole as the RFID tag representation, optimizing signal fidelity within metallic environments. Tailored for low-frequency RFID use, the multiple-array AMC addresses UHF bandwidth limitations by incorporating slots, ensuring a compact design and expanded bandwidth. This proposed structure enhances performance of the dipole which includes return loss, directivity, and bandwidth.
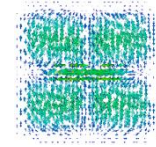
OS25-9 Effect of Fiber Orientation on the Mechanical Performance of Natural Fiber Polymer Composite Bicycle Frame using Finite Element Analysis

Kok Sem Too, Cik Suhana Hassan, Nor Fazilah Abdullah, Ammar Abdulaziz Majeed Al-Talib
(UCSI University, Malaysia)

In this research, the performance of oil palm empty fruit bunch (OPEFB) fibre-reinforced epoxy composite with varying fibre orientation and stacking sequence as the material for mountain bike frame was studied utilizing ANSYS software. The choice of OPEFB fibre was motivated by the fact that the waste by-product of oil palm extraction in Malaysia alone might reach 70-80 million tons per year, with 90% of oil palm biomass lost as waste. The properties of epoxy OPEFB composite in principal 1, 2, and 3 directions were calculated using Whitney and Riley estimates. 10 stacking sequences and five loading conditions were taken. The results have shown that the fibre orientation of epoxy OPEFB composite on the bicycle frame had little effect on the performance contrary to the number of plies in the laminate or number of laminates which had major effects.
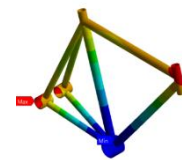
OS25-10 Modelling of Short-Circuit Protection for A Residential Grid-Connected BESS

Kong De Kang, Farah Adilah Jamaludin, Rodney H.G. Tan
(UCSI University, Malaysia)

This research paper presents the power protection study on a grid-connected Battery Energy Storage System (BESS) in a typical Malaysia low-voltage (LV) residential network. The BESS model and control algorithm is developed in MATLAB/Simulink environment to provide a platform for short circuit fault analysis of BESS. The BESS model can charge and discharge its energy with an algorithm-controlled bidirectional AC/DC converter. The paper also presents two cases for BESS short circuit.The protection level on BESS is also optimized by introducing a time-delay characteristics model to coordinate circuit breakers in compliance with standards outlined in IEEE Std 1375-1998, IEC/EN 60898-2, IEC/EN 60947-2, UL 1077, and CSA 22.2. No.235. As a conclusion, BESS was interrupted within stipulated time. In case of internal short circuit protection failure, a backup protection will act on isolating BESS from the grid provided that the discharged current exceeds system capacity.
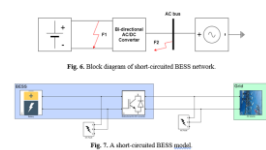
OS26 Navigating the Digital Frontier: Innovations in the Age of Industry Revolution 4.0 (11)
Chair Wei Hong Lim (UCSI University, Malaysia)
Co-Chair Takao Ito (Hiroshima University, Japan)

OS26-1 An Intelligent Cargo/Warehouse Management System

Zhongheng Sun[1], Zhou Yue[1], Xun Sun[1], Wenzhuo Fan[1], Wenxuan Zhou[1]
([1]Fuzou University, China)

This paper presents a low-cost and user-friendly warehouse management system developed using Arduino and ESP8266 hardware. The system accurately monitors temperature, humidity, and harmful gas concentration, and acts as an intrusion detector in a warehouse environment. It also includes functions for managing and detecting goods using RFID chips and for automatic delivery using unmanned intelligent vehicles. Data is uploaded to a cloud-based relational database in real time. The system provides a cost-effective and efficient solution for warehouse management.

OS26-2 A Comprehensive Approach to Design and Implement an IoT-Enabled Intelligent Shopping Cart System with Obstacle-Aware Navigation and Enhanced Customer Engagement for Elevated Consumer Experiences

Yao Chen[1], Jiacheng Du[1], Bo Peng[1], Ningfei Wang[1], Zehan Huang[1], Wei Hong Lim[2], Sew Sun Tiang[2], Mastaneh Mokayef[2], Chin Hong Wong[1]
([1]Fuzhou University, China) ([2]UCSI University, Malaysia)

Supermarket shopping is a universal experience. This project focuses on enhancing this experience by designing an intelligent, user-friendly shopping cart and an interactive webpage. The cart features include accessing item information, automatic movement, and obstacle detection. Users can view their itemized costs and control the cart's movements through the webpage. Utilizing an ultrasonic module, the cart autonomously navigates and follows specific customers, incorporating automatic barcode scanning for product identification. Connectivity is achieved through a WiFi module using the MQTT protocol, facilitating server communication. This development exemplifies Internet of Things applications, showcasing how internet connectivity can significantly increase convenience in everyday activities.

OS26-3 Design of a Four-Port Flexible UWB-MIMO Antenna for Wearable and IoT Applications

Jia Wei Tan, Sew Sun Tiang, Kim Soon Chong, Mohammad Arif Ilyas, Mastaneh Mokayef, Zhi Ying Yeoh, Wei Kang Lai, Wei Hong Lim (UCSI University, Malaysia)

This paper introduces a compact, four-port MIMO antenna for ultrawideband (UWB) applications, measuring 60×60 mm². Printed on a single-layer flame resistant (FR-4) substrate (permittivity 4.3, thickness 1.6mm), the antenna features four microstrip cells, each orthogonal to its neighbor for improved isolation. It includes a rectangular patch with staircase slits and a stepped feed line. Square stubs at the top center enhance isolation. The antenna boasts a significant return loss (-43.75dB), wide impedance bandwidth (1.967-12GHz), and isolation below -15dB. Its envelope correction coefficient (ECC) is under 0.02 with a moderate 4.4dBi gain. Although its 63% radiation efficiency could be enhanced, the antenna's ultrawide bandwidth and compactness make it suitable for UWB wearable IoT applications.
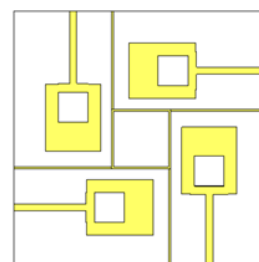
OS26-4 Investigate Power Efficiency in PLECS and MATLAB Software by Designing USB 5W Charger

Zhi Ying Yeoh, Kim Soon Chong, Sew Sun Tiang, Mohammad Arif Ilyas, Jia Wei Tan, Wei Kang Lai, Wei Hong Lim (UCSI University, Malaysia)

In power electronics, accurately assessing simulation tools is key for precise, reliable electronic system designs. This study compares MATLAB Simulink and PLECS in modeling a 5W USB charger's power characteristic. The charger, using an AC-DC full bridge rectifier and DC-DC flyback topology, delivers a stable 5VDC at 1A. The analysis focuses on power efficiency and thermal characteristics, incorporating real-life components for detailed insights. Results show PLECS, specialized in power electronics, surpasses MATLAB in accuracy and consistency. This research aids in understanding simulation tools' effectiveness, guiding engineers and researchers in power efficiency evaluations for electronic systems.

OS26-5 Design and Simulation and Performance of Grid Connected Photovoltaic System for Small, Tall Building in Malaysia

Wei Kang Lai, Kim Soon Chong, Sew Sun Tiang, Mohammad Arif Ilyas, Jia Wei Tan, Zhi Ying Yeoh, Wei Hong Lim (UCSI University, Malaysia)

In Malaysia's rapidly urbanizing landscape, sustainable energy for small, tall buildings is increasingly vital. This study addresses this need through the design, simulation, and performance analysis of grid-connected photovoltaic systems tailored for these unique structures. Utilizing AutoCAD and PVSyst for design and simulation, the research details rooftop array dimensions, PV panel wiring, and system components aligned with MS 1837-2018 standards. Results indicate the designed system can save 526.70MWh annually, costing RM 339,306.98, with a 9.8% return on investment (ROI) and a 13-year breakeven point, emphasizing its sustainability and economic viability.
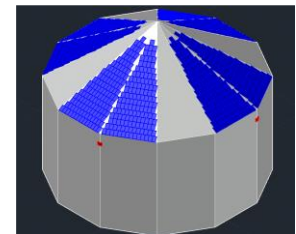
OS26-6 A Design of Dual-Band Coplanar Waveguide (CPW) Printed Antenna for 1.9 - 3.6GHz Applications

Jia Wei Tan, Sew Sun Tiang, Kim Soon Chong, Mohammad Arif Ilyas, Mastaneh Mokayef, Zhi Ying Yeoh, Wei Kang Lai, Wei Hong Lim (UCSI University, Malaysia)

A dual-band coplanar waveguide (CPW) printed antenna for IoT application is proposed, mounted on a low-profile RT/duroid 5880 substrate with dieletric constant of 2.2, loss tangent of 0.0009 and a standard height of 0.787mm. This design aims to cover the major frequency bands from LTE to Bluetooth/Wi-Fi band/WiMax/Zigbee, Extended IMT and 5G whereby the bandwidth range between 1.7 GHz to 3.6 GHz. The antenna is miniaturized through the CPW technique and has a rectangular size of 60 x 30 x $1.187mm^3$. The design and simulation of the result records a return loss of -25.12dB, peak gain of 4.23dBi, voltage standing wave ratio (VSWR) close to 1, omnidirectional radiation, and current distribution. Radiation efficiency reaches approximately 94%, with a total efficiency of 89.4% between 1.9GHz-3.6GHz.
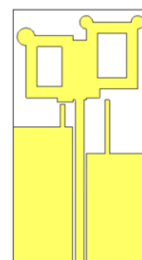
OS26-7 Optimized Convolutional Neural Network Towards Effective Wafer Defects Classification

Koon Hian Ang[1], Koon Meng Ang[1], Chin Hong Wong[2,3], Abhishek Sharma[4], Chun Kit Ang[1], Kim Soon Chong[1], Sew Sun Tiang[1], Wei Hong Lim[1]
([1]UCSI University, Malaysia) ([2]Fuzou University, China) ([3]Maynooth University, Ireland)
([4]Graphic Era Deemed to be University, India)

Semiconductor defect inspection is crucial for yield improvement but is hindered by manual inspection's subjectivity and error. This paper employs Convolutional Neural Networks (CNNs) for automated wafer defect classification, addressing the challenges of time-intensive training and complex hyperparameter tuning. We propose the Arithmetic Optimization Algorithm (AOA) to efficiently optimize CNN hyperparameters like momentum, initial learning rate, maximum epochs, and L2 regularization. Our method reduces the trial-and-error in hyperparameter tuning. Using the AOA-optimized ResNet-18 model, our simulations show superior performance in defect classification compared to the unoptimized model, demonstrating its effectiveness and practical potential.

OS26-8 Tackling Photovoltaic (PV) Estimation Challenges: An Innovative AOA Variant for Improved Accuracy and Robustness

Rayan Mohammed Noor Mohammed Bakhit[1], Abhishek Sharma[2], Tiong Hoo Lim[3], Chin Hong Wong[4,5], Kim Soon Chong[1], Li Pan[1], Sew Sun Tiang[1], Wei Hong Lim[1]
([1]UCSI University, Malaysia) ([2]Graphic Era Deemed to be University, India) ([3]Universiti Teknologi Brunei, Brunei Darussalam) ([4]Fuzou University, China) ([5]Maynooth University, Ireland)

Optimizing photovoltaic (PV) cell/module modeling is key to advancing solar power and achieving net zero carbon goals. Challenges in accurate PV parameter estimation arise from environmental variability, aging, and incomplete manufacturer data. Traditional Arithmetic Optimization Algorithm (AOA) often struggles with premature convergence due to imbalanced exploration and exploitation. This paper presents an enhanced AOA variant, incorporating chaotic maps and oppositional-based learning to better balance the optimization process. Our extensive simulations show that this improved AOA variant significantly enhances accuracy and robustness in PV cell/module parameter estimation compared to the conventional method.

OS26-9 Deep Learning in Manufacturing: A Focus on Welding Defect Classification with CNNs

Tin Chang Ting[1], Hameedur Rahman[2], Tiong Hoo Lim[3], Chin Hong Wong[4,5], Chun Kit Ang[1], Mohamed Khan Afthab Ahamed Khan[1], Sew Sun Tiang[1], Wei Hong Lim[1]
([1]UCSI University, Malaysia) ([2] Air University, Pakistan) ([3]Universiti Teknologi Brunei, Brunei Darussalam) ([4]Fuzou University, China) ([5]Maynooth University, Ireland)

Welding is integral to modern manufacturing, yet the complex process often leads to defects, impacting the quality of the final product. Recent advances in deep learning, particularly Convolutional Neural Networks (CNNs), have shown remarkable results in applications like defect recognition. This study evaluated AlexNet, ResNet-18, ResNet-50, ResNet-101, MobileNet-v2, ShuffleNet, and SqueezeNet for their effectiveness in identifying welding defects, using accuracy, precision, sensitivity, specificity, and F-score as metrics. The dataset covered defects like cracks, lack of penetration, porosity, and a no-defect class. Our analysis shows that most of these architectures deliver promising results in accuracy, sensitivity, specificity, precision, and F1-score, highlighting their potential in defect recognition.

OS26-10 Enhancing Global Optimization Performance of Arithmetic Optimization Algorithm with a Modified Population Initialization Scheme

Tin Chang Ting[1], Hameedur Rahman[2], Meng Choung Chiong[1], Mohamed Khan Afthab Ahamed Khan[1], Cik Suhana Hassan[1], Farah Adilah Binti Jamaludin[1], Sew Sun Tiang[1], Wei Hong Lim[1]
([1]UCSI University, Malaysia) ([2]Air University, Pakistan)

Arithmetic Optimization Algorithm (AOA) is widely used to solve global optimization problems. However, it often faces premature convergence challenges in complex optimization scenarios. A key factor affecting AOA's performance is the solution quality of the initial population. The conventional initialization scheme, despite its prevalence, lacks reliability in ensuring high-quality solutions due to inherent stochastic processes. To address this issue, we propose a modified initialization scheme that improves initial population quality by integrating chaotic maps and oppositional-based learning. Through extensive simulation studies, we demonstrate that the enhanced AOA, equipped with this new initialization scheme, exhibits superior performance in solving a range of benchmark functions with improved accuracy.

OS26-11 Enhancing Precision Object Detection and Identification for Autonomous Vehicles through YOLOv5 Refinement with YOLO-ALPHA

Guandong Li[1], Yanzhe Xie[1], Yuhao Lu[1], Jingzhen Fan[1], Yuankui Huang[1], Zongyan Wen[1], Wei Hong Lim[2], Chin Hong Wong[1]
([1]Maynooth University, Ireland) ([2]UCSI University, Malaysia)

Advancing swiftly in contemporary society, the rapid growth of autonomous driving technology suggests its potential adoption across continents. The realization of fully autonomous driving relies on proficiently detecting, classifying, and tracking road objects such as pedestrians and vehicles. This research employs the YOLOv5 neural network, enhancing it with YOLO-ALPHA. Modifications, encompassing freeze and attention mechanisms, serve to refine accuracy and expedite training. Furthermore, adjustments to the activation function aim to stabilize precision and recall. The integration of a FCN based on semantic segmentation theory contributes to improved accuracy in detecting road conditions during autonomous driving. Consequently, this enables the successful and highly accurate functionality of automatic identification.
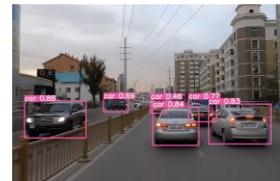
OS27 Post-narratological Approaches to Cognition in Humans and Robots (5)
Chair: Jumpei Ono (Aomori University, Japan)
Co-Chair: Hiroki Fxyma (Kobe University, Japan)

OS27-1 A Game Framework Based on the Disinformation Warfare in Russo-Ukrainian War

Jumpei Ono (Aomori University, Japan), Takashi Ogata (Iwate Prefectural University, Japan)

From 2022, when Russia began its invasion of Ukraine, to the present, the number of speeches and information directly/indirectly related to Russo-Ukrainian war has continued to increase in mass communication venues such as SNS and TV. This study proposes a game that simulates the spread of disinformation by focusing on the information published by Ukraine as a list of disinformation speakers, especially on the topics they talk about. The game is intended to provide players with psychological immunity against disinformation through the simulation of disinformation dissemination.
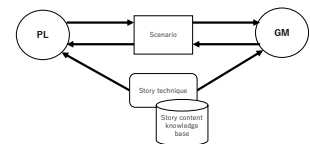
OS27-2 Comparative Analysis of Eye Tracking between Veteran and Novice during Radiological Interpretation

Yuka Naito, Jun Nakamura (Chuo University, Japan),
Yoshinobu Ishiwata (Yokohama City University, Japan)

With the development of AI, AI may substitute for physicians in radiological interpretations in the future. The purpose of this paper is to identify factors that could form the foundation for algorithms enabling AI to explore pathological findings. This study focuses on the "angles of gaze trajectory" and conducts a comparative analysis between veteran and novice. The results indicate commonalities and variances in the angles of gaze trajectory between veteran and novice, particularly around 10, 350, and 180 degrees. Both Veteran and Novice have a relatively higher frequency of gaze movement around 10 and 350 degrees compared to other angles. Veteran has a relatively higher frequency of gaze movement around 180 degrees compared to other angles, while novice has a relatively lower frequency around 180 degrees.
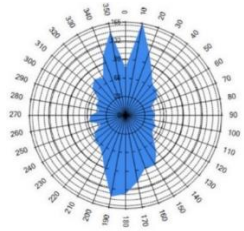
OS27-3 Development of Notification System to Prevent Working Productivity from Declining Caused by Increased Carbon Dioxide Concentration

Yohei Kamoda, Jun Nakamura (Chuo University, Japan)

In association with the change of work style to remote, the home environment should be considered in terms of carbon dioxide concentration (CDC) which has negative effects on the human body, such as less cognitive abilities. To challenge this problem, the author developed an alert system that supports remote workers to be notified of an increase in CDC, by combining Raspberry Pi, CDC sensor, and Slack. As a result, the developed alert system was able to support the user in keeping the CDC in the room below the set threshold. In addition, the system shows that a significant increase in CDC can be observed in the room when it comes to insufficient ventilation.

OS27-4 Visualization of the Skilled Physician's Gaze Characteristic during Diagnosis

Taiki Sugimoto, Jun Nakamura (Chuo University, Japan)
and Yoshinobu Ishiwata (Yokohama City University, Japan)

This study examines gaze movement differences in diagnosis between skilled and unskilled physicians, aiming to identify factors influencing diagnostic speed and enhance artificial intelligence's diagnostic capabilities. Results reveal that experienced physicians, on average, spent 61% less diagnostic time than beginners, covering 49% of the gaze distance on the X and Y-axes. (comparison of the cumulative distance of gaze on the screen). Despite increased movement on the Z-axis (comparison of the scrolling speed of the CT scan results), skilled physicians moved 2-3 times faster, effectively narrowing attention and identifying specific areas.

OS27-5 Comparative Analysis of Methods for Visualizing the Sensory Experience of Food and Beverages

Hiroki Fxyma (Kobe University, Japan)

This research explores and compares various methodologies employed in the visualization of taste experiences associated with food and beverages. The study delves into sensory analysis techniques, technological innovations, and artistic representations used to convey the complex and subjective nature of flavor. Through a comprehensive review of existing literature and practical experimentation, the paper evaluates the efficacy of different approaches in capturing and communicating the nuanced aspects of taste, aroma, and texture. The findings aim to contribute to the advancement of sensory science, culinary arts, and consumer research by providing insights into the strengths, limitations, and potential applications of diverse taste visualization methods. As the interdisciplinary field of gastronomy continues to evolve, this research serves as a valuable resource for professionals, researchers, and enthusiasts seeking to enhance our understanding and appreciation of the multisensory aspects of food and beverage consumption.

GS abstracts
GS1 Machine learning& Neural Network & Artificial Life (7)
Chair Hiroki Tamura (University of Miyazaki, Japan)

GS1-1 Deep Learning Based Prediction of Heat Transfer Coefficient Using Spectrogram Images from Boiling Sound

Fuga Mitsuyama, Ren Umeno, Tomohide Yabuki, Tohru Kamiya (Kyushu Institute of Technology, Japan)

Recently, cooling methods based on water boiling have attracted attention as a thermal solution for electronic equipment. In this situation, it is necessary to measure Heat Transfer Coefficient (HTC) to design more efficient cooling systems. In this paper, we propose a method to predict of the HTC from boiling sound data using deep learning techniques. Measuring HTC requires a variety of sensors and modifications to the equipment, which can be costly. Predicting HTC by sound alone can therefore reduce these costs. The accuracy improved by 1.12% compared to the conventional method through the development of Convolutional Neural Network (CNN) incorporate Convolutional Block Attention Module (CBAM).

GS1-2 A Study on Classification of Faulty Motor Sound Using Convolutional Neural Networks

Jamil Md Shafayet, Praveen Nuwantha Gunaratne, Hiroki Tamura (University of Miyazaki, Japan)

This paper is focused towards determining the use of audio classification methods to detect faulty motors by their sound, both with an abnormal noise and noise-free scenarios, so that the requirement of human inspection can be reduced in factories. Here, the soundwaves of motors, having both normal and faulty mechanical parts were acquired using an electronic stethoscope. The audio files were recorded in '.wav' format which later split into chunks. Noise reduction is widely used in audio classification tasks. Among various available methods, the Autoencoder was implemented for noise reduction. Our proposal method is to use various soundwave analysis methods for generating features such as in Mel-Frequency Cepstral Coefficient, Mel-spectrogram, and Spectrogram, and comparing the resulting accuracy. Convolutional neural networks are employed for the purpose of classification. Our proposed method resulted in approx. 97% accuracy in classifying 2 classes having 93 samples of motors and 2.23% loss in noise reduction with autoencoder.

GS1-3 Design and Development of a Flexible Active Ankle Joint Orthosis for Locomotion Assistance

Praveen Nuwantha Gunaratne, Hiroki Tamura (University of Miyazaki, Japan)

Active ankle joint orthosis is an artificial wearable device that is closely fitted to the human ankle and combines human intelligence and intentions with the powered robot joints. However, most of the existing designs have only focused on one degree of freedom (DOF) of the ankle joint: namely plantarflexion-dorsiflexion. Alternatively, a modular power-assist flexible active ankle orthosis is proposed, addressing anthropomorphic architecture of design, and supporting multiple DOFs of ankle for locomotion assistance of the physically weak individuals. The proposed ankle joint orthosis was designed with the aid of CAD tools followed by fabrication of a working model. Electrically powered systems with gear transmissions were adapted to support the principal motions during normal Gait. In parallel, the design of the control algorithm was carried out based on the EMG signals. The future advancements have been focused on developing a novel control method that provides sufficient flexibility to assist a wide variety of lower-limb motions.

GS1-4 A study on the Real-Time Biomechanical Analysis of Lamber Burden Utilizing Stereoscope Cameras

Taufik Hidayat Soesilo, Praveen Nuwantha Gunaratne, Hiroki Tamura (University of Miyazaki, Japan)

Lamber region is susceptible to strain and stress due to various physical activities and occupational task. To study about the lamber burden, researchers have developed many tools but most of existing design can only use static imaging or doesn't provide real-time update. Our proposed system records and examines the change of body posture in real time by utilizing the capabilities of stereoscope cameras. By using MediaPipe algorithms, 2D keypoints representing body joint can be extracted from images. Afterward, using the Direct Linear Transform (DLT) the corresponding 3D keypoints can be calculated using obtained 2D keypoints. With the 3D keypoints the body angles can be computed and used to calculate the wight of the lumbar burden using the digital human model software JACK-like calculation. Finally, our proposed system reached to its aim to study about lamber burden and adjust the person needs in real-time.

GS1-5 Verification of Determination Possibility using Convolutional Autoencoder for Machine Tool Abnormality Detection

Yuta Sumoto, Praveen Nuwantha, Hiroki Tamura (University of Miyazaki, Japan)

The purpose of this research is to clarify the cause of failure and to improve the accuracy of abnormality detection by predicting the noise added to the machine tool from the results of the convolutional autoencoder. Data obtained from an acceleration sensor mounted on the machine tool are reconstructed using a convolutional autoencoder, and the average absolute error is calculated. The maximum value of the average absolute error of the training data is used as the threshold value for abnormality detection. Multiple simulated data with different amplitude values based on a composite sine wave with a specific frequency, white noise, and random numbers within a specified amplitude value were verified. In this paper, the characteristics of each type of noise and the parameters of the optimal model were examined from error rate and error distribution.
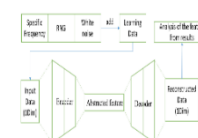
GS1-6 Basic Research for High-speed Heart Sound Determination using AI

Riku Nakashima, Praveen Nuwantha Gunaratne, Hiroki Tamura (University of Miyazaki, Japan)

In recent years, many researchers focus on enhancing the current situation of the home healthcare system. In this paper, we developed a system that uses AI to quickly determine whether auscultation sounds are normal or not. In our analysis, heart sounds were imaged, and their abnormalities were identified using machine learning tools. The proposal approach uses a YOLO_v7 model based anomaly detection and the Wavelet Transform was employed to analyze the acquired heart sound data. Our proposal system resulted with an 89% accuracy rate on the basis of 5 seconds of heartbeat data, in various recording environments.
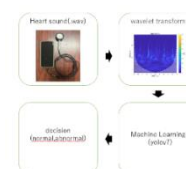
GS1-7 Prediction of High-Energy Electron Flux at Geosynchronous Orbit using a neural network technique

Ami Iwabu, Kentaro Kitamura (Kyusyu Institute of Technology, Japan)

The radiation belt, where the high-energy particles are predominant in near earth space from the low earth orbit (LEO) to the geostationary orbit (GEO), sometimes causes satellite malfunction. Therefore, the objective of this study is to predict the high-energy electron flux at GEO with the energy above 2 MeV after 24 hours with higher accuracy for the safety satellite operation in terms of the space weather science. In this study the various kinds of solar wind data from satellite observations and ground geomagnetic observation data in 1999 were used for the Recurrent Neural Network (RNN). Prediction results were evaluated by the prediction efficiency, which is derived from both predicted and actual variation data. As a result, the prediction using combined data of solar wind and geomagnetic data shows highest prediction efficiency of 0.72.

GS1-8 Image Gradient-based Monocular Visual-Inertial Odometry

Tae Ihn Kim[1]. Jae Hyung Jung[2], Chan Gook Park[2]
([1]Hyundai Motor Company, Republic of Korea), ([2]Seoul National University, Republic of Korea)

This paper presents an image gradient-based monocular visual-inertial odometry (VIO) algorithm, using image gradient measurements, robust to illumination change. Our proposed algorithm follows the multi-state constraint Kalman filter (MSCKF) framework, a sliding windowed extended Kalman filter (EKF). We expand the measurements from the reprojected feature locations on the image coordinates to the corresponding image gradients in the measurement selection. The iterated EKF and low-pass pyramid are adapted to reduce the linearization error in the MSCKF measurement update process. We verify that our proposed algorithm outperforms both conventional indirect and direct MSCKF-based VIO algorithms by evaluating the pose estimation performance using a real-world dataset, including illumination change.
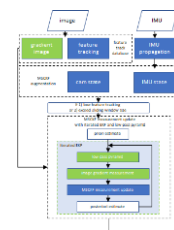
GS2 Image Processing I (5)
Chair Yui Tanjo (Kyushu Institute of Technology, Japan)

GS2-1 A Method for Embedding Multiple Photographic Images in a Photographic Image

Naoki Kouno, Kanya Goto, Toru Hiraoka (University of Nagasaki, Japan)

A method has been proposed for embedding another photographic image (image B) in a photographic image (image A).   An image (image A') is generated by embedding information of image B in image A, and then an image (image B') is restored by extracting information from image A'. The conventional methods can only embed one photographic image in one photographic image. Therefore, we extend the conventional method and propose a method for embedding multiple photographic images in one photographic image. In our method, as more images B are embedding in image A, the image quality of image A' deteriorates, but the image quality of images B' dose not deteriorate. To verify the effectiveness of our method, experiments using various photographic images were performed.

GS2-2 Generation of Flowing-Line Images Using Vertical and Horizontal Smoothing Filters

Karin Kuroki, Toru Hiraoka (University of Nagasaki, Japan)

We propose a non-photorealistic rendering method for automatically generating flowing-line images from photographic images. Flowing-line images consist of unidirectional lines with smooth curves. Our method is executed by an iterative calculation using vertical and horizontal smoothing filters. To verify the effectiveness of our method, we conducted an experiment using various photographic images to confirm that flowing-line patterns can be generated on the entire image. Additionally, we conducted an experiment to visually examine how flowing-line patterns generated by changing the values of the parameters in our method change.

GS2-3 Human Motion Recognition from Multiple Directions and Its Gait Cycles Analysis

Miki Ooba, Yui Tanjo (Kyushu Institute of Technology, Japan)

It is crucial for individuals to keep walking and stay healthy to prevent receiving nursing care. This paper proposes a method of recognizing walk motions and analyzing the gait cycle of a human focusing on his/her posture. We use 43 structural features defined from human joint coordinates obtained using OpenPose and 18 figural features from human domain images and their difference images. The feature vector containing these 61 features is used for the recognition of walk motion by Random Forest. In the experiment, we applied the method to recognizing six types of motions and analyzed the walk gait cycles of five persons, and obtained satisfactory results.

GS2-4 A Method of Improving the QOL of the People with Visual Impairment by MY VISION

Shun Kitazumi, Yui Tanjo (Kyushu Institute of Technology, Japan)

Visually impaired people face several difficulties in indoor activities, such as spending excessive time in locating objects. This paper proposes a method for assisting object acquisition by detecting desired objects and guiding users to them. The method requests a user to express the object he/she wants to acquire verbally and utilizes speech recognition to detect the specified object. Subsequently, the system guides in voice the user's hand to the location of the desired object. The performance of the method is experimentally shown. The method contributes to enhancing the comfort of indoor activities of visually impaired and, in this way, improves their quality of life.

GS2-5 Human Behavior Segmentation and Recognition Using a Single-camera

Jing Cao, Yui Tanjo (Kyushu Institute of Technology, Japan)

In recent years, elderly people living alone account for a large proportion of the elderly population, and the issue of safety has also been a matter of great concern for the public. Considering the importance of monitoring the behavior and activities of the elderly and detecting abnormal movements, this paper proposes a method that can segment human behavior into each action and identify the action from the videos taken by a single camera. It uses features that can represent the shape of the human area in the depth direction, as well as the features such as motion direction and speed. The performance and effectiveness of the method are verified by experiments.

GS3 Image Processing II (3)
Chair Seiji Ishikawa (Kyushu Institute of Technology, Japan)

GS3-1 Online Classroom Student Engagement Analysis using Enhanced YOLOv5

Shuai Wang, Abdul Samad Shibghatullah (UCSI University, Malaysia)

The rise of online education has highlighted the urgency of addressing cyberbullying in virtual classrooms. This paper presents an innovative method for early cyberbullying detection through the analysis of students' engagement and emotional responses in online classrooms. The SFER-YOLOv5 model, a fusion of Student Facial Expression Recognition and an enhanced YOLOv5 object detection model, introduces key optimizations, including Soft NMS for Non-Maximum Suppression, integration of a Channel Attention (CA) module, and use of Enhanced Intersection over Union (EIOU) as the bounding box regression loss function. This approach proactively identifies reduced engagement and emotional irregularities, providing a framework to mitigate cyberbullying in online classrooms.
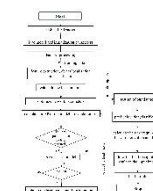
GS3-2 A Method of Recognizing Body Movements Based on a Self-viewpoint Video

Iichirou Moribe, Yui Tanjo (Kyushu Institute of Technology, Japan)

The most critical human sensory function resides in vision. This study focuses on utilizing visual information, specifically self-perspective footage, to identify individual movements. Existing researches require third-party filming to recognize body movements and states. The proposed method, on the other hand, simply attaches a camera to the human head and enables the recognition of the subject's actions. Consequently, it becomes easier to monitor daily movements of a human and gather his/her data on   body kinetics. This approach would be beneficial in scenarios involving individuals engaging in risky behavior or, during a certain emergency, providing valuable assistance.

GS3-3 Supporting Safe Walk of a Visually Impaired Person at a Station Platform based on MY VISION

Shintaro Yamada, Yui Tanjo, Seiji Ishikawa (Kyushu Institute of Technology, Japan)

When individuals with visual impairment go out, public transportation such as trains and buses is commonly used. However, many of them experience accidents, such as falling from train platforms or tripping due to unexpected contact with other passengers. To solve this problem, we propose method using the MY VISION system which detects the obstacle that may pose risks to individuals with visual impairment. The proposed method detects obstacles such as passengers pillars and platform edges at train stations. We employ an RGB-D camera for capturing frontal view of a user, use depth images to detect the edge of obstacles and level differences, and give warning to the visually impaired user based on the distance between him/her and the detected obstacle. Experimental results show satisfactory performance of the method.

GS4 Robotics (6)
Chair Jiwu Wang (Beijing Jiaotong University, China)

GS4-1 Development and evaluation of a learning support robot for vector learning

Kosei Machida, Shinichi Imai (Tokyo Gakugei University, Japan)

Vectors studied in high school are a new concept that differs from those studied in previous scalars. Therefore, it has been pointed out that difficulties arise in the conceptual formation of vectors. This study aims to develop a learning support robot that can visualize vector information by robot movements, and to acquire vector concepts for students through classroom practice using the robot. The robot operates on a piece of imitation paper, marking the start and end points and connecting them with arrows to visualize the vectors. In a class using the robot, the student predicts the sum of the vectors, and the robot confirms it. A questionnaire was administered before and after the class. The results revealed significantly higher mean values when comparing the pre- and post-assessments.

GS4-2 Feasibility Study on Methods to Measure the Strain on Young Children's Bodies.

Sachiko Kido, Praveen Nuwantha Gunaratne, Hiroki Tamura (University of Miyazaki, Japan)

This paper investigates the effective methods of measuring the burden on the infant's body. Initially, the "the burden on the body and the lumbar spine burden are the same" was defined. The calculations and comparisons were carried out based on the two methods: the AnyBody Modeling System and the smartphone application Yo-bukun (a lumbar spine burden measurement application). Videos of a five-year-old girl squatting and balancing were filmed, and the lumbar burden was calculated using the AnyBody Modelling System. In addition, while the video was being filmed, an iPhone running the Yo-bukun lumbar burden measurement application was placed close to the heart in order to measure lumbar burden and the amount of burden was calculated. The results were compared with the respective average values and validated.

GS4-3 An Integration of Contact Force Models with Multibody Dynamics Analyses for Human Joint Mechanisms and Effects of Viscoelastic Ground Contact

Shintaro Kasai[1], Dondogjamts Batbaatar[2], Hiroaki Wagatsuma[1]
([1]Kyushu Institute of Technology, Japan) ([2]Mongolian University of Science and Technology, Mongolia)

In human movement and rehabilitation analyses, human joint dynamics is a key to consider the incorporation of spring-damper components, flexible bodies and contact forces analytically. In the present study, an analytical method for human gaits were introduced to integrate those essential elements, and viscoelastic properties of musculoskeletal system were modeled with the absolute nodal coordination formula (ANCF) method representing flexible body motions. A contact force model simulates interactions between different body segments and the environment. The proposed system is applied to a slider crank mechanism, demonstrating its capabilities in human joint motion analysis using integrated dynamic model within the framework of multibody dynamics (MBD), which realizes dynamic/inverse dynamics for human biomechanics.

GS4-4 Haptic Sensation Enhancement via the Stochastic Resonance Effect and Its Application to Haptic Feedback for Myoelectric Prosthetic Hands

Yoshitaka Mizumoto, Taro Shibanoki (Okayama University, Japan)

In this paper, we propose a highly realistic haptic feedback method for myoelectric prosthetic hands. Although several haptic feedback methods have been studied, this study attempts to create realistic sensory feedback through prosthetic hands by not only using feedback methods but also by improving the haptic sensation based on the stochastic resonance effect. In the experiment, contact information obtained from a microphone attached to the fingertip of a prosthetic hand was transferred through a vibrotactile stimulator near the elbow socket, and white noise vibration was also applied near the elbow socket to verify the improvement of tactile sensitivity of the fingertips. The results demonstrated the possibility of transmitting tactile information for myoelectric prostheses through sensory enhancement of the user.
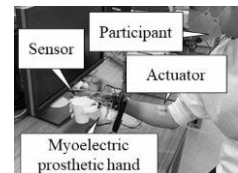
GS4-5 PID Parameter Tuning of a Low-Cost DC Motor Speed Control for Mobile Robot Application

Munkh-Erdene Ayurzana, Erkhembayar Gankhuyag, Naranbaatar Erdenesuren, Dondogjamts Batbaata
(Mongolian University of Science and Technology, Mongolia)

Precise control of DC motors is vital for robotics, industrial automation, and mechatronics. Traditional PID (Proportional-Integral-Derivative) control methods, while widely used, often require offline parameter tuning, which can be time-consuming and suboptimal for real-time applications. This paper proposes a GUI (Graphical User Interface) based approach for PID tuning of a DC motor model. The proposed method utilizes a MATLAB-based parameter estimation model with GUI to continuously monitor and update PID controller parameters based on real-time data from the Arduino-controlled DC motor setup.

GS4-6 Reinforcement Learning DDPG Algorithm Based Wheeled Mobility Aid Robot Control Methods

Junkai Li, Mohd Rizon Mohamad Juhari, Tiang Sew Sun (UCSI University, Malaysia)

Ensuring stability and precise trajectory tracking is crucial when employing wheeled walker robots to enhance mobility for individuals with limited walking ability. We propose a novel trajectory tracking method for wheeled walking robots by combining the Deep Deterministic Policy Gradient (DDPG) algorithm in reinforcement learning with a Proportional Integral Differential (PID) controller. We verify the effectiveness of the research scheme and control strategy through joint simulation experiments. The results demonstrate that the DDPG-based PID controller can automatically adjust parameters to ensure trajectory accuracy and exhibits a strong anti-interference capability.
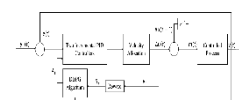
GS5 Applications I (4)
Chair Jiwu Wang (Beijing Jiaotong University, China)

GS5-1 Unsupervised image registration based on Residual-connected DRMINE for diagnostic metastatic bone tumors

Shogo Baba[1], Tohru Kamiya[1], Takashi Terasawa[2], Takatoshi Aoki[2]
([1]Kyushu Institute of Technology, Japan), ([2]University of Occupational and Environmental Health, Japan)

Computed tomography (CT) scans produce more than 100 images in an examination, which imposes a significant burden on radiologists and entails a potential risk of misdiagnosis. In this study, we focus on the preliminary stage of Computer-Aided Diagnosis (CAD) development specialized in bone metastasis extraction, with a particular emphasis on rigid registration. We propose a novel rigid registration technique by augmenting DRMINE, which estimates mutual information using neural networks, with skip connections and normalization. The proposed method was influenced by the capture area, but it indicated the potential to provide stable registration as the standard deviation decreased for all Full Width at Half Maximum (FWHM).
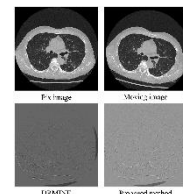
GS5-2 Developing a smart Belt for Monitoring Elderly Activities Based on Multi-Modal Sensors Integration and Internet of Things

Abdul Jalil, Pujianti Wahyuningsih, Najirah Umar, Muhammad Risal, Suwatri Jura, A. Edeth Fuari Anatasya
(Universitas Handayani Makassar, Indonesia)

This study aims to develop a smart belt for monitoring elderly activities at home based on the integration of multi-modal sensors and the Internet of Things (IoT). The multi-modal sensors used in this study to collect the elderly information are the IMU sensor, vibration sensor, push-button, and ESP32 to process the data. Furthermore, the IoT framework used to transmit the elderly information data from the smart belt to the family's smartphone is a Blynk. In this study, the smart belt can monitor the elderly activities when walking, sitting down, lying down, and sleeping, after that, can give a security warning when the smart belt detects the elderly doing abnormal activities. This study's results show that the smart belt is effectively used to monitor elderly activities to help families when taking care of the elderly at home.
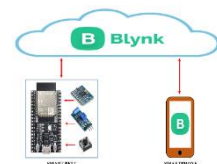
GS5-3 A penalized motion detection model for extracting ionospheric echoes from low signal-to-noise ratio Ionogram video images

Yuu Hiroshige[1], Akiko Fujimoto[1], Akihiro Ikeda[2], Shuji Abe[3], Akimasa Yoshikawa[3]
([1]Kyushu Institute of Technology, Japan) ([2]National Institute of Technology, Kagoshima College, Japan)
([3]Kyushu University, Japan)

Measuring the altitude distribution of electron density in the upper atmosphere, known as the ionosphere, using High-Frequency radio wave reflections often causes the low signal-to-noise ratio of ionospheric echoes due to radio frequency interference. We propose a model for converting low-signal-to-noise-ratio ionospheric echo video images (Ionogram) into noise-reduced images using image processing techniques, for tracing the ionospheric echoes from Ionogram. The proposed method consists of three processing parts: noise removal optimized for individual Ionogram images, extraction of ionospheric echoes by penalized background subtraction technique, and fine-tuning of ionospheric echo signals using a minimum spanning tree algorithm. The proposed model successfully reproduces fine Ionograms with 98% recall and 99% precision.

GS5-4 Verification Experiments on the Lower Back Burden caused by Posture and Environment during Lifting Operations

Tomoka Kimura, Yutaro Fujino, Sachiko Kido, Praveen Nuwantha Gunaratne, Hiroki Tamura
(University of Miyazaki, Japan)

One of the measures to prevent back pain is the use of appropriate posture corrections. In generally, the Squat method (method of keeping the knees bent and the waist as straight as possible) is recommended over the Stoop method (method of keeping the knees straight and the waist bent). The results of previous studies using acceleration and muscle potentials have shown that the lifting with knees in a kneeling position reduces the amount of load born by the lower back. However, few studies discuss the case where there is an obstacle between the subject and the object to be lifted and further, the scenarios where the object is away from the body's center of gravity. Therefore, this research focus on analyzing two types of movements using the AnyBody musculoskeletal mechanics analysis software and the Delsys surface EMG and verifying the amount of lower back burden when there is an obstacle between the lifting object and the subject. This paper presents the verification results.
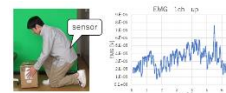
GS6 Applications II (4)
Chair Masayuki Fujiwara (Kyushu Institute of Technology, Japan)

GS6-1 Development of Smartphone Application for Calculating the Low Back Pain Risk

Seigo Imura, Praveen Nuwantha, Hiroki Tamura (University of Miyazaki, Japan)

This study primarily relies on a smartphone application, developed within our research institute known as Yo-bukun, for the real-time estimation of lumbar burden. The Yo-bukun is capable of estimating the lumbar burden of a subject (the user) by placing an app. installed smartphone in the subject's chest pocket, while the subject is lifting/ relocating an object. The subject's movements are assessed through sensors embedded in the smartphone and certain aspects of their physical information initially fed into the app. are used in the estimation formula for determining lumbar burden. In the current scenario, Yo-bukun lacks the capability to ascertain whether the user is holding an object or not; consequently, it can only estimate the lumbar burden for limited cases of a subject holding an object. To address such limitations, the proposed system integrates voice recognition to facilitate lumbar burden estimation, considering the presence or absence of an object. Further, it was incorporated with the capability to recalculate lumbar burden after measurements, enabling prospective studies.

GS6-2 A Computational Approach for Global Trade Analysis Sensitive to Free Trade Agreement Circumstances: A Case Study Focusing on the Great Mekong Subregion

Ahmad Altaweel[1], Bo-Young Lee[2], Masayuki Fujiwara[1], Jang-Sok Yoon[2], Hiroaki Wagatsuma[1]
([1]Kyushu Institute of Technology, Japan) ([2]Logistics Revolution Korea, Korea)

In the Global Trade Analysis Project (GTAP), GDP and economic statistical indices can be analyzed for forecasting future trends of them in multiple countries by using the GTAP database and GEMPACK utilities, which realize the numerical calculation based on the concept of Computable General Equilibrium (CGE) models. Even by such well-prepared tools with the official database, an appropriate forecasting is still difficult due to the sensitivity to Free Trade Agreement (FTA) circumstances. FTA scenarios with a uniform tariff reduction can be assumed in general, while an abrupt and unreasonable change may occur in the simulation depending on the network structure of trading countries and upper and lower bounds of tariffs in the time course. In the present study, we focused on the Great Mekong subregion (GMS) and explored possible methods to calculate substantially.
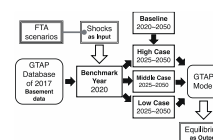
GS6-3 Trigger circuit design and system integration for simultaneous measurement of human EEG, motion, and gaze

Masayuki Fujiwara[1], Phan Hoang Huu Duc[1], Laurent Bougrain[2], Patrick Hénaff[2], Hiroaki Wagatsuma[1]
([1]Kyushu Institute of Technology, Japan) ([2]Université de Lorraine, France)

The simultaneous measurement of human EEG, motion, and gaze has the potential to lead to the discovery of new scientific insights. In order to achieve these simultaneous measurements, it is necessary to manage triggers and time information between measurement devices, as well as to correct time offsets. However, the management of accurate triggers and time information requires the design of a dedicated circuit board and the integration of TTL signal voltage information. In this study, we report on the fabrication of a trigger circuit and an experimental system using it to solve these problems. We created a home-made trigger circuit board for voltage integration and combined it with a commercially available microcomputer to realize an integrated trigger circuit and measurement system.
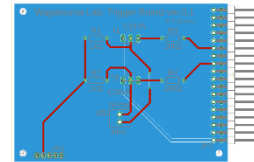
GS6-4 Terminal Synergetic Controller for Car's Active Suspension System Using Dragonfly Algorithm

Tinnakorn Kumsaen[1], Sorn Simatrang[2], Arsit Boonyaprapasorn[3], Thunyaseth Sethaput[4]
([1]Khon Kaen University, Thailand) ([2]Nacres Co., Ltd, Thailand) ([3]Chulachomklao Royal Military Academy, Thailand) ([4]Thammasat University, Thailand)

This research introduces a terminal synergetic controller (TSC) designed for the active suspension system of automobiles through the implementation of the dragonfly algorithm (DA). The proposed controller aims to enhance the dynamic performance of a car's suspension using the DA in tuning the system parameters. The stability of the designed controller is proved through the application of Lyapunov stability theory. Through iterative optimization processes, the TSC approach seeks to achieve an optimal balance between ride comfort and vehicle handling. The simulation results demonstrate that the proposed controller enhances convergence properties and alleviates the presence of chattering. The results indicate that the proposed approach with the optimal parameters provided insights into its potential application in improving the overall suspension system.

GS7 Applications III (6)
Chair Kasthuri Subaramaniam (UCSI University, Malaysia)

GS7-1 Rehabilitating Flood-Damaged Cars for Sustainable Car Rental Services: A Web-Based System

Pon Xiao Qi, Abdul Samad Shibghatullah, Kasthuri Subaramaniam (UCSI University, Malaysia)

The web-based system aims to redefine transportation norms by unlocking the potential of flood-damaged cars. It seeks to reutilize these cars, providing affordable and sustainable transportation options while minimizing landfill waste. The system development approach that the researcher used for this system is Rapid Application Development (RAD) Model. The reason that the researcher chose this methodology is because it enables her to develop the system under limited time while ensuring the quality of the system. There is four stages in Rapid Application Development (RAD) methodology life cycle, which is requirement planning phase, user design phase, rapid construction phase and cutover phase.
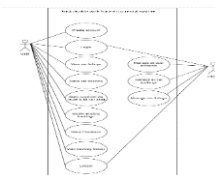
GS7-2 Optimizing E-Invoicing Rollout: Adaptive E-Invoicing Rollout (AER) Framework for Navigating Malaysia's Digital Transformation

Koh Chee Hong, Abdul Samad Shibghatulla (UCSI University, Malaysia)

This study presents the Adaptive E-Invoicing Rollout (AER) Framework, developed to guide Malaysia's impending e-Invoicing mandate. It is crafted through a meticulous literature review and expert consensus, highlighting crucial variables such as integration capability, adoption rate, readiness, and compliance. These elements are central to the predictive model, designed to forecast and facilitate successful e-Invoicing implementation. The findings from our extensive review and model application confirm the framework's effectiveness in diverse business environments, demonstrating its adaptability and strategic value. The AER Framework significantly contributes to digital finance strategies, providing an innovative, empirically based tool for not only Malaysia but also other regions in the globe pursuing digital economic transformations.
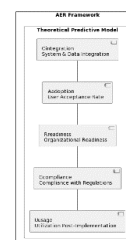
GS7-3 App Alert System for Smart Phones

Chee Kin Hoe, Kasthuri Subaramaniam, Abdul Samad Shibghatullah (UCSI University, Malaysia)

¥This development is to develop a porotype mobile app that notifies the user to take their phones after shutting down the car. This development is being done to gather feedback from participants who frequently forget to take their phones such as dementia sufferers, careless youngsters and users who park in a dark space. As for the results of the survey, it was extensively researched and analyzed to find out the problems and requirements of users for the app. For the methodology, after extensive consideration the researcher decided to use Modified Waterfall model a derivative from the famous waterfall model. The main factors that influenced the decision is the suitability of the methodology, the confidence of the researcher to use the methodology and the time constraint to develop the app. There are 6 phases in the modified waterfall model which are Requirement Analysis, System Design, Implementation, Testing, Deployment and Maintenance. After developing the app, the user acceptance test was done to see whether how the general public would accept the app. For the result, the app was accepted by most of the participants and they would like the researcher to do further development to improve and maintain the app. Based on the results of the development the researcher can conclude that this app can benefit users in helping them to remember to take their phones from their car. Rather than using the old fashion way which is remember to take the phone. With this app, user can rest assured as they will be always be reminded to take their phones from their car.
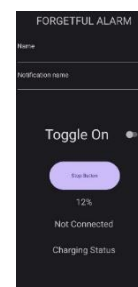
GS7-4 Developing Hand Gesture Recognition System in Interpreting American Sign Language

Kong Seh Chong, Kasthuri Subaramaniam, Ismail Ahmed Al-Qasem Al-H (UCSI University, Malaysia)

This research is going to study and gather more information on the hand gesture recognition system in order to bridge the communication between non-signers and signers by implementing the sign language into the system. The system may help people with physical disabilities such as deaf, mute and etc, to have an efficient and direct communication with others without the needs of interpreters. Through the studying of the hand gesture recognition system, there are two main frameworks which are OpenCV and MediaPipe that hold significant value in the system as it is mostly used in computer vision related problem. With this system developed, it can help many signers to avoid the problem of indirect interaction and also enable those who want to learn sign language by practicing hand gesture through the system. The system development approach that the researcher used for this system is Agile Methodology. The reason that the researcher chooses this methodology is because the development of this system mostly deals with the datasets as well as the accuracy of the detection and recognition of hand gestures. Hence, an ability to adapt to changes of agile methodology is needed if there are any problems happen in the system. There are seven of the stages in Agile Methodology life cycle, which are planning phase, designing phase, implementation phase, testing phase, deployment phase, review phase and launch phase. All these seven phases of the Agile methodology will be having some deliverables as their output.
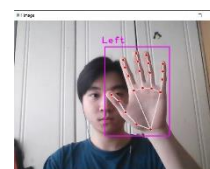
GS7-5 Miniature Enterprise Resource Planning

Adim Khalid Aldireejah, Kasthuri Subaramaniam, Ghassan Saleh（UCSI University, Malaysia）

Organizations aim for maximum efficiency, acquiring a competitive edge, and quickly completing their goals in today's competitive business environment. The achievement of these business objectives depends heavily on enterprise resource planning (ERP). ERP is a software program that combines several elements needed for a company's administration and operation, spanning several sectors like supply chain management, manufacturing, and finance. ERP makes work easier, saves time and money, and gives managers operational insights by automating corporate activities within a single platform. As a result, ERP has become essential for businesses of all sizes that want to advance and maintain their position as industry leaders. This research proposal intends to analyze the key traits, implementation concerns, and challenges associated with ERP systems while also outlining potential fixes. By creating a convenient and affordable mini-ERP solution, the initiative further intends to increase the accessibility of ERP systems to small and medium-sized enterprises (SMEs). The study of current ERP systems, their development, and their advantages and disadvantages will form the basis of the research. It will also investigate the difficulties SMEs have in using common ERP solutions.

GS7-6 Developing Cloud-based Sportswear Website

Lim Wei Yee, Kasthuri Subaramaniam, Raenu Kolandaisamy (UCSI University, Malaysia)

This research is going to develop a cloud-based sportswear website that can handle high volumes of traffic during events or seasonal sales, assist customers in selecting the right size for their sportswear, the interface should have options, size charts, fitting instructions, and recommendations. A study on the existing system will be conducted, design, develop and evaluate on the sportswear website also will be carried out. The system development approach used on this website is Rapid Application Development (RAD) model. The reason to choose RAD is because this method is time efficient, flexibility and adaptability. With RAD, functional software is delivered quickly. We can quickly build prototypes, iterate on them, and gather feedback from stakeholders with a website project. There are four of the steps in Rapid Application Development (RAD) methodology life cycle, which is define the requirements, prototypes, rapid construction and feedback gathering, and cutover.

AUTHORS INDEX

Notation of session name

PS: Plenary Session IS: Invited Session, OS: Organized Session, GS: General Session

Note: 33/90 = (page no. in Technical Paper Index) / (page no. in Abstracts)

Cao	Jing	GS2-5	23/106			OS23-2	24/90
Celiker	Baris	OS10-1	28/66			OS23-3	24/91
Chang	Heng-En	OS5-3	39/59			OS23-4	24/91
Chang	Meng-Wei	OS1-4	21/52			OS23-5	24/91
Chang	Yun-Hao	OS1-2	21/51	Dhanush	R.	OS18-2	40/78
		OS1-4	21/51	Dobrokvashina	Alexandra	OS10-6	28/67
Chang	Zhencheng	OS4-8	44/57	Du	Jiacheng	OS26-2	45/97
		OS4-9	44/58	Duan	Suqing	OS13-3	26/71
Chebotareva	Elvira	OS11-3	47/68			OS13-6	26/72
Chen	Chih-Hung	OS1-6	21/52	Dudorov	Evgeny	OS5-6	39/60
Chen	Nai-Yu	OS1-1	21/51	Duvacher	Maxime	OS6-3	43/61
		OS1-5	21/52				
Chen	Yao	OS26-2	45/97	[E]			
Chen	Yi-Wei	OS5-1	38/58	Edgar A.	Martínez-	OS5-5	39/59
Chen	Yu-Le	OS5-2	38/58		García		
Cheng	Jiale	OS13-1	26/71			OS11-1	47/68
Chiong	Meng Choung	OS26-10	46/100	Elshaikh	Mohamed	OS21-2	41/87
Chong	Kim Soon	OS6-2	43/60	Elamvazuthi	Irraivan	OS6-4	43/61
		OS26-3	45/97	Era	Junya	OS20-10	30/84
		OS26-4	45/98	Erdenesuren	Naranbaatar	GS4-5	24/108
		OS26-5	45/98	Eremin	Alexander	OS10-5	28/67
		OS26-6	45/98	Eryomin	Alexander	OS5-6	39/60
		OS26-7	46/99				
		OS26-8	46/99	[F]			
Chong	Kong Seh	GS7-4	49/112	Fadilla	Reno	OS14-3	36/74
Chong Jun Lynn	Esther	OS18-7	40/79	Faizullin	Ramil	OS5-5	39/59
				Fan	Jingzhen	OS26-11	46/100
Chou	Pei-Wen	OS1-1	21/51	Fan	Wenzhuo	OS26-1	45/97
		OS1-5	21/52	Fu	Wenqi	OS13-5	26/72
Chuang	Li-Min	OS1-6	21/52	Fuengfusin	Ninnart	OS15-1	22/74
				Fujimoto	Akiko	GS5-3	37/109
[D]				Fujinaga	Takuya	OS20-14	31/86
D. Ahmad	Aeryn	OS21-3	42/87			OS20-15	31/86
Dai	Fengzhi	OS2-1	25/52	Fujino	Kazuma	OS12-1	34/70
		OS2-3	25/53	Fujino	Yutaro	GS5-4	37/110
Daimaru	Yoshiyuki	OS20-9	30/84	Fujiwara	Masayuki	GS6-2	42/110
Darusman	Abu Hassan	OS6-4	43/61			GS6-3	42/111
Desa	Hazry	PS2	20/50	Furukawa	Tomoyasu	OS20-10	30/84
		OS23-1	24/90			OS20-11	30/85

115

Fxyma	Hiroki	OS27-5	33/102			OS25-9	48/96
						OS26-10	46/100
[G]				Hasebe	Kodai	OS22-6	38/89
Gamberov	Timur	OS10-3	28/66			OS22-7	38/89
		OS11-4	47/69			OS22-8	38/90
Gankhuyag	Erkhembayar	GS4-5	24/108	Hattori	Shimon	OS17-2	32/77
Gao	Yuxuan	OS24-3	20/92			OS17-3	32/77
Gong	Haoran	OS3-8	27/55	Hattori	Mai	OS17-5	32/78
		OS3-9	28/55	Hattori	Tetsuo	OS17-2	32/77
Gong	Ying	OS4-3	44/56			OS17-3	32/77
Gonzales	Andrea Tantay	OS18-1	40/78			OS17-4	32/77
Goto	Kanya	GS2-1	23/105			OS17-5	32/78
Gunaratne	Praveen	GS1-2	34/103	Hayashi	Eiji	OS16-1	39/75
		GS1-3	34/103			OS16-2	39/76
		GS1-4	35/104			OS16-3	39/76
		GS1-5	35/104			OS16-4	40/76
		GS1-6	35/104			OS20-12	30/85
		GS4-2	23/107	Hayashi	Takamasa	OS16-2	39/76
		GS5-4	37/110			OS16-4	40/76
		GS6-1	42/110	Hayashi	Toshihiro	OS17-5	32/78
Guo	Jr-Hung	OS5-1	38/58	He	Hao	OS2-3	25/53
						OS3-3	27/54
[H]						OS3-4	27/54
Hamakawa	Fumito	OS22-5	37/89	Hénaff	Patrick	GS6-3	42/111
		OS22-6	38/89	Hidayab	Muzaiyanah	OS6-2	43/60
		OS22-7	38/89	Hieda	Hiroo	OS12-1	34/70
		OS22-8	38/90	Hiroshige	Yuu	GS5-3	37/109
Hamasuna	Ota	OS22-5	37/89	Hiramitsu	Tatsuhiro	OS14-3	36/74
		OS22-6	38/89	Hiraoka	Toru	GS2-1	23/105
Haque	Ashraful	OS25-8	48/96			GS2-2	23/105
Harada	Kensuke	OS14-1	36/73	Hoe	Chee Kin	GS7-3	49/112
		OS14-2	36/73	Hoh	Ian	OS25-6	48/95
Hassan	Cik Suhana	OS7-1	31/62	Hong	Koh Chee	GS7-2	49/112
		OS7-2	31/62	Hong	L. Y.	OS23-3	24/91
		OS7-3	31/62	Hori	Yoshiki	OS16-1	39/75
		OS7-4	31/63	Hsia	Kuo-Hsien	OS5-2	38/58
		OS25-1	47/93			OS11-6	47/69
		OS25-4	48/94	Hu	Mengyuan	OS4-5	44/57
		OS25-5	48/95	Huang	Mei	OS8-5	36/64

116

Huang	Jiawei	OS20-4	29/83	Ito	Takao	OS21-1	41/86
Huang	Yuankui	OS26-11	46/100			OS22-4	37/89
Huang	Yumei	OS3-8	27/55			OS22-7	38/89
		OS3-9	28/55			OS22-8	38/90
		OS4-8	44/57	Ito	Tsutomu	OS22-7	38/89
		OS4-9	44/58			OS22-8	38/90
Huang	Zehan	OS26-2	45/97	Iwabu	Ami	GS1-7	35/104
Hung	Chung-Wen	OS5-3	39/59				
		OS5-4	39/59	[J]			
Hussain	A. -S. T.	OS23-4	24/91	Jamaludin	Farah Adilah	OS25-10	49/96
		OS23-5	24/91			OS26-10	46/100
Hussin	Eryana	OS25-8	48/96	Jamil	Md Shafayet	GS1-2	34/103
				Jelban	Ahmad	OS25-4	48/94
[I]				Jia	Yingmin	OS8-1	35/63
Ikeda	Akihiro	GS5-3	37/109			OS8-4	36/64
Ikeda	Satoshi	OS22-4	37/87	Jiao	Jianhao	OS4-4	44/56
		OS22-5	37/89	Jothi	Neesha	OS18-1	40/78
		OS22-6	38/89			OS18-7	40/79
		OS22-7	38/89	Jung	Jae Hyung	GS1-8	35/105
		OS22-8	38/90				
Imai	Shinichi	GS4-1	23/107	[K]			
Imai	Yoshiro	OS17-5	32/78	K.	Indraah	OS18-8	41/80
Imura	Seigo	GS6-1	42/110			OS18-9	41/80
Ishak	Nurfadzillah	OS23-1	24/90			OS18-10	41/80
		OS23-2	24/90			OS18-11	41/81
Ishigami	Ryoichi	OS9-1	22/64	Kaishita	Wataru	OS20-10	30/84
Ishii	Kazuo	OS20-3	29/82	Kakinoki	Kanta	OS9-4	22/65
		OS20-4	29/83	Kamoda	Yohei	OS27-3	33/101
		OS20-5	29/83	Kamei	Keiji	OS20-9	30/84
		OS20-6	30/83	Kamiya	Tohru	GS1-1	34/103
		OS20-13	31/85			GS5-1	36/109
		OS20-14	31/86	Kanaoka	Daiju	OS15-4	23/75
Ishikawa	Seiji	GS3-3	25/107	Kang	Jinwu	OS24-4	21/93
Ishiwata	Yoshinobu	OS27-2	33/101			OS24-5	21/93
		OS27-4	33/101	Kasai	Shintaro	GS4-3	24/108
Islamova	Renata	OS11-2	47/68	Katayama	Daigo	OS20-14	31/86
Isomoto	Kosei	OS15-3	22/74	Katayama	Tetsuro	OS9-1	22/64
		OS15-4	23/75			OS9-2	22/65
Isozaki	Hideo	OS12-4	34/70			OS9-3	22/65

117

		OS9-4	22/65	Kong	De Kang	OS25-10	49/96
Kawaguchi	Akira	OS16-3	39/76	Koshiba	Mamiko	OS12-1	34/70
Kawahara	Tomorou	OS20-9	30/84	Kotani	Naoki	OS12-3	34/70
Kawakami	Yusuke	OS17-2	32/77	Kouno	Naoki	GS2-1	23/105
		OS17-3	32/77	Koyama	Yoshihiro	OS16-3	39/76
		OS17-4	32/77	Kuan	Chyntia	OS25-7	48/95
Kawashima	Mizuki	OS15-6	23/75	Kudo	Akihiro	OS22-2	37/88
Khan	M.K.A Ahamed	OS6-1	43/60			OS22-3	37/88
		OS6-2	43/60	Kumar	Prashant	OS14-2	36/73
		OS6-3	43/61	Kumsaen	Tinnakorn	GS6-4	42/111
		OS6-4	43/61	Kuremoto	Shun	OS12-1	34/70
		OS6-5	43/61	Kuremoto	Takashi	OS12-1	34/70
		OS7-1	31/62	Kurogi	Shuichi	OS22-5	37/89
		OS7-2	31/62	Kuroki	Karin	GS2-2	23/105
		OS7-3	31/62	Kyasudeen	Muhammad	OS21-1	41/86
		OS7-4	31/63		Farris		
		OS26-9	46/99				
		OS26-10	46/100	[L]			
Kido	Sachiko	GS4-2	23/107	Lai	Chun-Chi	OS5-2	38/58
		GS5-4	37/110	Lai	Wei Kang	OS26-3	45/97
Kim	Tae Ihn	GS1-8	35/105			OS26-4	45/98
Kimura	Kenji	OS20-6	30/83			OS26-5	45/98
		OS20-7	30/83			OS26-6	45/98
		OS20-8	30/84	Lavrenov	Roman	OS10-6	28/67
Kimura	Tomoka	GS5-4	37/110			OS11-5	47/69
Kita	Yoshihiro	OS9-2	22/65	Lee	Bo-Young	GS6-2	42/110
		OS9-3	22/65	Lee	L. J.	OS23-5	24/91
		OS9-4	22/65	Li	Chu-Fen	OS1-3	21/51
Kitamura	Kentarou	GS1-7	35/104	Li	Fangyan	OS3-1	27/53
Kitazumi	Shun	GS2-4	23/106	Li	Guandong	OS26-11	46/100
Koh	Pei Cong	OS18-6	40/79	Li	Haibo	OS13-8	26/72
Kolandaisamy	Raenu	GS7-6	49/113	Li	Hongbing	OS11-4	27/69
		OS18-8	41/80	Li	Huahao	OS4-1	44/56
		OS18-9	41/80	Li	Jung-Shian	OS1-1	21/51
		OS18-10	41/80			OS1-2	21/51
		OS18-11	41/81			OS1-3	21/51
		OS18-12	41/81			OS1-4	21/51
Kolin	Nikita	OS11-3	47/68			OS1-5	21/52
Komori	Ryohei	OS20-12	30/85	Li	Junkai	GS4-6	24/108

118

Li	Lin	OS8-5	36/64
Li	Mingxuan	OS4-8	44/57
		OS4-9	44/58
Li	Mingyue	OS4-4	44/56
Li	Na	OS24-4	21/93
Li	Zixin	OS24-3	20/92
Liang	Xin	OS13-4	26/71
Lim	Tiong Hoo	OS26-8	46/99
		OS26-9	46/99
Lim	Wei Hong	OS6-1	43/60
		OS6-3	43/61
		OS26-2	45/97
		OS26-3	45/97
		OS26-4	45/98
		OS26-5	45/98
		OS26-6	45/98
		OS26-7	46/99
		OS26-8	46/99
		OS26-9	46/99
		OS26-10	46/100
		OS26-11	46/100
Lim	Wei Yee	GS7-6	49/113
Liu	Chun-Liang	OS5-4	39/59
Liu	Fangyv	OS2-3	25/53
		OS3-3	27/54
		OS3-4	27/54
Liu	I-Hsien	OS1-1	21/51
		OS1-2	21/51
		OS1-3	21/51
		OS1-4	21/51
		OS1-5	21/52
Liu	Mengda	OS4-7	44/57
Liu	Yi-Ting	OS5-2	38/58
Liu	Yufei	OS20-3	29/82
Lu	Yuhao	OS26-11	46/100
Lv	Lei	OS13-4	26/71

[M]

M Noor	Muhammad Badri	OS21-1	41/86
M.Tahir	Noor Idayu	OS25-1	47/93
		OS25-2	48/94
		OS25-3	48/94
		OS25-4	48/94
		OS25-5	48/95
		OS25-6	48/95
		OS25-7	48/95
Ma	Dongpo	OS13-9	26/73
		OS13-10	27/73
Ma	Shuangshuang	OS2-1	25/52
Mabu	Shingo	OS12-1	34/70
Machida	Kosei	GS4-1	23/107
Magid	Evgeni	OS5-5	39/59
		OS5-6	39/60
		OS10-2	28/66
		OS10-4	28/67
		OS11-1	47/68
		OS11-4	47/69
		OS11-6	47/69
Mahmoud E.A.	Zeiad	OS25-5	48/95
Marco	Rozilyn	OS14-2	36/73
Matoba	Osamu	OS17-2	32/77
		OS17-3	32/77
Matsumoto	Sho	OS9-1	22/64
Mazher	Moona	OS7-1	31/62
		OS7-2	31/62
		OS7-3	31/62
		OS7-4	31/63
Mei	Yuping	OS3-5	27/54
		OS3-6	27/55
		OS3-7	27/55
Meng	Tianbin	OS24-1	20/92
Minato	Gaku	OS20-2	29/82
Mirzoian	Viktoriia	OS10-4	28/67
Mitsuyama	Fuga	GS1-1	34/103
Mizoe	Eiji	OS20-10	30/84
		OS20-11	30/85

Mizutani	Akinobu	OS15-2	22/74	Nakamura	Akira	OS14-1	36/73
		OS15-4	23/75	Nakamura	Jun	OS27-2	33/101
Mizumoto	Yoshitaka	GS4-4	24/108			OS27-3	33/101
Mo	Lipo	OS8-4	36/64			OS27-4	33/101
Mohamed	Rizon	GS4-6	24/108	Nakashima	Riku	GS1-6	35/104
Mohammad Arif	Ilyas	OS26-3	45/97	Narayanan	Ganesh	OS18-4	40/79
				Niederer	Steven	OS7-1	31/62
		OS26-4	45/98			OS7-2	31/62
		OS26-5	45/98			OS7-3	31/62
		OS26-6	45/98			OS7-4	31/63
Mokayef	Mastaneh	OS6-1	43/60	Nishida	Naoya	OS20-11	30/85
		OS6-2	43/60	Nishida	Yuya	OS20-13	31/85
		OS6-3	43/61	Noor	Rayan	OS26-8	46/99
		OS6-4	43/61	Mohammed Bakhit	Mohammed		
		OS6-5	43/61	Nur Muhammad	Afifi	OS25-6	48/95
		OS7-1	31/62				
		OS7-2	31/62				
		OS7-3	31/62	**[O]**			
		OS7-4	31/63	Ogata	Takashi	OS27-1	33/100
		OS26-2	45/97	Ogawa	Shintaro	OS16-2	39/76
		OS26-3	45/97			OS16-4	40/76
		OS26-6	45/98	Ohno	Asako	OS17-5	32/78
Mokhtar	Norrima	OS18-6	40/79	Ohara	Makoto	OS12-4	34/70
		OS21-1	41/86	Ohe	Kaoru	OS22-1	37/88
		OS21-3	42/87			OS22-5	37/89
Moribe	Iichirou	GS3-2	25/106			OS22-6	38/89
Moshkin	Vladimir	OS5-6	39/60	Okawachi	Yuto	OS16-2	39/76
Muhammad Risal		GS5-2	36/109			OS16-4	40/76
Muthusamy	Hariharan	OS21-2	41/87	Okazaki	Naonobu	OS9-2	22/65
Mustapha	Ain Atiqa	OS25-3	48/94			OS9-3	22/65
		OS25-4	48/94			OS9-4	22/65
		OS25-7	48/95	Okuda	Haruhisa	PS3	29/50
Mustafin	Maksim	OS10-4	28/67	Onaga	Haruki	OS9-2	22/65
				Ono	Jumpei	OS27-1	33/100
[N]				Ooba	Miki	GS2-3	23/105
Nagayoshi	Masato	OS12-2	34/70	Oshima	Tatsuya	OS22-1	37/88
Naito	Yuka	OS27-2	33/101	Ozaki	Tomoaki	PS1	20/50
Najirah Umar		GS5-2	36/109				

120

[P]			
Pak	Alexander	OS10-5	28/67
Pan	Li	OS26-8	46/99
Paramasivam	Sivajothi	OS6-5	43/61
Park	Chan Gook	GS1-8	35/105
Patan	R.	OS23-4	24/91
Peng	Bo	OS26-2	45/97
Peng	Tzu-En	OS1-2	21/52
		OS1-4	21/51
Peng	Yizhun	OS13-1	26/71
		OS13-2	26/71
		OS13-3	26/71
		OS13-4	26/71
		OS13-5	26/72
		OS13-6	26/72
		OS13-7	26/72
		OS13-8	26/72
		OS13-9	26/73
		OS13-10	27/73
Permyakov	Alexander	OS5-6	39/60
Phan	Duc	GS6-3	42/111
Phiri	Gershom	OS6-1	43/60
Pramanta	Dinda	OS15-1	22/74
Pujianti Wahyuningsih		GS5-2	36/109
[Q]			
Qi	Pon Xiao	GS7-1	49/111
Qin	Jianfeng	OS13-7	26/72
Qayyum	Abdul	OS6-1	43/60
		OS6-2	43/60
		OS6-3	43/61
		OS6-5	43/61
		OS7-1	31/62
		OS7-2	31/62
		OS7-3	31/62
		OS7-4	31/63
[R]			
Rahman	Hameedur	OS26-9	46/99
		OS26-10	46/100
Rajagopal	Heshalini	OS18-3	40/78
		OS18-5	40/79
		OS18-6	40/79
		OS18-8	41/80
		OS18-9	41/80
		OS18-10	41/80
		OS18-11	41/81
		OS18-12	41/81
		OS21-1	41/86
Rajagopal	Renuka	OS18-3	40/78
Rajapakse	R.P.C. Janaka	OS17-4	32/77
Ramis	Kulmukhametov	OS11-6	47/69
Rathinam	Manoj	OS18-3	40/78
Razzak	Imran	OS7-1	31/62
		OS7-2	31/62
Ren	Xiang	OS2-4	25/53
		OS2-5	25/53
Riajuliislam	MD	OS22-7	38/89
Rizwan	Amar	OS25-1	47/93
		OS25-3	48/94
[S]			
S	Amutha	OS18-2	40/78
S	Akshay	OS18-3	40/78
Sadyiko c a	Karina	OS11-5	47/69
Safin	Ramil	OS10-2	28/66
		OS11-6	47/69
Sakamoto	Makoto	OS22-2	37/88
		OS22-3	37/88
		OS22-4	37/89
		OS22-5	37/89
		OS22-6	38/89
		OS22-7	38/89
		OS22-8	38/90
Salman	Riham	OS11-2	47/68
Sendari	Siti	OS21-1	41/86
Saleh	Ghassan	OS18-6	40/79

121

Takeoka	Kenta	OS12-3	34/70			GS3-2	25/106
Tamaki	Hisashi	OS12-2	34/70			GS3-3	25/107
Tamukoh	Hakaru	OS15-1	22/74	Terasawa	Takashi	GS5-1	36/109
		OS15-2	22/74	Tew	Hwa Hui	OS25-2	48/94
		OS15-3	22/74	Thamadharan	Kavitha	OS18-1	40/78
		OS15-4	23/75	Thota	Tejaswini	OS18-2	40/78
		OS15-5	23/75	Tian	Shaokai	OS13-5	26/72
Tamura	Hiroki	GS1-2	34/103	Tiang	Sew Sun	GS4-6	24/108
		GS1-3	34/103			OS6-1	43/60
		GS1-4	35/104			OS6-3	43/61
		GS1-5	35/104			OS26-2	45/97
		GS1-6	35/104			OS26-3	45/97
		GS4-2	23/107			OS26-4	45/98
		GS5-4	37/110			OS26-5	45/98
		GS6-1	42/110			OS26-6	45/98
Tan	Chi Jie	OS16-2	39/76			OS26-7	46/99
		OS16-4	40/76			OS26-8	46/99
Tan	Jia Wei	OS26-3	45/97			OS26-9	46/99
		OS26-4	45/98			OS26-10	46/100
		OS26-5	45/98	Ting	Tin Chang	OS26-9	46/99
		OS26-6	45/98			OS26-10	46/100
Tan	Rodney	OS25-1	47/93	Titan	Janthori	OS16-2	39/76
		OS25-10	49/96			OS16-4	40/76
Tanaka	Takeshi	OS17-5	32/78	Tominaga	Ayumu	OS16-2	39/76
Tanaka	Toshiki	OS17-1	32/77			OS16-4	40/76
		OS17-2	32/77			OS20-12	30/85
		OS17-3	32/77	Tominaga	Moeko	OS20-10	30/84
Tanaka	Yuichiro	OS15-2	22/74			OS20-11	30/85
		OS15-4	23/75			OS20-14	31/86
Tanev	Ivan	OS17-1	32/77	Tomioka	Toyoaki	OS22-4	37/89
Taniguchi	Rie	OS19-3	37/82	Too	Kok Sem	OS25-9	48/96
Tanveer	M. H.	OS23-3	24/91	Tsoy	Tatyana	OS5-5	39/59
		OS23-4	24/91			OS10-1	28/66
		OS23-5	24/91			OS10-3	28/66
Taniguchi	Yasutaka	OS20-1	29/82			OS10-5	28/67
		OS20-2	29/82			OS11-1	47/68
Tanjo	Yui	GS2-3	23/105			OS11-2	47/68
		GS2-4	23/106			OS11-6	47/69
		GS2-5	23/106	Tsuji	Tokuo	OS14-3	36/74

[U]			
Umeno	Ren	GS1-1	34/103
ul Husnain	Anees	OS21-1	41/86
Urano	Shingen	OS20-1	29/82
[W]			
Wadi harun	Sulaiman	OS21-3	42/87
Wagatsuma	Hiroaki	GS4-3	24/108
		GS6-2	42/110
		GS6-3	42/111
Wan	Weiwei	OS14-2	36/73
Wan Ahmad	Wan Khairunizam	OS21-2	41/87
Wang	Chun-Chieh	OS5-3	39/59
Wang	Depeng	OS4-10	45/58
Wang	Haiquan	OS3-5	27/54
Wang	Jiashuai	OS2-1	25/52
Wang	Jiaxin	OS2-3	25/53
		OS3-3	27/54
		OS3-4	27/54
Wang	Jiqiang	OS8-2	35/63
Wang	Jiwu	OS24-1	20/92
		OS24-2	20/92
		OS24-3	20/92
		OS24-4	21/93
		OS24-5	21/93
Wang	Limei	OS13-1	26/71
Wang	Ningfei	OS26-2	45/97
Wang	Peng	OS4-5	44/57
		OS4-6	44/57
		OS4-7	44/57
Wang	Qikun	OS4-6	44/57
		OS4-7	44/57
Wang	Shengfeng	OS4-6	44/57
Wang	Shuai	GS3-1	25/106
Wang	Shuxin	OS13-8	26/72
Wang	Siyi	OS13-6	26/72

Wang	Xiaolong	OS24-4	21/93
Wang	Xuran	OS4-3	44/56
Wang	Zipei	OS2-4	25/53
		OS2-5	25/53
Watanabe	Keisuke	OS20-1	29/82
		OS20-2	29/82
Watanabe	Konosuke	OS20-1	29/82
		OS20-2	29/82
Wen	Zongyan	OS26-11	46/100
Wong	Chin Hong	OS6-2	43/60
		OS26-2	45/97
		OS26-7	46/99
		OS26-8	46/99
		OS26-9	46/99
		OS26-11	46/100
Wu	Qihua	OS24-4	21/93
Wu	Tai-Hsuan	OS5-4	39/59
Wu	YingCheng	OS1-3	21/51
[X]			
Xia	Yuntian	OS13-2	26/71
		OS13-6	26/72
Xiang	Yande	OS4-10	45/58
Xiao	Ziyue	OS3-9	28/55
		OS4-8	44/57
		OS4-9	44/58
		OS4-10	45/58
Xie	Jiahao	OS3-8	27/55
		OS3-9	28/55
Xie	Shengke	OS18-12	41/81
Xie	Yanzhe	OS26-11	46/100
[Y]			
Yabuki	Tomohide	GS1-1	34/103
Yamaba	Hisaaki	OS9-2	22/65
		OS9-3	22/65
		OS9-4	22/65
Yamada	Shintaro	GS3-3	25/107
Yamaguchi	Hiromasa	OS15-2	22/74

124

Developing a High-Speed Working Motion of the Multi Robot in DENSO

Tomoaki Ozaki

DENSO CORPORATION Advanced Research and Innovation Center,
500-1, Minamiyama, komenoki-cho, Nisshin-shi, Aichi-ken, 470-0111, Japan
Email: tomoaki.ozaki.j4a@jp.denso.com

Abstract

The number of labor force is expected to decrease due to the declining birthrate and aging population. Although many efforts to solve this social issue using automation technology by robots around the world have been implemented, most of the current applications of robots are still repetitive works such as picking and placing in mass production lines in factories, and small progress of the application of robots for high-mix low-volume production lines where the operations are is frequently changed. In this paper, we discuss the reasons and the potential solutions for autonomous control technologies including the AI and deep learning. Moreover, the high-speed working motion of the multi robot developed by DENSO will be presented in this paper.

Keywords: Unmanned Manufacturing, Motion Planning, Deep Neural Networks (DNN), Intelligent robotics

1. Introduction

The decline in the working-age population due to aging is progressing worldwide, becoming a serious social issue. In particularly Japan, we have been facing on a severe decrease in the working-age population, with estimates suggesting a decrease to 52.75 million people (a 29.2% decrease from 2021) by 2050 [1]. To address this issue, DENSO is actively intensifying its research and development efforts in automation technology utilizing robots. At DENSO's factories, we aim to achieve fully unmanned manufacturing and significantly reduce production lead times through the utilization of AI in process design for automation by 2035. We are also advancing the development of robots for automated harvesting of crops in agriculture field and automated disassembling of vehicles in the field of circular economy. As the utilization of robots advances in various sectors of society, such as production and agriculture, our group company DENSO WAVE has developed the human-collaborative robot COBOTTA PRO. COBOTTA PRO is a human-collaborative robot that combines the speed (2,500mm/s) and precision (repeatability ±0.04mm) comparable to industrial robots. It maximizes its operation speed when there are no workers nearby and minimizes the distance required for deceleration and stoppage when a worker approaches, achieving a good balance between productivity and safety aspect.

However, there are still several important issues that need to be addressed for further utilization of robots. Firstly, one of these issues is related to the design of robot movements. The work of teaching and programming for robot motion design requires specialized technicians to physically operate the robot and refine its movements, resulting in significant time and effort when introducing robots. Although the cost of the robot itself is becoming affordable in the market, it cannot be ignored due to its motion design cost when frequent motion changes are made in low-volume, high-mix production. The same issue arises in scenes such as agriculture field or vehicle disassembling outside of factory, where it is difficult to fix the motion of robots due to their variability.

The second issue is related to adapting to changes and variations in the workpieces. In the teaching playback method commonly used in repetitive tasks in factories, it is difficult to manage changes and variations in the workpieces, making it challenging to accommodate misalignment or stacking of the workpieces. Efforts to address these issues have progressed at a practical level by utilizing vision systems. The vision system uses 2D or 3D sensor cameras and recognition algorithms to recognize the position and posture of the workpiece and the surrounding environment, and the robot system as a whole uses this information to generate robot movements according to the situation. However, the range of workpieces that can be handled is still limited, and the recognition capabilities for complex shapes, reflective properties, transparent objects, and even indefinite shapes are still insufficient. Furthermore, the construction of these vision systems requires the actual objects of the workpieces for adjustment, which involves significant time and effort.

Additionally, tasks that involve handling flexible materials are particularly challenging for robots. The tasks such as handling wire harnesses and arranging components that undergo significant shape changes due

to interactions with the robot and the environment are difficult for robots to predict and respond to, making them one of the tasks where automation has not yet advanced in factories.

2. Literature Survey

Efforts to automate teaching are primarily progressing through the development of motion planning techniques. Motion planning is the technology that automatically generates a path connecting the start and end points of robot motion. Traditional manual teaching requires significant effort to create robot paths that are collision-free and have the shortest possible motion time. Motion planning techniques are predominantly based on the sampling-based methods, with Rapidly-exploring Random Trees (RRT [2]) being devised and subsequently improved upon [3], [4], [5]. Currently, these techniques have reached a practical stage where they are implemented within robot simulators and used for pre-verification of motion in simulation environments. However, there are issues related to the trade-off between computation time and the quality of generated paths, as well as the variability in the computation time required for path generation. Further technological advancements are expected to enable real-time motion generation, particularly in complex environments with narrow spaces or multiple robots sharing the workspace, where computation time significantly increases. One promising approach for generating high-quality paths with fast and constant computation time is to utilize Deep Neural Networks (DNN) as an alternative to traditional computation methods. For example, Chi, C., et al. have employed diffusion models to acquire DNN models for generating robot motion paths, which have demonstrated superior performance in terms of discreteness and multimodality compared to conventional imitation learning [6]. As improvements in learning stability and environmental robustness progress, expectations for practical implementation are increasing.

One example of addressing changes and variations in work pieces is the utilization of AI in vision systems. The Dex-Net (The Dexterity Network [7]) project has constructed a large-scale database for object grasping and developed models that predict the gropability of given objects by training DNN using the database as training data. This project utilizes not only real-world data but also synthetic point clouds generated from 3D models, significantly reducing the cost of creating training data. The scalability of both DNN models and training data is expected to enhance generalization, leading to improved performance for unknown objects.

In the Dactyl project by OpenAI, they demonstrate an example of acquiring dexterous manipulation skills using deep reinforcement learning, where a robot with five human-like fingers adeptly manipulates a cube [8]. The Deep Learning techniques often require a large amount of training data to achieve high performance, which presents a challenge for real-world applications. However, Dactyl addresses this issue by employing a technique called domain randomization, which diversifies the physical behavior in the simulation environment. By training solely in the simulated space with varied physics, it is able to acquire movements that can be applied Denso's in the real world. While the current examples in the Dactyl project primarily focus on manipulating rigid objects, it is anticipated that similar techniques will be explored to extend the acquisition of movements for flexible objects in the future.

3. Research in Denso

In order to replace robots with more advanced tasks in the future, it is essential to have coordination among multiple robot arms as a single robot arm has its limitations. However, when trying to apply RRT-based motion planning to multiple robots, as mentioned earlier, the search space expands exponentially with the increase in the number of robots, resulting in a significant increase in computation time. To address this challenge, we incorporated the expertise of skilled teaching engineers into the exploration algorithm and effectively constrained the search range, significantly reducing computation time while maintaining path optimality. Fig. 1 illustrates the relationship between the number of controlled robots and the computation time required to output the initial solutions for the generated paths. With the conventional RRT-based method, the computation time exceeded 31 seconds for two robots, over 32 minutes for three robots, and reached more than 87 hours for four robots (estimated computation time for four robots). However, with the method we developed, the computation time was reduced to approximately 0.4 seconds for two robots, 0.7 seconds for three robots, and 1.6 seconds for four robots.
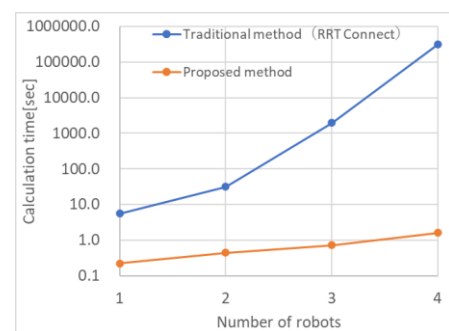
Fig. 1. Relationship between computation time and number of robots in route generation

Furthermore, by adding post-processing for path optimization, we reduced the robot motion time. Through practical verification in actual pick-and-place operations (In the environment shown in Fig. 2, we conducted a demonstration of placing 75 CDs to three designated locations using three robots.), we confirmed a 23.8% reduction in robot operation time compared to the results

of skilled teaching engineers (The results of skilled teaching engineers: 48.9 seconds, Automatic Generation by Motion Planning: 37.2 seconds).

2-1. View of the actual machine environment
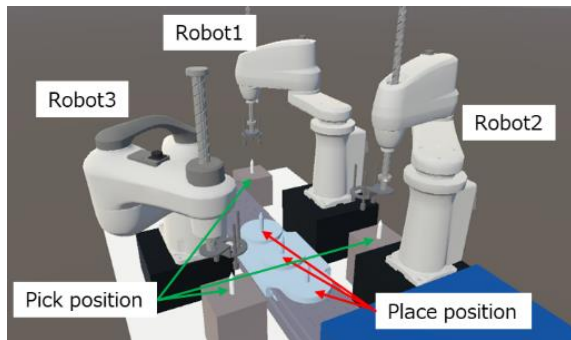
2-2. Overview of the demonstration environment
Fig. 2. The demonstration of pick-and-place operations with multiple robots

Table 1 shows a more detailed comparison. In a multi-robot operation, the amount of movement (joint movement) as well as the degree to which the robot is able to operate simultaneously greatly affect its working time. When we compared these results, we found that there was no significant difference in joint movement between the teaching result and the auto-generated result (The auto-generated result showed approximately 3.7% less movement compared to the teaching result.), but the auto-generated result showed a greater improvement of over 30% in the simultaneous movement rate compared to the teaching result.

In this development case study, we have shown that we can automate advanced teaching tasks that handle multiple robots and generate highly optimal paths compared to skilled teaching engineers. In the future, we intend to make it possible to automatically generate placement of robots and peripheral equipment, and operation procedures, leading to fully automated start-up of equipment.

Table. 1. Comparison of the results of motion generation by teaching and automatic generation algorithms

1-1. Comparison of joint movement

	Joint movement [rad]		Reduction rate [%]
	(a) Teaching method	(b) Proposed method (Automatic generation)	Reduction rate from (a) to (b)
Robot 1	6363.3	6268.7	1.49
Robot 2	5835.7	5352.7	8.28
Robot 3	4209.9	4176.7	0.79
Total	16408.9	15798.1	3.72

1-2. Comparison of simultaneous operation rates of multiple robots

	(a) Teaching method	(b) Proposed method (Automatic generation)	Improvement value from (a) to (b)
Total operation time [sec]	48.9	31.7	-
Maximum possible concurrent operation time [sec]	46.6	29.3	-
Actual Concurrent Operation Time [sec]	28.2	26.8	-
Rate of concurrent operation [%]	60.6	91.4	30.8

4. Conclusion

To expand the applications of robots, the development of technology that automatically generates their movements is desired. Motion planning is an effective technique for automatically generating optimal paths for robots to reach their work points. In this paper, we introduced examples of automatically generating movements, specifically for cases where multiple robots operate simultaneously with high motion speed, to demonstrate the effectiveness of motion planning. The technology for accurately recognizing and manipulating the position and orientation of target objects is still under active research. We are confident that the advancement of deep learning-based techniques will contribute to solving these challenges. Unlike language and images, the cost of generating large amounts of training data is an issue in robot motion. However, by utilizing synthetic point clouds generated from 3D models and physical simulations, successful examples of robot motion

generation applicable to real space are emerging, and further performance improvements can be expected.

References

1. Cabinet Office. (2022). 2022 edition of the white paper on aging society. Retrieved from https://www8.cao.go.jp/kourei/whitepaper/w-2022/zenbun/pdf/1s1s_01.pdf [In Japanese].
2. LaValle, S. M., & Kuffner, J. J. (2001). Rapidly-exploring random trees: Progress and prospects. In B. R. Donald, K. M. Lynch, & D. Rus (Eds.), Algorithmic and computational robotics: New directions (pp. 293-308). A. K. Peters.
3. Kuffner, J. J., & LaValle, S. M. (2000). RRT-connect: An efficient approach to single-query path planning. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '00) (Vol. 2, pp. 995-1001). San Francisco, CA.
4. Kavraki, L. E., Svestka, P., Latombe, J.-C., & Overmars, M. H. (1996). Probabilistic roadmaps for path planning in high-dimensional configuration spaces. IEEE Transactions on Robotics and Automation, 12(4), 566-580.
5. Harada, K. (2014). ロボットの動作計画における最適化 [Optimization in robotic motion planning]. 日本ロボット学会誌 [Journal of the Robotics Society of Japan], 32(6), 508-511. [In Japanese].
6. Chi, C., Feng, S., Du, Y., Xu, Z., Cousineau, E., Burchfiel, B., & Song, S. (2023). Diffusion policy: Visuomotor policy learning via action diffusion. Retrieved from https://arxiv.org/abs/2303.04137
7. Mahler, J., & Goldberg, K. (n.d.). Dex-Net: Deep learning for dexterous robotic manipulation. Berkeley AUTOLAB. Retrieved from https://berkeleyautomation.github.io/dex-net/
8. Andrychowicz, M., Baker, B., Chociej, M., et al. (2018). Learning dexterous in-hand manipulation. Retrieved from https://arxiv.org/abs/1808.00177

Authors Introduction

Mr. Tomoaki Ozaki

He is the Chief of Intelligent Robotics R & I Section in AI R & I Division at DENSO CORPORATION. He graduated from the Graduate School of Information Technology, Kyushu Institute of Technology in 2001. He is currently working on research and development related to the application of AI in factory automation, with a specific focus on autonomous control of robot arms.

Experimenting with Variable Arm Quadrotors: Realizing Dynamic Configurations for Enhanced Flight Performance

Hazry Desa

Centre of Excellence for Unmanned Aerial System (COE-UAS), Universiti Malaysia Perlis, Block E, Pusat Perniagaan Pengkalan Jaya, Jalan Kangar – Alor Setar, 01000 Kangar, Perlis, Malaysia

Muhammad Azizi Azizan

Centre of Excellence for Unmanned Aerial System (COE-UAS), Universiti Malaysia Perlis, Block E, Pusat Perniagaan Pengkalan Jaya, Jalan Kangar – Alor Setar, 01000 Kangar, Perlis, Malaysia
Email: hazry@unimap.edu.my

Abstract

This paper introduces two innovative concepts for variable arms designed for a quadrotor, enabling precise control of its movement through manipulation of the bending moment via varying arm lengths. The primary objective of this research is to develop and identify the most suitable variable arm configuration that facilitates smooth and stable quadrotor movement. The study delves into two concept designs that are well-suited for the quadrotor application. By employing a suitable variable arm, the quadrotor's maneuverability can be effectively regulated based on the bending moment adjustments made possible by the variable arm. Ultimately, the paper presents the design and performance testing of two types of variable arms. The obtained results confirm that the variable arm of the electric actuator with linear guide-Type 2 exhibits smooth and stable movement.

Keywords: Variable Arm, Quadrotor, Dynamic Configuration, Flight Performance

1. Introduction

The quadrotor, a type of UAV (unmanned aerial vehicle), holds great potential for overcoming terrestrial challenges [1], leading to its widespread use in various industrial and commercial applications such as geographic mapping, surveillance, agricultural tasks like fertilization or pesticide application, and aerial photography [2]. This continuous utilization has spurred the development of the quadrotor to enhance its stability, performance, and multimodal capabilities [3]. Typically, quadrotors are designed with four rotors, positioned at the vertices of a square frame [4]. To understand how the quadrotor operates, Newton's Third Law of Motion comes into play, stating that every action elicits an equal and opposite reaction [5]. When the quadrotor's motor rotates the propellers, it generates a downward force on the air, as explained by the Bernoulli Principle [6]. In a conventional quadrotor, adjusting motor speeds in the four motors allows the control of thrust and facilitates the required movements [7]. However, as the size and weight of the quadrotor increase, more thrust is needed to lift it, often necessitating larger propellers or faster motors, which in turn lead to higher power consumption and reduced flying time [8]. This conventional design approach can be both costly and inefficient [9].

Considering the issues of power consumption and cost, limited research has explored alternative methods to generate more thrust besides varying motor speed and propeller size [10]. One such study by Wu in 2018 proposed controlling the quadrotor through changes in both motor speed and rotor blade pitch angle, affecting the produced thrust and power consumption [11], [12]. Additionally, a simulation analysis in 'Effects of Variable Arm Length on UAV Control Systems' in 2020 highlighted the impact of varying arm lengths on the quadrotor's bending moment. Increasing the arm length resulted in higher bending moments and, consequently, increased thrust production [13]. To address these challenges and explore new possibilities, this project aims to design and test two concept designs for variable arms, ensuring smooth and stable extension and retraction.

2. Design of variable arm
2.1. *Concept design*

Fig. 1 shows three concept designs of variable arms in both the retract and extend positions. Fig. 1(a)

corresponds to a concept design previously introduced in research [14]. However, when a quadrotor's arm experiences thrust, it may undergo deflection and vibration caused by the motor, impacting the arm's extension. To counter this issue, the use of a rod shaft and linear guide method for the variable arm proves effective in preventing arm deflection and ensuring smooth movement. The design depicted in Fig. 1(a) presents challenges in linear translation control due to the possibility of high deflection and the lack of motion guidance. On the other hand, the designs shown in Fig. 1(b) and Fig. 1(c) employ different methods for arm movement control. Fig. 1(b) employs a stepper motor with a screw shaft while Fig. 1(c) utilizes an electric actuator with a linear guide for variable arm control.

a.   Variable arm concept

b.   Variable arm of stepper motor with screw shaft and rod shaft (type 1)
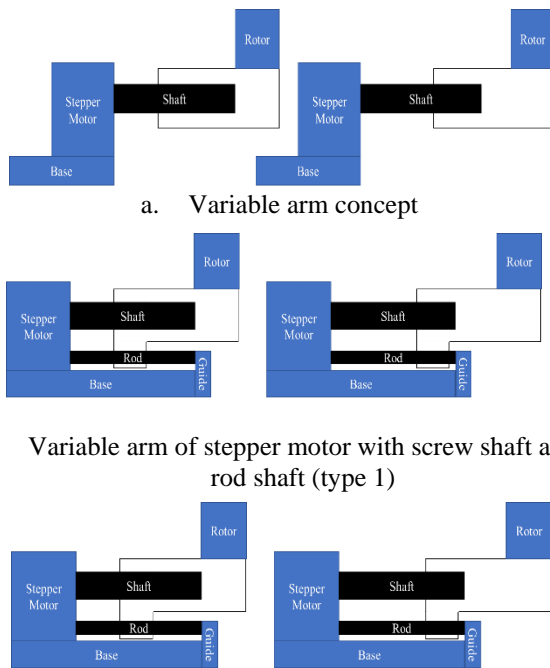
c.   Variable arm of electric actuator with linear guide (type 2)

Fig. 1. Variable arm quadrotor design concept.

2.2. *Variable arm of speed motor with screw shaft and rod shaft (type 1)*

According to Eq. (1), the bending moment is determined by multiplying the force by the length. In conventional quadrotors, maneuverability relies on varying forces to influence the bending moment and consequently rotate the quadrotor's body. This means that the same force can produce different bending moments depending on the length of the variable arm. By manipulating the bending moment, the quadrotor can be directed and moved in a specific direction. For the purpose of testing, variable arms capable of moving within a range of 20 mm to 30 mm have been designed.

$$M = FL \tag{1}$$

Fig. 2 illustrates the variable arm of type 1, utilizing a stepper motor with a screw shaft and rod shaft, in both the extend and retract positions. The 3D drawing is used to calculate the shaft's deflection using Eq. (2) and Fig. 3(a) assists in determining the moment of inertia. The results display the maximum deflection when the variable arm moves from its retracted position to a fully extended position, as depicted in Fig. 3(b).
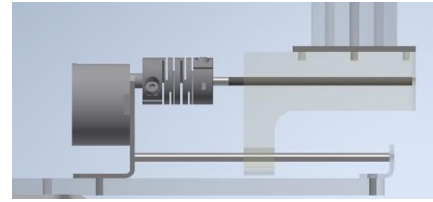
a.   Retract position
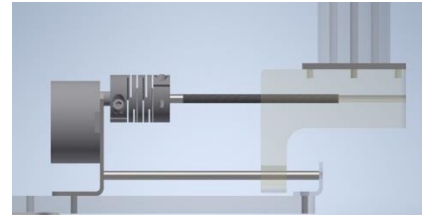
b.   Extend position

Fig. 2. 3D modeling of variable arm steeper motor with screw shaft and rod shaft (type 1).

$$\delta_{max} = \frac{PL^3}{3EI} \tag{2}$$

$$I_x = \frac{\pi r^4}{4} \tag{3}$$

Analyzing Fig. 3(b), it becomes evident that the maximum deflection increases with an elongating length under a constant force. The maximum deflection amounts to approximately 0.001 mm, positioned 66.5 mm away from the support of the bracket design. Given its small magnitude, this deflection does not significantly impact the turning of the screw shaft caused by the deflection of the shaft.
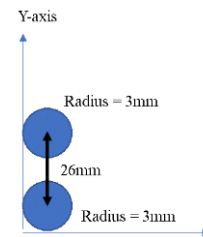
a.   Cross section of screws shaft and rod shaft
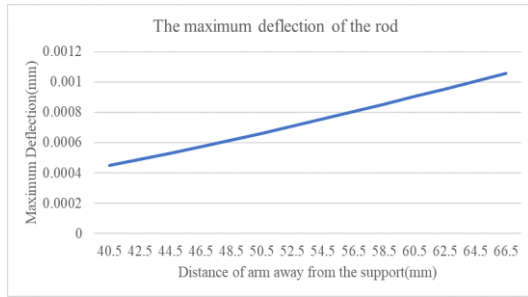
b. Maximum deflection of the shafts
Fig. 3. Deflection analysis.

Furthermore, Table 1 presents the specifications of the stepper motor utilized, indicating an RPM range of approximately 15 to 20. With the use of a screw shaft of 0.5 mm pitch, the variable arm is able to move at the speed of 10 mm/min.

Table 1. The planning and control components.

Technical Parameter	Value
Operating Voltage (V)	5
Operating Current (mA)	240
Step Angle	5.625/64
Reduction Ratio	1/64
Phase	4
Frequency (Hz)	100
Friction torque (gf.cm)	600-1200
Pull in torque (gf.cm)	300
Coil	Unipolar 5 lead coil
Decent Torque (mN.m)	34.3
Speed (RPM)	15-20

2.3. *Variable arm of electric actuator with linear guide (type 2)*

Fig. 4 presents the variable arm of type 2, which utilizes an electric actuator with a linear guide, in both the retracted position (Fig. 4a) and extended position (Fig. 4b). Referring to Table 2 reveals that the dynamic load of the linear guide is approximately 140 kgf. Considering the pull exerted by the brushless motor on the variable arm, it results in the moment of P acting on the linear guide.

a. Retract position
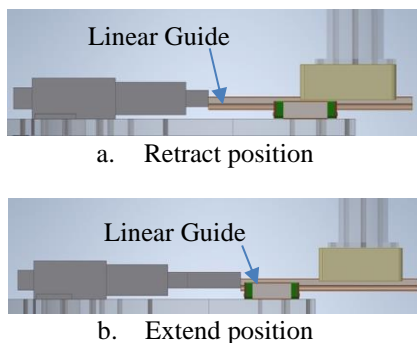
b. Extend position
Fig. 4. Variable arm of electric actuator with linear guide (type 2).

Table 2. Specification of MGN7H linear guide.

Technical Parameter	Value
Dynamic Load (kgf)	140
Static Load (kgf)	200
Moment of R (kgf.m)	0.78
Moment of P (kgf.m)	0.49
Moment of Y (kgf.m)	0.49

Given the small size of the 2 kg quadrotor, a force of 2 kgf is required to maintain equilibrium, while around 3 to 4 kgf thrust is needed to lift the quadrotor. In the design configuration, the center of the brushless motor is positioned 26 mm away from the center of the linear guide, enabling the calculation of the bending moment resulting from the thrust force applied to the linear guide, as per Eq. 1. The linear guide proves to be suitable for use, considering a maximum load of 4 kgf and a maximum bending moment of 0.104 kgf.m, leading to very minimal deflection of the linear guide.

Table 3 presents the specifications of the electric actuator used. As indicated in the table, the speed of the electric actuator is approximately 8 mm/s. Moreover, considering the force of 4 kgf, which equates to around 40 N, so the specification of the electric actuator is suitable for the variable arm application.

Table 3. Specification of LA-YR type electric actuator.

Technical Parameter	Value
Input Voltage (V)	12
No Load Speed (mm/s)	8
Load Push Capacity (N)	90
Load Pull Capacity (N)	90
Static Damping (N)	90
Stroke Length (mm)	30

3. Testing on variable arms

Based on the concept designs presented in the previous section, two types of variable arms have been developed, as depicted in Fig. 5 and Fig. 6. The first type is variable arm of stepper motor with screw shaft and rod shaft referred to as type 1. It comprises a total of 11 components, including a motor driver to control the stepper motor, a bracket to secure the stepper motor, the stepper motor itself, a coupling to connect the stepper motor's shaft with the screw shaft, the screw shaft, a brushless motor, the variable arm, a bracket to hold the rod shaft, the rod shaft, a frame, and a linear bushing that ensures smooth movement of the variable arm along the rod shaft.

On the other hand, the second type, variable arm of electric actuator with linear guide known as type 2, is intended for an electric actuator with a linear guide. It consists of 6 parts: an MDD3A driver to control the electric actuator, the electric actuator, a linear guide, a frame, the variable arm, and a brushless motor.
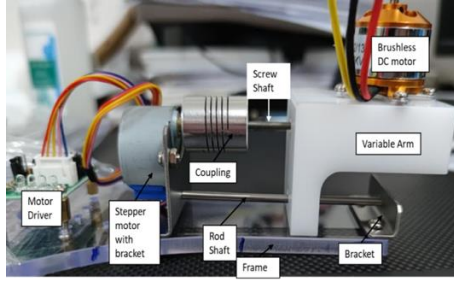
Fig. 5. Variable arm of stepper motor with screw shaft and rod shaft (type 1)
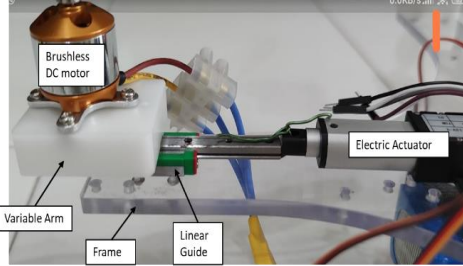
Fig. 6. Variable arm of electric actuator with linear guide (type 2)

Through the utilization of a mobile app equipped with a vibration meter, the vibration of the variable arm in motion is assessed. The phone is placed on the variable arm during the evaluation, and measurements are taken both with and without the operation of a brushless motor. The results are illustrated in Fig. 7 and Fig. 8, displaying the vibration levels with and without the motor's operation for both variable arms.

According to the graphs, it was evident that the brushless motor caused higher vibration, indicating that the motor generates significant vibrations. Additionally, a comparison between the two types of variable arms revealed that type 1 exhibited higher vibration than type 2.
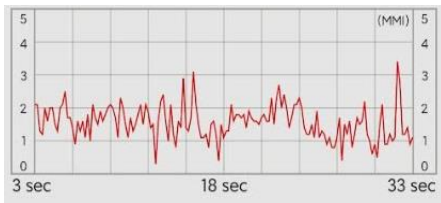
a.   Without operation of brushless motor
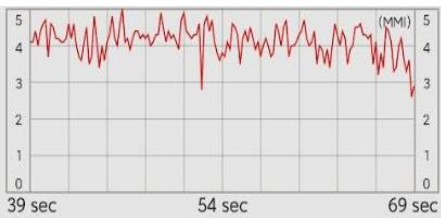
b.   With operation of brushless motor

Fig. 7. Vibration data for variable arm of stepper motor with screw shaft and rod shaft (type 1)
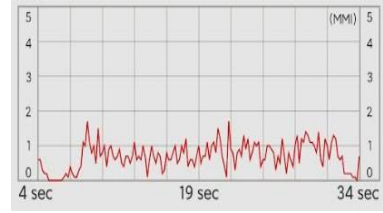
a.   Without operation of brushless motor
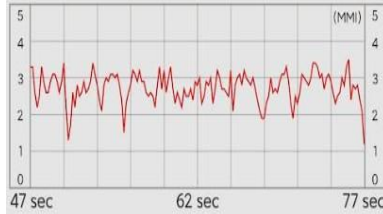
b.   With operation of brushless motor

Fig. 8. Vibration data for variable arm of electric actuator with linear guide (type 2)

4. Conclusion

Table 4. Comparison between 2 designs – type 1 and type 2.

Design	type 1	type 2
Cost	Low	High
Weight (g)	139	96
Movement speed	Slow	Fast
Stability	High vibration	Low vibration

Based on the observations made during the motion of the two variable arms, the results presented in Table 4 indicate that using type 2 would be a more costly option. This is because electric actuators tend to be expensive when compared to stepper motors. However, it should be noted that the electric actuator demonstrated smooth and rapid movement, reaching approximately 8 mm/s, which made it well-suited for quadrotor control. Additionally, the electric actuator's advantage lies in its lightweight nature, as it utilizes fewer components, thereby minimizing the overall weight impact on the quadrotor.

Type 2 indeed offers a cost-effective solution. However, it comes with certain drawbacks. One major concern is the increased weight due to the presence of multiple brackets and shafts, which can adversely affect the overall performance of the system. Moreover, the movement of the variable arm driven by the stepper motor is not stable, mainly due to the high vibration experienced along the arm and its slow-moving speed. This instability can be attributed to the variable arm's elevated position, leading to a higher center of gravity, with the screw and rod shaft situated at the center of the variable arm. Consequently, this setup causes vibrations that oscillate the arm to the left and right. As a result, the thrust force provided by the brushless motor becomes unstable, impacting the system's overall performance.

Additionally, the pitch of the M3 screw shaft is merely 0.5 mm. As a consequence, with each revolution of the

stepper motor, the variable arm can only move by 0.5 mm. This limited displacement capability results in the variable arm moving at a slow speed of only 10 mm/min, making it less suitable for certain applications that require faster movement.

In summary, type 2 is suitable to use as it provides fast motion and stability which is suitable for quadrotors that are sensitive to motion. Besides type 1, it will be suitable to use when a high-pitch screw shaft or higher RPM motor is used so that it can be moved faster. However, there is room for improvement in the design of type 1. One approach is to consider replacing the rod shaft with a linear guide or exploring other methods to enhance stability. Such modifications could help address the issues related to vibration and slow movement, making it a more suitable option for certain scenarios.

Acknowledgements

References

1. Y. H. Tan and B. M. Chen, Underwater Stability of a Morphable Aerial-Aquatic Quadrotor with Variable Thruster Angles, IEEE International Conference on Robotics and Automation (ICRA), 2021, pp. 314–320.
2. H. Shraim, A. Awada and R. Youness R, A Survey on Quadrotors: Configurations, Modeling and Identification, Control, Collision Avoidance, Fault Diagnosis and Tolerant Control, IEEE Aerospace and Electronic Systems Magazine, Vol. 33(7), 2018, pp. 14-33.
3. J. Goslinski, W. Giernacki and S. Gardecki S, Introduction of the Flying Robots into the Human Environment: An Adaptive Square-Root Unscented Kalman Filter for a Fault Tolerant State Estimation in a Quadrotor, IEEE International Conference on Intelligent Environments, 2014, pp. 117-123.
4. G. E. U. Faelden, J. M. Z. Maningo, R. C. S. Nakano, A. A. Bandala and E. P. Dadios, A Neural Network Approach to a Cooperative Balancing Problem in Quadrotor-Unmanned Aerial Vehicles (QUAVs), International Conference on Humanoid, Nanotechnology, Information Technology, Communication & Control, Environment and Management (HNICEM), 2016, pp.1-5.
5. M. F. Ahmed and Y. S. Narayan, Fabrication and Testing of Quadcopter Prototype for Surveillance, International Journal of Mechanical & Production Engineering Research and Development, 2018, pp. 99-105.
6. S. A. Khan, Z. Mehmood and Z. Afshan, Design, Analysis and Topology Optimization of a Landing Gear Strut for a Quadcopter Upon Impact, International Conference on Applied and Engineering Mathematics, 2021, pp. 37-42.
7. Y. R. Tang and Y. Li Y, Realization of the Flight Control for an Indoor UAV Quadrotor, International Conference on Information and Automation (ICIA), 2014, pp. 1278-1283.
8. A. V. Javir, K. Pawar, S. Dhudum, N. Patale and S. Patil, Design Analysis and Fabrication of Quadrotor, Journal of The International Association of Advance Technology and Science (JIAATS), Vol. 16, 2015.
9. N. Y. Kamil, D. Hazry, K. Wan and Z. M. Razlan, Trajectory Tracking Based on Arm's Length Variation, Journal of Theoretical and Applied Information Technology, Vol. 79 (3), 2015, pp. 528-536.
10. N. Y. Kamil, D. Hazry, K. Wan, M. Z. Razlan and A. O. Khaldoon, Payload Capability of VAL – Quadrotor Based on PID Controller, International Journal of Mechanical and Mechatronics Engineering, Vol. 16(2), 2016, pp. 22-29.
11. M. Cutler and J. P. How, Analysis and Control of a Variable-Pitch Quadrotor for Agile Flight, Journal of Dynamic Systems, Measurement, and Control, Vol. 137(10), 2015.
12. S. Sheng and C. Sun, Control and Optimization of a Variable-Pitch Quadrotor with Minimum Power Consumption, Energies, Vol. 9(4), 2016.
13. N. Y. Kamil, D. Hazry, K. Wan and M. Z. Razlan, A Novel VAL: Quadrotor Control Technique for Trajectory Based on Varying the Arm's Length, ARPN Journal of Engineering and Applied Sciences, Vol. 11(5), 2016, pp. 9195-9204.
14. M. Rizon, C. K. Ang, M. I. Solihin, M. Z. Razlan, D. Hazry, S. A. Bakar, K. Wan and I. Zunaidi, Effects of Variable Arm Length on UAV Control Systems, Journal of Robotics, Networking and Artificial Life, Vol. 7 (2), 2020, pp. 91-97.

Authors Introduction

Dr. Hazry Desa

He obtained his PhD in Materials Science and Production Engineering (Robotics) from Oita University and currently holds the position as a Head at the Centre of Excellence for Unmanned Aerial Systems (COE-UAS) at Universiti Malaysia Perlis (UniMAP).

Dr. Muhammad Azizi Azizan

He received his PhD in Civil Engineering from the Universiti Malaysia Perlis. He is currently a Senior Lecturer in the same institution. He is Head of Project Integration & Management (PIM) at Centre of Excellence for Unmanned Aerial System (COEUAS), Universiti Malaysia Perlis.

Artificial Intelligence and Technologies of Arm-type and Mobile Robots in Industry

Haruhisa Okuda

Mitsubishi Electric Corporation, 8-1-1 Tsukaguchi-honmachi, Amagasaki, Hyogo, 661-8661, Japan
Email: Okuda.Haruhisa@ct.MitsubishiElectric.co.jp

Abstract

In recent years, labor shortage has become a serious issue in industrial fields. Various technologies including robot and information processing system to realize flexible work like humans are effective solutions to this issue. Artificial intelligence technology of arm-type robots equipped with 3D sensors and force sensors has been applied in the manufacturing field to cope with different intelligent and highly precise tasks in Mitsubishi Electric. In addition, various technologies to expand the scope of application to the service field, as well as to realize highly functional delivery with mobile robots is under development. Furthermore, IoT technology is also being used for easy and quick on-site implementation and efficient operation. This paper introduces these initiatives with actual examples.

Keywords: Industry, Arm-type robot, Service, Mobile robot, Artificial intelligence, IoT, 3D sensor, Force sensor

1. Introduction

In recent years, the decline in the working population is accelerating worldwide as the birthrate declines and the population ages and labor shortage has become a serious issue in industrial fields. Many intelligent technologies including robot system and information processing system to realize flexible work like humans are effective solutions to this issue, because there is a serious shortage of skilled and experienced workers. Based on these ideas, Mitsubishi electric corp. has developed the intelligent technologies of arm-type robots equipped with 3D sensors and force sensors in the manufacturing field to cope with intelligent and highly precise tasks [1], [2], [3]. By combining artificial intelligent (AI) technology with these technologies, we have developed technologies such as automatic parameter adjustment technology to achieve optimum motion adjustment quickly, and technology for safe operation by determining abnormal conditions [4].

To spread various applications with intelligent robots, we have developed the teaching-less robot system using intuitive I/F technology to facilitate teaching operations that previously required a high level of expertise and skilled experience. This system includes technologies to generate optimum routes and movements, interference checking algorithm with the surrounding area, and series of technologies for bin picking of random stacked objects using 3D sensor.

There is a growing demand for automation of transport operations not only in factories, but also in various facilities such as logistics, public buildings, commercial facilities, and hospitals. For this reason, multiple technologies to expand the scope of application to the service field, as well as to realize highly functional transportation with mobile robots is under development. In addition, as a future technology that will further expand the scope of robot applications, we are also developing technologies for remote robot operation by humans for non-standardized tasks and troubleshooting. Specifically, we have been developing technologies to convert force information, into visual information that is easily recognized by humans, to decompose summary instructions into robot actions, and to operate multiple types of robots with a common interface.

Furthermore, the increase in the number of target devices monitored in real time with advancement of communication technology and the increase in computer capacity have led to remarkable progress in 3D simulation technology. These technologies are also being used for easy and quick on-site implementation and efficient operation. In this paper, these initiatives with actual examples are described.

2. Teaching-less Robot System
2.1. *Overview of the system*

The use of machines and robots in mass production is increasing due to a shrinking labor force caused by an aging population and declining birth rate. However, there are still challenges in implementing robots in certain production lines, such as those involved in making "Bento" boxes and ready-made dishes. These lines require expertise and time-consuming adjustments due to the soft and irregular shapes of the food products and frequent menu changes. We have developed a teaching-less robot system that utilizes technology and research data accumulated over the years [5]. This system includes arm robots, controllers, 3D vision sensors, force sensors,

a tablet PC and a master PC for control. Task assignments can be easily inputted using a tablet PC and 3D sensor camera. The 3D sensor camera senses the area around the arm robot, allowing for static environment recognition and accurate measurement of spaces. 2D vision sensors installed on the production line detect the conditions of the items moving on the line. Task assignments can be set using voice and touch inputs, and the movement path of the arm robot is displayed using augmented reality. This teaching-less robot system eliminates the need for expert knowledge and time-consuming programming, as it automatically generates optimum movements based on assigned start and endpoints. Autonomous programs generate various approaches to determine the optimum and fastest speed for movements with reducing the need for test runs and saving time. The 3D sensors monitor the production area and ensure smooth workflow. The sensors identify slanted "Bento" boxes, allowing the robot to determine their position and add ingredients accordingly. These technologies not only control the arm robot but also facilitate the expansion of production lines with multiple robots.

2.2. *Surrounding Environment Modeling*

In order to acquire the 3D model of the robot's surrounding environment, we have used our original simultaneous localization and mapping (SLAM) method which use only an inexpensive camera (Fig. 1). This camera can be used for interference detection during path generation and displayed on the operation teaching device. We have achieved processing time by simplifying 3D models and reducing data size (approximately 60% reduction).
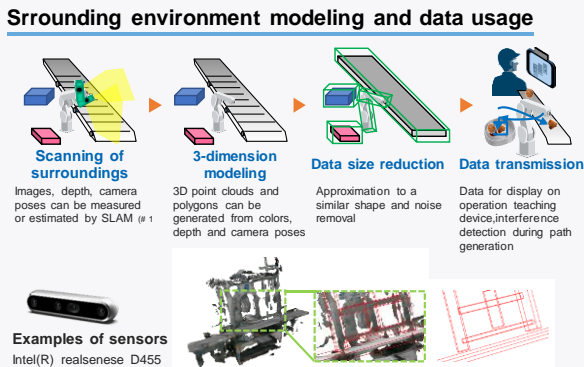
Fig. 1. Surrounding Environment Modeling

2.3. *Intuitive UI for easy robot teaching*

We have developed the intuitive user interface (UI) system which use tablet-touching and voice-controlled operation to realize easy teaching operation of robots. Various specifications such as the workpiece sizes and positions can be confirmed beforehand by checking the simulated movies using the augmented reality (AR) technology. After the motion instruction is completed,

the robot's motion can be confirmed in advance on a tablet PC, thus ensuring the safety of the operator (Fig. 2).
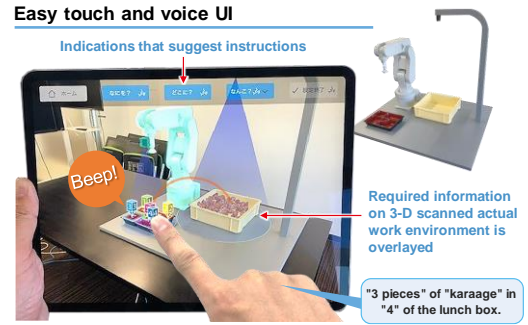
Fig. 2. Intuitive UI for easy robot teaching

2.4. *Automatic Generation of Optimum Path*

We have developed the automatic calculation method which generate optimum path without interfering to peripheral equipment by simply specifying start and end points. It can derive the avoidance route that minimizes the operating time by considering the amount of movement of each joint and allowable torque (Fig. 3).
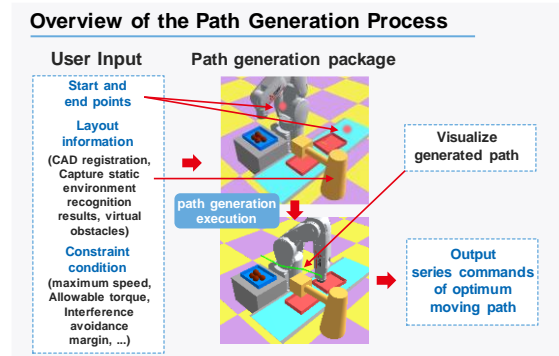
Fig. 3.　Automatic Generation of Optimum Path

2.5. *Irregular Shape Object Recognition*

To packing food objects using random bin-picking method with 3D sensor, each object is distinguished according to segmented region, and each grasping position is estimated with high accuracy at high-speed using the deep learning (DL) based AI algorithm (Fig. 4).
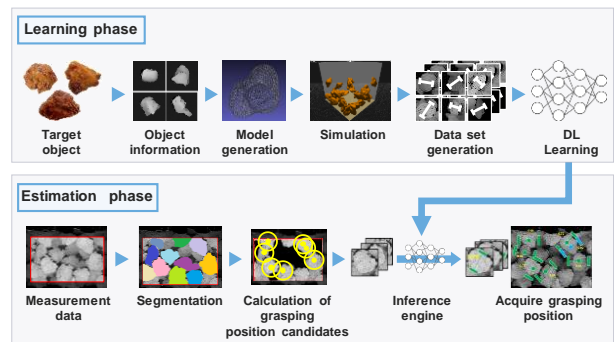
Fig. 4.　Irregular Shape Object Recognition

Prior learning of the DL network for determining optimum grasping position according to the shape of each object is operated. Then this algorithm calculates segmented region of each object using the trained DL network.

3. Robot Technologies in Service Field

3.1. *Mobile Robots for Transportation*

The demand for the Autonomous Mobile Robot (AMR) is increasing due to the decrease in the working population and the increase in demand for various types of transport (factories, distribution warehouses, commercial facilities, hospitals etc.). To meet these demands, we have been developing the system that can be redesigned to variation of transport configurations. We have also developed the series of technologies to enable robots to run safely and autonomously in response to changes in facility rules and conditions through the control system that controls multiple robots. In addition to these technologies, we have also realized autonomous vertical and horizontal movement in coordination with facilities such as elevators and access control systems [6], [7] (Fig. 5).

We have been also conducting actual experiments with mobile robots in various situations (e.g., transporting books in reading rooms, making announcements while moving, and transporting items in backyards) as a field test of mobile robots that can coexist with people (Fig. 6).
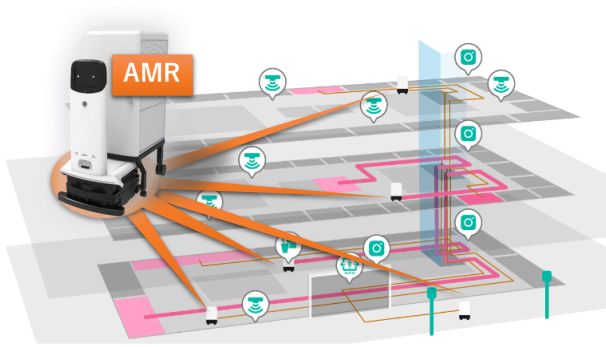
Fig. 5.   Integrated Control System for AMR
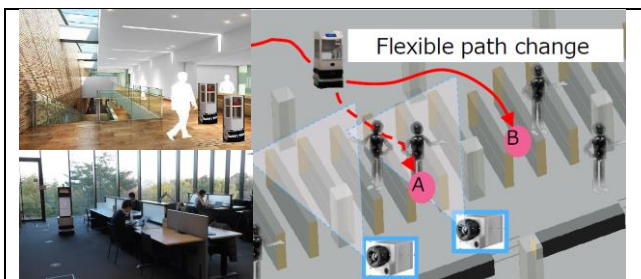
Fig. 6.   Field Test image of AMR

3.2. *Remote Robotics for Dexterous Operation*

The technologies for remotely operating robots are useful for individual operations in non-standardized tasks and for troubleshooting operations in automated tasks. Force-haptic information is often important in these tasks, but feedback of these information to the operator typically requires a complex operation system. We proposed the "Visual Haptics" architecture that converts force-haptics information into visual ones that are easily precepted by humans (Fig. 7). The cognitive process in this method utilizes visual and force-haptic cross modalities. We have showed its effectiveness through quantitative experiments [8].
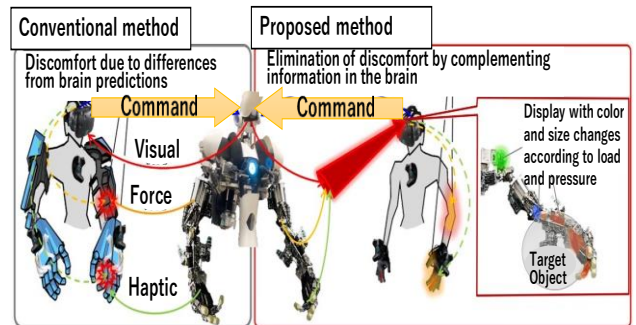
Fig. 7.   Visual Haptics" architecture

Furthermore, we have showed its effectiveness in various situations through demonstration systems [9] (Fig. 8).

Fig. 8.   Operation Test Scene

4. IoT Platforms for Data Integration

Mitsubishi Electric is ready to provide the digital transformation (DX) services including multiple IoT platforms that organically connect data from facilities, equipment, mobile applications functioning in various fields such as manufacturing, logistics, business facilities, hospitals, etc., with the aim of contributing to area value enhancement and advanced operational management. The various data collected in these distributed platforms can be used to accelerate the creation of new value-added products and seamless indoor/outdoor services (Fig. 9).

Haruhisa Okuda
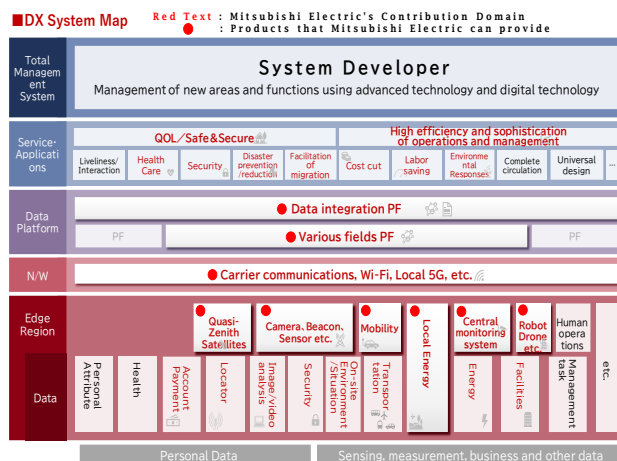
Fig. 9.   DX System including IoT Platforms

5. Conclusion

At the first in this paper, the teaching-less robot system was described, which is using intuitive I/F, optimum routes and movements generation, interference checking algorithm and series technologies for 3D random bin picking. At the second, as the robot technologies in service field, the system that realize the flexible and safety control of multiple AMR and the remote robotic technologies for dexterous operations are showed with actual examples. At the last, the integrated IoT platforms for connecting data from facilities, equipment and mobile applications functioning in various fields is described. Various solutions using these initiatives will be provided in the near future.

Acknowledgements

These technologies, demonstration system and products have been researched and developed by many colleagues in Mitsubishi Electric Corporation and collaborative research scientists. I owe them a great deal of gratitude.

References

1. H. Okuda, R. Haraguchi, Y. Domae, and K. Shiratsuchi, "Novel intelligent technologies for industrial robot in manufacturing – Architectures and applications", Proc. International Symposium on Robotics (ISR2016), pp.570-575, 2016.
2. H. Okuda, Y. Domae, Y. Kitaaki, K. Sumi, Y. Kimura, H. Takauji and S. Kaneko, "Industrial Robot System utilizing 3-D Vision Sensing for Flexible Cable Handling", 16th Korea-Japan Joint Workshop on Frontiers of Computer Vision (FCV2010), O5-2, pp.271-276, 2010.
3. K. Tanaka, H. Okuda, "Innovation on Manufacturing Generated by Intelligent Technologies", Proc. The 2018 International Conference on Artificial Life and Robotics (ICAROB 2018), pp.11-14, 2018.
4. Mitsubishi Electric Press Release, "Mitsubishi Electric and AIST Develop AI Technology for Real-time Control of FA Equipment",

5. https://www.mitsubishielectric.com/news/2021/pdf/1125.pdf, 2021-11-25.
6. Mitsubishi Electric Press Release, "Mitsubishi Electric Develops Teaching-less Robot System Technology", https://www.mitsubishielectric.com/news/2022/0228.html, 2022-2-28.
7. M. Yamazumi, T. Hirai, S. Kobori, T. Asahina, "A coordinated control between a robot and a facility for service robots in a building", Proc. of JSME Annual Conf. on Robotics and Mechatronics (ROBOMECH) 2021, 2A1-H07, 2021, DOI: 10.1299/jsmermd.2021.2A1-H07. (in Japanese)
8. M. Yamazumi, D. Mizuno, T. Koshimoto, T. Yamamoto, "Development of Mobile Robot Management System cooperated with Indoor Dynamic Map Management System", Proc. of JSME Annual Conf. on Robotics and Mechatronics (ROBOMECH)2022, 1P1-Q04. 2022, DOI:10.1299/jsmermd.2022.1P1-Q04. (in Japanese)
9. M. Haruna, N. Kawaguchi, M. Ogino, and T. Koike-Akino, "Comparison of Three Feedback Modalities for Haptics Sensation in Remote Machine Manipulation," IEEE Robot. Autom. Lett., 2021, DOI: 10.1109/LRA.2021.3070301.
10. CEATEC2022, "Evolutionary tele-operation service platform utilizing human cognitive characteristics",, 2022. (Online in Japanese)

Authors Introduction

Dr. Haruhisa Okuda

He received B.E. and M.E. from the Dept. of Precision Engineering, Kyoto University in 1991 and 1993 respectively. In 1993, he joined Mitsubishi Electric Corp., where he has been engaged in the research and development of 2-D and 3D object recognition system and AI technology for industrial robots at mainly Advanced Technology R&D Center. He received the Ph.D. degree in 2007. He has been concurrently serving as a visiting professor at the Graduate School of System Informatics, Kobe University since 2019.He is a member of the RSJ.

The Detecting Abnormal Operations in ICS Using Finite-State Machines

Pei-Wen Chou

Department of Electrical Engineering / M.S. Degree Program on Cyber-Security Intelligence,
National Cheng Kung University
No.1, University Rd., East Dist., Tainan City 701401, Taiwan

Nai-Yu Chen

Department of Electrical Engineering / M.S. Degree Program on Cyber-Security Intelligence,
National Cheng Kung University
No.1, University Rd., East Dist., Tainan City 701401, Taiwan

Jung-Shian Li

Department of Electrical Engineering / Institute of Computer and Communication Engineering,
National Cheng Kung University
No.1, University Rd., East Dist., Tainan City 701401, Taiwan

I-Hsien Liu[*]

Department of Electrical Engineering / Institute of Computer and Communication Engineering,
National Cheng Kung University
No.1, University Rd., East Dist., Tainan City 701401, Taiwan
E-mail: pwchou@cans.ee.ncku.edu.tw, nychen@cans.ee.ncku.edu.tw, jsli@ cans.ee.ncku.edu.tw,
ihliu@cans.ee.ncku.edu.tw[*]
www.ncku.edu.tw

Abstract

In 2021, a water treatment facility in Florida, USA, fell victim to an external malicious attack. in this incident, malicious actors attempted to manipulate the quantities of specific chemicals to impact water quality and safety. Given the intricacies of abnormal operation detection in Industrial Control Systems and the advantages of finite-state machine, we endeavored to apply this approach for the detection of abnormal ICS (Industrial Control System) operations. We conducted a series of tests using the dam control system cybersecurity testbed established by TWISC@NCKU, Taiwan. The results indicate that our approach effectively enhances the efficiency of identifying non-standard operational behaviors, enabling maintenance personnel to promptly identify anomalies.

Keywords: FSM, ICS Security, PLC, Dam Gate Testbed

1. Introduction

In 2021, a water treatment facility in Florida, USA, faced an external malicious attack, where attackers attempted to manipulate the quantities of specific chemicals, severely jeopardizing water quality and safety. [1] How to effectively detect abnormal operation of ICS has become an important challenge?

The integration of FSM methods involves comprehensive monitoring and analysis of system states [2], such as utilizing finite state machines. Taking equipment malfunction as an example, the awareness that its operation may deviate from normalcy prompted the design of a method for detection, employing FSM to monitor changes in PLC states. This amalgamation of state and anomaly detection enables swift identification and response to potential issues, ensuring the stable operation of the system.

This research focused on applying FSM methods to detect abnormal operations within ICS, aiming to enhance detection efficiency and accuracy. This research uses the dam control system cyber-security testbed built by TWISC@NCKU to conduct relevant research and verification. The results show that the scheme is feasible and effective, and the method can be effective at a very low cost. It can identify non-standard operating behaviors and help personnel identify abnormal situations immediately.

2. Methodology

2.1. *Programmable Logic Controller*

A Programmable Logic Controller (PLC) [3] automates control by executing instructions stored in its memory. Crucial in industries, it interfaces with systems like Finite State Machines (FSM) for more efficient monitoring.

PLC states indicate its operations during automation. "DI" and "DO" signify Digital Input and Output, reflecting signal statuses (e.g., switches). "AI" and "AO" represent Analog Input and Output, detailing continuous signal conditions. Together, they showcase the PLC's performance across various inputs and outputs.

Its architecture consists of a CPU, input, and output modules. The CPU processes logic and manages device communication. Inputs gather data from sensors, while outputs control actuators. This robust design ensures stable operation in demanding industrial environments.

2.2. *Finite-State Machine*

In recent years, Finite State Machines (FSMs) have been widely applied in various fields, including software development [2] and machine learning [4]. The strength of FSM methodology lies in its intuitive system model, aiding in the comprehension, design, and testing of intricate system behaviors. Its simplicity and scalability render it an effective tool for describing and controlling system behaviors.

In the context of anomaly detection, FSMs are employed for modeling normal system behavior by defining states and transitions. (Fig.1) This approach involves monitoring system operations, detecting state transitions deviating from expected behavior, and providing real-time alerts. The integration of FSM-based anomaly detection is widely adopted to enhance the efficiency of system behavior modeling and anomaly detection [5], highlighting the pivotal role of FSMs in comprehending system behavior and addressing challenges in anomaly detection.
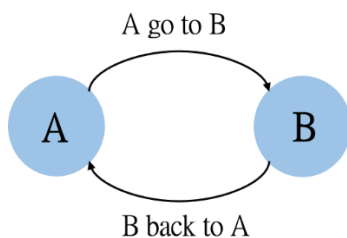
Fig. 1. Finite-State Machine

2.3. *Modbus/TCP*

Modbus is widely employed as a communication protocol in industrial environments, facilitating information exchange among industrial devices such as PLCs and sensors [6]. Despite the convenience of plaintext data transmission provided by Modbus/TCP, it inherently poses potential security risks, making it susceptible to unauthorized access and cyber attacks.

In the context of my research, using Modbus/TCP as an example, although it facilitates communication among industrial devices, careful consideration of associated security risks is crucial. The use of plaintext data transmission introduces the risk of information exposure, thereby compromising the integrity and confidentiality of the system. Therefore, during the implementation of Modbus/TCP, robust security measures must be implemented to mitigate potential risks and enhance the resilience of critical infrastructure.

Simultaneously, by utilizing Modbus commands to read the memory addresses of PLCs [7], specific states such as DI and DO can be understood, with these states stored in specific memory variable addresses. Through Modbus queries, real-time insight into the current status of the PLC can be obtained. This underscores the importance, in practical applications, of handling Modbus communication security cautiously to ensure that the system's operation remains resilient against potential risks and cyber threats.

2.4. *Critical infrastructure testbed*

The research through a series of tests conducted at the Dam Control System Cybersecurity Testbed established by TWISC@NCKU in Taiwan [8]. This simulation validation ensures that the dam system accurately and promptly responds to abnormal conditions during actual operations. The simulation of gate control scenarios on the testbed verifies the correct response of gate operations to simulated abnormal events during actual operations, ensuring the stable operation of the dam system under abnormal conditions. This integrated validation on the testbed contributes to ensuring the safety and reliability of dam systems, enhancing the efficiency of responding to abnormal events during actual operations.

3. Construct the PLC status set based on continuous discovery

3.1. *System architecture*

In response to cybersecurity and abnormal detection issues in critical infrastructure, this paper proposes the design of an experimental platform for simulating the abnormal detection system of dam gates. (Fig. 2) Each gate is operated by a Programmable Logic Controller (PLC), which is equipped with registers for accessing relevant instructions and data. Consequently, this

information can reflect the current state of the environment, such as the requirement for the abnormal indicator light to be in the off state before any operations can take place.
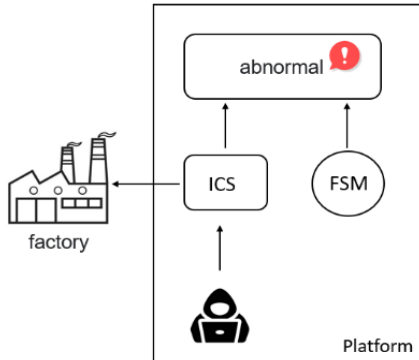
Fig. 2. Anomaly Detection Diagram

3.2. State set construction process Virtual Dam Gate Abnormal Detection testbed

The testing platform employs the Modbus communication protocol to scan the state of Programmable Logic Controllers (PLCs). The scanning interval is set at 0.1 seconds, with a 1-second pause after each scan to simulate real operating conditions. This operation is repeated infinitely to observe changes in two key variables: scan frequency and state variation. Through an infinite number of scans, we will record the results of each observation, laying the foundation for subsequent discussions to thoroughly analyze potential variations and system behaviors. The objective of this testing platform is to gain a comprehensive understanding of the abnormal detection performance of the virtual dam gate system.

```
Start:
    Setting Target PLC

While true:
    # For each iteration
    While PLC's Reg is not END:
        Reading Memory Address
    End Loop

    Save PLC's State n
    Sleep n

    If n > 1:
        Compare PLC's State n & n-1
    End If
End Loop
```

4. Experiment Results

This system effectively communicates the various states and operational stages of the gate through distinct light variations. When the gate is in remote monitoring

mode, the remote light is illuminated, indicating normal system operation. However, if floodgate discharge is required, personnel must adhere to regulations and physically attend the site, prompting a switch to on-site mode with the power light activated.

Following the commencement of operations, the ascending light is activated if gate opening is necessary, signifying the gate's upward movement. Upon reaching a specified height while fully opening the gate, the ascending light ceases, and both the on-site and power lights illuminate, denoting complete gate opening. Upon achieving full gate openness at the base, the fully open light illuminates, the ascending light extinguishes, and the descending light activates, indicating an impending gate descent. Operation Process Diagram in Fig. 3.

However, during operation, anomalies such as short circuits or system malfunctions might occur, potentially causing gate loosening or jamming during ascent, leading to an overload condition. In such instances, the system should promptly react, for instance, by activating the corresponding warning light to alert operators to perform maintenance or emergency procedures.

In summary, this system proficiently communicates the diverse statuses and phases of gate operations through distinctive light cues. Additionally, it promptly issues warnings during abnormal situations, ensuring operational safety and manageability. Experimental results as shown in Fig .4.
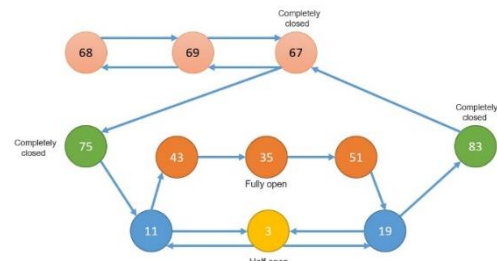
Fig. 3. Operation Process Diagram
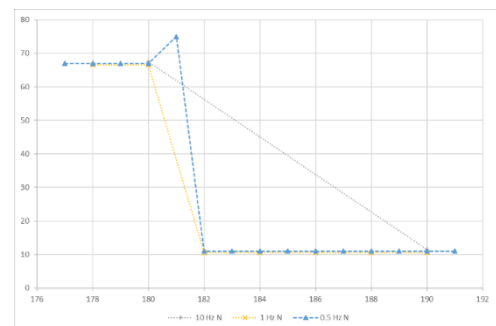
Fig. 4. Experimental results

5. Conclusion

Given the complexity of abnormal operation detection in industrial control systems (ICS), we opted to

Pei-Wen Chou, Nai-Yu Chen, Jung-Shian Li, I-Hsien Liu

employ the Finite State Machine (FSM) method for anomaly detection in ICS operations. Through a series of tests conducted at the Dam Control System Cybersecurity Testbed established by TWISC@NCKU in Taiwan, our results demonstrated the effectiveness of our approach in improving the efficiency of identifying non-standard operational behaviors, enabling maintenance personnel to promptly recognize anomalies. This study not only constitutes a technical exploration but also emphasizes addressing potential anomalies in unique scenarios. In summary, this research provides a comprehensive and effective approach to ICS security, particularly in the realm of abnormal operation detection.

Acknowledgements

References

1. CNN, "Someone tried to poison a Florida city by hacking into the water treatment system, sheriff says," 2021. [Online]. Available: https://www.cnn.com/2021/02/08/us/oldsmar-florida-hack-water-poison. [Accessed 30 Oct. 2023].
2. I. Y. Smolyakov and S. A. Belyaev, "Design of the Software Architecture for Starcraft Video Game on the Basis of Finite State Machines," 2019 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus), Moscow, Russia, 28-30 Jan., 2019.
3. T. Imanto, and A. Adriansyah. "Performance analysis of profinet network in plc-based automation system," 2nd 2020 International Conference on Broadband Communication, Wireless Sensors and Powering (BCWSP), Yogyakarta, Indonesia, 28-30 Sep. 2020.
4. S. Wang, J. F. Balarezo, S. Kandeepan, A. Al-Hourani, K. G. Chavez, and B. Rubinstein, "Machine learning in network anomaly detection: A survey," *IEEE Access*, vol. 9, pp. 152379-152396, 2021.
5. F. Farahmandi and P. Mishra, "FSM anomaly detection using formal analysis," 2017 IEEE 35th International Conference on Computer Design (ICCD), Boston, USA, 5-8 Nov., 2017.
6. Y. Wang, Y. Wang, Z. Zhu, and Q. Wang, "Modbus TCP protocol in industrial control system Research on anomaly detection method," 2021 IEEE International Conference on Electrical Engineering and Mechatronics Technology (ICEEMT), Qingdao, China, 2-4 Jul., 2021.
7. X. Li, F. Meng, and X. Zheng, "Automatic Control System of Sluice Based on PLC, MCGS and MODBUS Communication," 2021 7th Annual International Conference on Network and Information Systems for Computers (ICNISC), Guiyang, China, 23-25 Jul., 2021.
8. M.-W. Chang, J.-S. Li, and I-H. Liu, "Cyber-Physical Security Testbed for Dam Control System", *Journal of Advances in Artificial Life Robotics*, Vol. 4, No. 2, pp. 63-66, 2023.

Authors Introduction

Ms. Pei-Wen Chou

She is a postgraduate of Cloud and Network Security (CANS) Lab, Institute of Computer and Communication Engineering, National Cheng Kung University in Taiwan. She received her B.B.A. degree from the Department of Healthcare Administration and Medical Informatics, Kaohsiung Medical University, Taiwan in 2022. Her interests encompass network security, blockchain, and industrial control systems.

Ms. Nai-Yu Chen

She was born in Taichung, Taiwan in 1998. She is acquiring the master's degree in Degree Program on Cyber-Security Intelligence, National Cheng Kung University in Taiwan. She received her B.B.A. degree from the Bachelor of BioBusiness Management, National Chiayi University, Taiwan in 2021. Her interests are ICS Security, Network-Based Intrusion and PLC.

Prof. I-Hsien Liu

He is an assistant professor in Department of Electrical Engineering, National Cheng Kung University, Taiwan. He obtained his Ph.D. in 2015 in Computer and Communication Engineering from the National Cheng Kung University. He teaches cybersecurity courses and his interests are Cyber-Security, Wireless Network, Group Communication, and Reliable Transmission. He is the deputy director of Taiwan Information Security Center @ National Cheng Kung University(TWISC@NCKU).

Prof. Jung-Shian Li

He is a full Professor in the Department of Electrical Engineering, National Cheng Kung University, Taiwan. He graduated from the National Taiwan University, Taiwan, with B.S. in 1990 and M.S. degrees in 1992 in Electrical Engineering. He obtained his PhD in 1999 in Computer Science from the Technical University of Berlin, Germany. He teaches communication courses and his research interests include cybersecurity, cloud computing and network management. He is currently involved in funded research projects dealing with cybersecurity and critical infrastructure protection. He is the director of Taiwan Information Security Center @ National Cheng Kung University.

Industrial Control System State Monitor Using Blockchain Technology

Yun-Hao Chang
Department of Electrical Engineering / Institute of Computer and Communication Engineering,
National Cheng Kung University,
No.1, University Rd., East Dist., Tainan City 701401, Taiwan

Tzu-En Peng
Department of Electrical Engineering / Institute of Computer and Communication Engineering,
National Cheng Kung University,
No.1, University Rd., East Dist., Tainan City 701401, Taiwan

Jung-Shian Li
Department of Electrical Engineering / Institute of Computer and Communication Engineering,
National Cheng Kung University
No.1, University Rd., East Dist., Tainan City 701401, Taiwan

I-Hsien Liu
Department of Electrical Engineering / Institute of Computer and Communication Engineering,
National Cheng Kung University
No.1, University Rd., East Dist., Tainan City 701401, Taiwan
E-mail: yhchang@cans.ee.ncku.edu.tw, tepeng@cans.ee.ncku.edu.tw, jsli@cans.ee.ncku.edu.tw,
ihliu@cans.ee.ncku.edu.tw[*]
www.ncku.edu.tw

Abstract

This paper introduces an innovative approach to enhance data verification and security in intelligent systems through the integration of blockchain technology. The proposed method amalgamates the transparency and decentralization inherent to blockchain with the command and oversight functionalities of PLC to ensure the utmost data integrity. The devised approach synergizes the decentralized attributes of blockchain with the control capabilities of PLCs, thus establishing robust safeguards for data integrity. Through the utilization of blockchain's tamper-resistant ledger, PLCs orchestrate data interactions and enforce real-time monitoring and control. The viability and efficacy of this innovative scheme are substantiated through empirical evaluations and simulations, conclusively affirming its practicality.

Keywords: Industrial Control System, Programmable Logic Controller, Blockchain, State Monitor

1. Introduction

In recent years, with the continuous development of industrial automation, Industrial Control Systems (ICS) have played a crucial role in modern production environments [1]. These systems possess complex architectures, and their proper functioning is paramount to ensuring the stability and efficiency of the production process. However, with the ongoing advancement of digital technology, industrial environments have become more susceptible to data tampering and security threats. Our research is based on the application of blockchain technology, which, with its decentralized and transparent nature, introduces a new level of data security. By integrating blockchain into Programmable Logic Controllers (PLC) ensures the availability and integrity of the data.

2. Related Work

The related work in our study revolves around the application of Industrial Internet of Things (IIoT) and blockchain technology.

2.1. *Industrial Control Systems and Blockchain*

Researchers have explored the integration of blockchain technology with Industrial Control Systems (ICS) to enhance security and privacy within the Industrial Internet of Things (IIoT). Notable studies, such

as Z.-H. Sun et al. [2] survey of enterprise and literature reviews to identify specific industrial requirements in various supply chain scenarios, and G. Puthilibai et al. [3] proposal of a secure wireless solution based on blockchain technology for IIoT, and W. Zhou and J. Jin's development and evaluation of a distributed access control system using blockchain and smart contract technologies [4], collectively contribute to advancing the understanding and application of blockchain in the industrial domain.

2.2. *Blockchain Consensus*

Blockchain technology originated from Satoshi Nakamoto's "Bitcoin: A Peer-to-Peer Electronic Cash System" [5] and is a decentralized database technology. Various consensus protocols, including Proof of Work (PoW), Proof of Stake (PoS), Practical Byzantine Fault Tolerance (PBFT), are chosen based on blockchain use cases, performance requirements, and trust levels. The application of PoA (Table. 1) in the Industrial Internet of Things (IIoT) is notable, providing features such as a simplified block verification process, resulting in lower block generation times, and consequently, improved overall efficiency in industrial environments. X. Chen et al. measured the latency performance of Internet of Things (IoT) applications on private Ethereum blockchains, focusing on two consensus algorithms: Proof of Work (PoW) and Proof of Authority (PoA) [6]. The Study show that PoA Ethereum network has a lower Block-Oriented Latency (BOL) than the PoW one due to a simpler block verification process [7].

Table 1. PoA Overview

Energy Efficiency	PoA consumes lower energy compared to other consensuss, making it suitable for applications that demand efficiency.
Transaction Throughput	PoA provides higher transaction throughput, supporting a large volume of real-time transactions, aligning with the needs of industrial control systems.
Applicability	Particularly well-suited for industrial control systems, meeting their strict requirements for efficiency, real-time capabilities, and security.

3. Methodology

3.1. *System Design*

By utilizing Proof of Authority (PoA) blockchain (Fig. 1) [8] as the record framework for the industrial control system. PoA blockchain employs an authority node consensus mechanism where designated nodes are responsible for generating and validating new blocks.
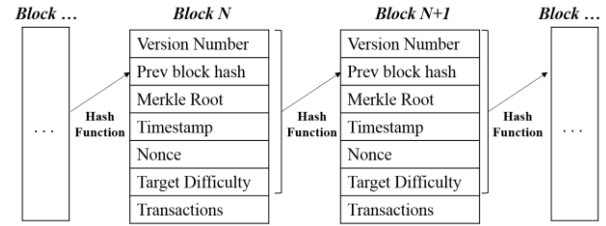
Fig.1. Blockchain Structure

Clients are also participants in the blockchain network, with their primary responsibility being the submission of transactions or information to the blockchain. Clients utilize the Remote Procedure Call (RPC) protocol to communicate with blockchain nodes. This mode of communication allows clients to send remote requests, such as transaction submission requests, while nodes process these requests, participate in the consensus mechanism, and provide feedback on the final outcomes to the clients. (Fig. 2).
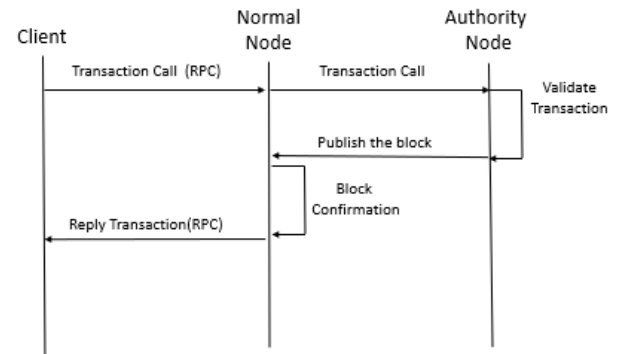
Fig.2. PoA Message Sequence Chart.

3.2. *Data Analysis*

By conducting tests in various scenarios using the same Genesis Block configuration and same mining setting to evaluate our blockchain performance [9]. Firstly, we will explore the multi-node scenario by establishing multiple nodes to simulate a decentralized environment, evaluating the system's performance in a decentralized setting. Subsequently, our focus will shift to the multi-authority scenario, where we increase the number of authority nodes to assess system performance in a multi-authority context. We will emphasize the measurement of Transactions Per Second (TPS) [10], employing different test cases and ensuring experiment repeatability to guarantee result stability. Ultimately, we will present the data distribution using box plots for intuitive comparisons, coupled with statistical analysis to gain deeper insights into the system's performance characteristics under various conditions.

4. Results and Discussion

4.1. *System Simulation*

By emulating the communication between the Human-Machine Interface (HMI) and the Programmable Logic Controller (PLC). The HMI, serving as an authoritative node, is responsible for block generation, transaction verification, as well as establishing connections with the database (DB). On the other hand, the PLC operates as a regular node, handling the synchronization and reception of transactions through client RPC. (Fig. 3).
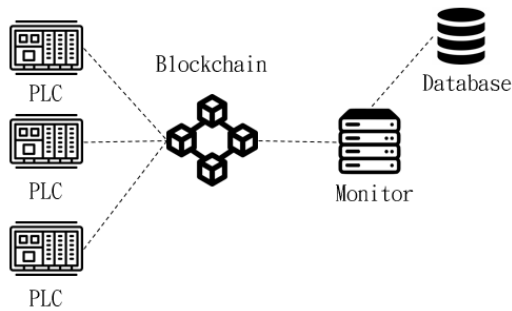
Fig. 3. Integration of blockchain with PLC.

4.2. *Experiment Result*

We measured the service rate variations of a PoA blockchain under the same conditions for other variables (Table 2). Specifically, we assessed the impact of different node counts (from 1 to 3) (Fig. 4) and different validator counts (from 1 to 3) (Fig. 5) on the Transactions Per Second (TPS). Each scenario underwent 30 test iterations, and each test iteration included 1000 transactions.

Table 2. experiment setting

Genesis block	
Period	15 secs
Epoch	30000
Difficulty	0x010
Mining Setting	
CPU	Intel(R) Core(TM) i5-8250U CPU
Thread	1
Software	
PoA blockchain	Geth/v1.10.26-stable-e5eb32ac/windows-amd64/go1.18.5
Client	Nethereum 4.18
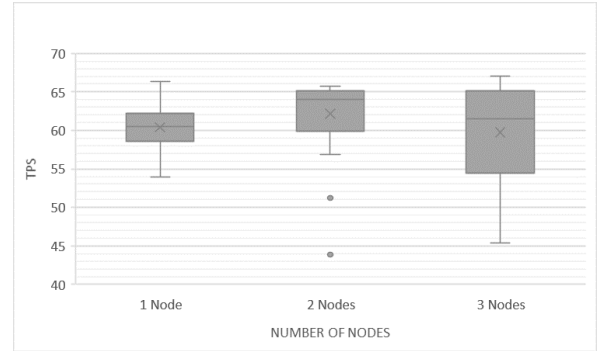
Fig. 4 TPS under different number of nodes
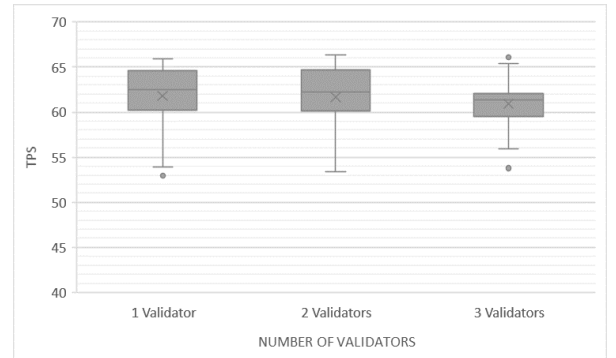
Fig. 5 TPS under different number of validators

The figures above illustrates that there is minimal variation in Transactions Per Second (TPS) across scenarios with 1-3 nodes and 1-3 validators. We hypothesize that the performance of Proof of Authority (PoA) may be more closely associated with the speed of transmission rather than the number of nodes.

5. Conclusion

In this study, we conducted practical experiments on the record-keeping functionality of the Proof of Authority (PoA) blockchain in a real-world Cyber-Physical System (CPS) environment. Additionally, we evaluated the performance of the PoA blockchain under different node counts and validator counts.

In the future, our focus will be on developing a comprehensive blockchain solution for CPS, addressing the undeniable cybersecurity needs of critical infrastructure. We will also continue to monitor the performance assessment of this solution.

Acknowledgements

Yun-Hao Chang , Tzu-En Peng, Jung-Shian Li, I-Hsien Liu

References

1. I-H. Liu, K.-M. Su, and J.-S. Li, "The Security Issue of ICS: The Use of IT Infrastructure," *Journal of Robotics, Networking and Artificial Life*, vol. 8, no. 1, p. pp. 29–32.

2. Z.-H. Sun, Z. Chen, S. Cao and X. Ming, "Potential Requirements and Opportunities of Blockchain-Based Industrial IoT in Supply Chain: A Survey," IEEE Transactions on Computational Social Systems, vol. 9, no. 5, pp. 1469-1483, 2022.

3. G. Puthilibai, T. Benil, S. Chitradevi, V. Devatarika, D. R. Ashwin Kumar and U. Padma, "Securing IIoT sensors communication using blockchain technology," 2022 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS), Chennai, India ,08 - 09 Dec., 2022.

4. W. Zhou and J. Jin, "A Blockchain-Based Access Control Framework for Secured Data Sharing in Industrial Internet," 2020 Eighth International Conference on Advanced Cloud and Big Data (CBD), Taiyuan, China, 5-6 Dec., 2020.

5. Satoshi Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," 2009. [Online]. Available: https://bitcoin.org/bitcoin.pdf. [Accessed 2 Nov. 2023].

6. X. Chen, K. Nguyen and H. Sekiya, "On the Latency Performance in Private Blockchain Networks," *IEEE Internet of Things Journal*, vol. 9, no. 19, pp. 19246-19259, 2022.

7. I-H. Liu, Y.-C. Tsai, C.-F. Li and J.-S. Li, "Cross-organizational Non-repudiation Industrial Control Log System Based on Blockchain," *Journal of Robotics,Networking and Artificial Life*, vol. 9, no. 3, pp. 240-244.

8. C.-H. Wu, I-H. Liu, J.-S. Li and C.-F. Li, "Device's Operation Tracking using Blockchain in Industrial Control System," ICAROB 2023, Oita, Japan, 9-12 Feb., 2023.

9. M. Schäffer, M. di Angelo and G. Salzer, "Performance and Scalability of Private Ethereum Blockchains," Business Process Management: Blockchain and Central and Eastern Europe Forum 2019., Vienna, Austria, 1–6 ,Sep., 2019.

10. Y.-C. Tsai, I-H. Liu and J.-S. Li, "Blockchain-based Verification Mechanism for Industrial Control System," ICAROB 2022, Oita, Japan. 20–23,Jan. , 2022.

Authors Introduction

Mr. Yun Hao Chang

He is acquiring a master's degree in the Department of Electrical Engineering /Institute of Computer and Communication Engineering, National Cheng Kung University, Taiwan. His interests are blockchain technology ,and Industrial Control Systems.

Mr. Tzu-En Peng

He is acquiring a master's degree in the Department of Electrical Engineering/Institute of Computer and Communication Engineering, National Cheng Kung University, Taiwan. He obtained his B.S. degree from the Department of Electrical Engineering, National Cheng Kung University, Taiwan in 2022. His interests are Cyber-Security and Industrial Control Systems.

Prof. Jung-Shian Li

He is a full Professor in the Department of Electrical Engineering, National Cheng Kung University, Taiwan. He graduated from the National Taiwan University, Taiwan, with B.S. in 1990 and M.S. degrees in 1992 in Electrical Engineering. He obtained his PhD in 1999 in Computer Science from the Technical University of Berlin, Germany. He teaches communication courses and his research interests include cybersecurity, cloud computing and network management. He is currently involved in funded research projects dealing with cybersecurity and critical infrastructure protection. He is the director of Taiwan Information Security Center @ National Cheng Kung University.

Prof. I-Hsien Liu

He is an assistant professor in Department of Electrical Engineering, National Cheng Kung University, Taiwan. He obtained his Ph.D. in 2015 in Computer and Communication Engineering from the National Cheng Kung University. He teaches cybersecurity courses and his interests are Cyber-Security, Wireless Network, Group Communication, and Reliable Transmission. He is the deputy director of Taiwan Information Security Center @ National Cheng Kung University(TWISC@NCKU).

Enhancing Dam Security and Water Level Alerting with Blockchain Technology

YingCheng Wu

Department of Electrical Engineering / Institute of Computer and Communication Engineering,
National Cheng Kung University
No.1, University Rd., East Dist., Tainan City 701401, Taiwan

Jung-Shian Li

Department of Electrical Engineering / Institute of Computer and Communication Engineering,
National Cheng Kung University
No.1, University Rd., East Dist., Tainan City 701401, Taiwan

Chu-Fen Li

Department of Finance, National Formosa University
No.64, Wunhua Rd., Huwei Township, Yunlin County 632301, Taiwan

I-Hsien Liu*

Department of Electrical Engineering / Institute of Computer and Communication Engineering,
National Cheng Kung University
No.1, University Rd., East Dist., Tainan City 701401, Taiwan
E-mail: ycwu@cans.ee.ncku.edu.tw, jsli@mail.ncku.edu.tw, chufenli@gmail.com, ihliu@cans.ee.ncku.edu.tw*
www.ncku.edu.tw, www.nfu.edu.tw

Abstract

Ensuring the security, monitoring, and timely alerting of water levels in dams is a major challenge. We use blockchain technology to enhance the security and monitoring of dam infrastructure, and also improving the alerting system for water level changes. The use of blockchain technology in dam infrastructure management provides a decentralized, transparent, and tamper-resistant platform for storing and managing data. This ensures the integrity and security of critical data related to dam operations and water levels. This research investigates the enhanced security, monitoring, and alerting capabilities that this integration offers, and aims to contribute to the improved security and efficiency of dam infrastructure, leading to more reliable operations and better protection against potential disasters.

Keywords: Water Dam, Blockchain, Infrastructure Security

1. Introduction

Water dams play an important role as fundamental components of critical infrastructure, fulfilling essential functions in water resource management, flood control, and energy production. Recent cyber incidents, exemplified by the reported data destruction attack on a Ukrainian power plant by Russian hackers on April 8, 2022 [1], underscore the vulnerabilities of critical infrastructure. This incident exemplifies the growing threats faced by critical infrastructure worldwide [2], emphasizing the need for innovative approaches to enhance the security, monitoring, and responsiveness of such infrastructure.

As demands on these dams increase due to population growth and climate change, ensuring security, effective monitoring, and timely response to potential issues of water dams becomes important [3]. The decentralized and tamper-resistant properties of blockchain present a possible solution to enhance the security and monitoring of dams to changing conditions. We propose the network architecture and operational workflow for the integrated system. The integration of blockchain technology aims to enhance the security, monitoring, and alerting systems of dam infrastructure, addressing current shortcomings and leading to more reliable operations and better protection against potential disasters.

2. Background

2.1. *Regulatory and Policy*

Strategies for easing security risks must consider regulatory and policy frameworks. This involves aligning security measures with industry standards, governmental

regulations, and cybersecurity best practices to ensure a comprehensive and effective approach. Regulations should contain scenarios involving abnormal water levels. Policies can ensure the implementation of advanced monitoring systems capable of detecting sudden or unexpected changes in water levels. This ensures timely responses to mitigate potential risks associated with extreme water level variations.

Regulatory frameworks must not only recognize the potential for abnormal water levels but also involve the importance of consistency in gate status and water level readings. These consistencies can arise from various factors, including technical problems, sensor inaccuracies, or even intentional manipulations. Addressing these issues in policies is an important step toward enhancing the resilience of dam infrastructure.

In the scenario of gate status and water level inconsistencies, dam operators face the challenge of accurately assessing the true conditions of the dam. A gate that is reported as closed in the control system, while physically open, can lead to an underestimation of the water level, posing risks of flooding downstream. On the contrary, a reported open gate when it is closed might result in an overestimation of water levels, potentially impacting water resource management.

Regulatory and policy considerations play an important role in addressing security risks in dam infrastructure. By incorporating events such as gate abnormal openings, water level fluctuations, equipment reading problems, and extreme weather events, regulatory frameworks contribute to a comprehensive and adaptive approach to safeguarding critical infrastructure.

2.2. *Data Integrity*

The potential for unauthorized access and manipulation of critical operational data poses a significant security risk. Unauthorized individuals or external malicious hackers may attempt to gain unauthorized access to sensitive information related to water levels or operational parameters. This external threat could lead to misleading dam operators, compromise decision-making processes, and introduce risks to the safety and functionality of the dam [4].

The risk also extends to internal employees who may intentionally manipulate operational data. This insider threat poses unique challenges as it involves individuals with legitimate access to the system. Internal data tampering can mislead dam operators, compromise decision-making, and lead to potential risks in the safety and functionality of the dam. Implementing measures to detect and prevent both external and internal data tampering is important for maintaining the integrity of

operational data [5]. Manipulating these important data sets could lead to inaccurate assessments, potentially resulting in inadequate responses to changing conditions and an increased vulnerability to operational failures.

2.3. *Blockchain Technology*

As technological advancements continue to appear in critical infrastructure, the integration of blockchain technology is seen as a possible solution for enhancing the security and transparency of dam operations. Blockchain technology enhances the security of dam infrastructure by providing a transparent and immutable record of all transactions and data related to the operation of gates, water levels, and other important parameters.

The decentralized property of blockchain ensures that data is distributed across a network of nodes, reducing the risk of a single point of failure or unauthorized access. The tamper-resistant property of blockchain ensures the integrity of data generated by sensors and monitoring systems. This feature is important in relation to intentional data manipulations that could compromise the reliability of water level readings.

3. DamChain Architecture

The DamChain architecture is designed to enhance the security and monitoring of dam infrastructure through the integration of blockchain technology. Our proposed architecture integrates the decentralized and tamper-resistant features of blockchain to build an architecture for managing important data related to dam operations and water levels.

3.1. *Network Architecture*

The network architecture of the integrated system is an important component depending on the design of the blockchain network. Choosing the right consensus mechanism is important. Moreover, incorporating blockchain technologies and a consensus mechanism can play an important role in implementing the blockchain network to the requirements of dam operations.

We designed the network architecture in Fig. 1. The network architecture demonstrates the application of blockchain technology in the management of dam infrastructure. By recording water level data and dam operations on the blockchain, the system achieves transparent information storage and protection against tampering. Operators can monitor the dam's status by observing data on the blockchain.
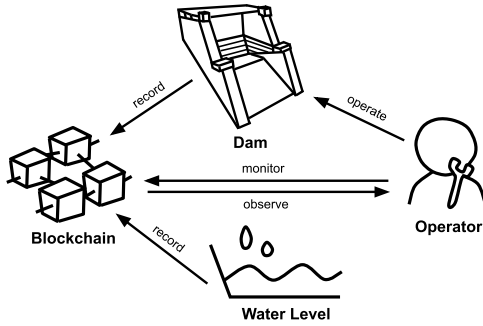
Fig. 1 Network Architecture

3.2. *Operational Workflow*

An essential aspect of the regulatory model and operational workflow is the continuous monitoring of compliance with established policies. This involves regular audits and reviews to identify any differences from the regulatory framework. The operational workflow involves the interaction between the operator, dam gate, water sensors, and the blockchain. The flowchart in Fig. 2 illustrates the process.
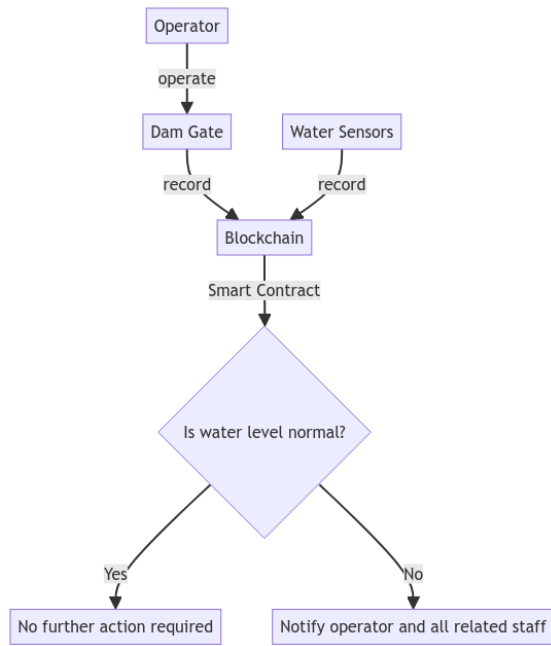
Fig. 2 Operational Workflow

In this workflow, the operator operates the dam gate, and the actions are recorded on the blockchain. Water sensors also record data, and a smart contract evaluates whether the water level is normal. If yes, no further action is required; if no, the operator and all related staff are notified.

4. Experiment

The integrated system uses Geth (Go Ethereum) as the blockchain implementation and is hosted on a system with the following specifications in Table 1.

Table 1. System Specifications

Processor	Intel Core i5-7500 CPU @ 3.40GHz
Operating System	Windows 10 Enterprise 64-bit
RAM	16GB

The experiment result is shown in Fig. 3. The result is in an average transaction throughput of 115.575 transactions per second, and shows the evaluation of the system ability to handle and record dam operations within the specified hardware and software environment.
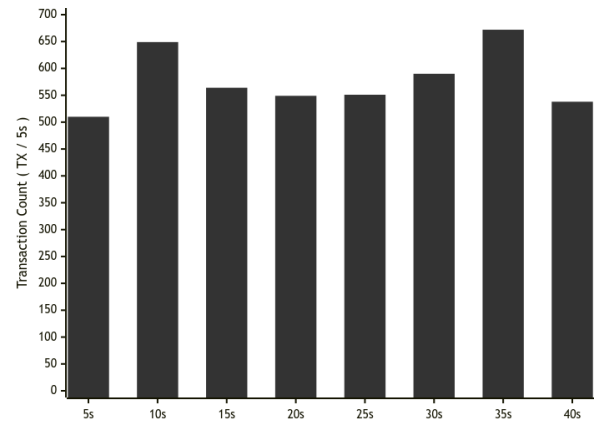
Fig. 3 Transaction Count

An additional experiment was conducted to evaluate the average response time of the system to water level alert events. The result indicates an average response time of 5.06425 seconds in Fig. 4. The experimental evaluation of the system provides reasonable insights into its performance and capabilities within a real-world dam management scenario.

Fig. 4 Response Time

YingCheng Wu, Jung-Shian Li, Chu-Fen Li, I-Hsien Liu
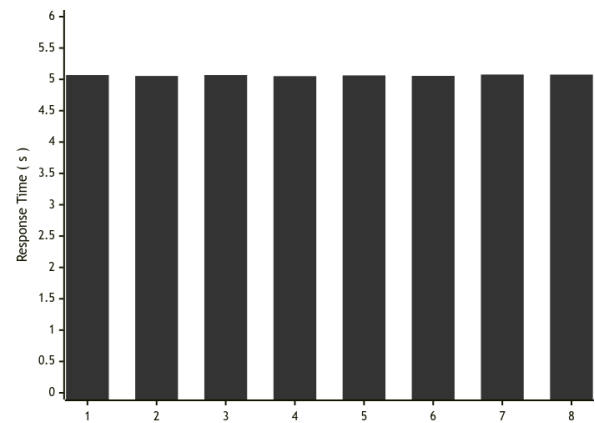
5. Conclusion

The study demonstrates the possible benefits of integrating dam infrastructure management with blockchain technology. The integration not only enhances security and monitoring but also shows the possibilities for resilient and efficient operations. The regulatory and policy frameworks underscored the importance of continuous monitoring to address unexpected challenges.

The integration system introduces how important data related to dam operations and water levels is managed. The blockchain technology ensures data integrity, providing an indisputable record of transactions. This is important in critical infrastructure scenarios to prevent unauthorized access or data tampering. Moreover, the implementation of the system addresses the need for water level alerting. These features enhance security and monitoring to dam security and water level alerting.

Acknowledgements

References

1. ESET Research, "Industroyer2: Industroyer reloaded This ICS-capable malware targets a Ukrainian energy company," 2022. [Online]. Available: https://www.welivesecurity.com/2022/04/12/industroyer2-industroyer-reloaded/. [Accessed 15 Nov 2023].
2. CNN, "Russian military-linked hackers target Ukrainian power company, investigators say," 2022 2022. [Online]. Available: https://edition.cnn.com/2022/04/12/politics/gru-russia-hackers-ukraine-power-grid/index.html. [Accessed 15 Nov 2023].
3. I-H. Liu, C.-H. Wu, J.-S. Li, C.-F. Li, "Utilizing Blockchain to Monitor the Functioning of Devices in Industrial Control Systems," *Journal of Advances in Artificial Life Robotics*, vol. 3, no. 4, pp. 205-208, 2023.
4. A. Parvizimosaed, H. Azad, D. Amyot and J. Mylopoulos, "Protection against Ransomware in Industrial Control Systems through Decentralization using Blockchain," 2023 20th Annual International Conference on Privacy, Security and Trust (PST), Copenhagen, Denmark, 21-23, Aug., 2023.
5. Y. Shah and S. Sengupta, "A survey on Classification of Cyber-attacks on IoT and IIoT devices," 2020 11th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), New York, NY, USA, 28-31, Oct., 2020.

Authors Introduction

Mr. YingCheng Wu

He is acquiring a master's degree in the Department of Electrical Engineering/Institute of Computer and Communication Engineering, National Cheng Kung University, Taiwan. He obtained his B.S. degree from the Department of Communication Engineering, National Taipei University, Taiwan in 2020. His interests are Cyber-Security and Blockchain.

Prof. Jung-Shian Li

He is a full Professor in the Department of Electrical Engineering, National Cheng Kung University, Taiwan. He graduated from the National Taiwan University, Taiwan, with B.S. in 1990 and M.S. degrees in 1992 in Electrical Engineering. He obtained his PhD in 1999 in Computer Science from the Technical University of Berlin, Germany. He teaches communication courses and his research interests include cybersecurity, cloud computing and network management. He is currently involved in funded research projects dealing with cybersecurity and critical infrastructure protection. He is the director of Taiwan Information Security Center @ National Cheng Kung University.

Prof. Chu-Fen Li

She is an Associate Professor in the Department of Finance at the National Formosa University, Taiwan. She received her PhD in information management, finance and banking from the Europa-Universität Viadrina Frankfurt, Germany. Her current research interests include intelligence finance, e-commerce security, financial technology, IoT security management, as well as financial institutions and markets. Her papers have been published in several international refereed journals such as European J. of Operational Research, J. of System and Software, International J. of Information and Management Sciences, Asia J. of Management and Humanity Sciences, and others.

Prof. I-Hsien Liu

He is an assistant professor in Department of Electrical Engineering, National Cheng Kung University, Taiwan. He obtained his Ph.D. in 2015 in Computer and Communication Engineering from the National Cheng Kung University. He teaches cybersecurity courses and his interests are Cyber-Security, Wireless Network, Group Communication, and Reliable Transmission. He is the deputy director of Taiwan Information Security Center @ National Cheng Kung University (TWISC@NCKU).

MiniDAM: A Dam Cybersecurity Toolkit

Tzu-En Peng

Department of Electrical Engineering / Institute of Computer and Communication Engineering,
National Cheng Kung University
No.1, University Rd., East Dist., Tainan City 701401, Taiwan

Meng-Wei Chang

Department of Electrical Engineering / Institute of Computer and Communication Engineering,
National Cheng Kung University
No.1, University Rd., East Dist., Tainan City 701401, Taiwan

Yun-Hao Chang

Department of Electrical Engineering / Institute of Computer and Communication Engineering,
National Cheng Kung University
No.1, University Rd., East Dist., Tainan City 701401, Taiwan

Jung-Shian Li

Department of Electrical Engineering / Institute of Computer and Communication Engineering,
National Cheng Kung University
No.1, University Rd., East Dist., Tainan City 701401, Taiwan

I-Hsien Liu

Department of Electrical Engineering / Institute of Computer and Communication Engineering,
National Cheng Kung University
No.1, University Rd., East Dist., Tainan City 701401, Taiwan
E-mail: tepeng@cans.ee.ncku.edu.tw, mwchang@cans.ee.ncku.edu.tw, yhchang@cans.ee.ncku.edu.tw,
jsli@cans.ee.ncku.edu.tw, ihliu@cans.ee.ncku.edu.tw[*]
www.ncku.edu.tw

Abstract

Testbeds, serving as simulations of real-world scenarios, are of paramount importance for research in cybersecurity related to critical infrastructure. In this paper, we aim to offer a comprehensive exploration of the MiniDAM and our testbed, introducing its physical settings based on real dam operational standards. Furthermore, a comparative analysis between the Secure Water Treatment (SWaT) testbed, MiniCPS, our testbed, and MiniDAM is presented. This paper also includes insights into dataset generation and the integration of other functionalities. The exposition of MiniDAM's features and capabilities serves as a foundation for enhancing resilience and provides valuable support for advancing research within the broader field of dam-related studies.

Keywords: Dam Testbed, Critical Infrastructure, Industrial Control System, Cyber-Physical System, Cybersecurity

1. Introduction

Critical infrastructures [1] have significantly enhanced our quality of life over years of development. Since the introduction of the Industry 4.0 [2] concept, ensuring the security of Cyber-Physical Systems (CPS) [3] has become a primary objective across various industrial sectors. CPS involves the seamless integration of computation, networking, and physical processes. Through CPS, we gain the capability to monitor both the physical processes and network traffic within a system, thereby enhancing overall system performance and resource allocation.

However, despite the advancements, numerous vulnerabilities persist within the CPS of critical infrastructures, posing potential risks to public safety [4]. Dam facilities, in particular, experience failures and security breaches annually. The cyberattack on the Bowman Avenue Dam in 2013 underscored the potential crisis posed by malicious actors targeting these systems. Conversely, Some failures are attributed to anomalies in

inflow, often triggered by extreme weather conditions. For instance, the collapse of the Laos Dam and the Sandford Dam failure in 2018 were both consequences of heavy rainfall.

The security of dam CPS can no longer be ignored after the tragedies that happened around the world. As a result, a testbed with a toolkit that contains both physical and network aspects of data for further research in a dam scenario is needed.

2. Research Background

Our research is primarily based on the concept of Cyber-Physical Systems (CPS), focusing on the investigation of the widely referenced SWaT Testbed [5] and MiniCPS [6]. Subsequently, we compare their features with our own testbed and toolkit.

2.1. *Cyber-Physical System*

CPS represents an innovative paradigm that integrates computational algorithms and physical processes to create intelligent and interconnected systems. These systems enable seamless communication and collaboration between the digital and physical worlds, allowing for real-time monitoring, control, and optimization of diverse applications.

R. Alguliyev et al. [7] illuminated the principle of CPS operation and philosophical issues of CPS raised, and also proposed a tree of attacks on CPS. J. Shi et al. [8] described and summarized the features of CPSs, then three classic cases were given with research challenges and suggestions for future work.

2.2. *SWaT Testbed & MiniCPS*

When it comes to testbeds and toolkits in the field of CPS, the SWaT Testbed and MiniCPS are frequently referenced.

2.2.1. *SWaT Testbed*

Secure Water Treatment (SWaT) is a water treatment testbed for research in the field of cybersecurity. SWaT consists of a six-stage process of water treatment with frequently used industrial components, such as Programmable Logic Controllers (PLCs) made by Allen-Bradley, Human Machine Interfaces (HMIs), Supervisory Control and Data Acquisition (SCADA) workstation, Historian, etc. The SWaT Dataset systematically generated from the testbed is also provided for CPS researchers to do further analysis and other related works.

2.2.2. *MiniCPS*

MiniCPS is a toolkit built on top of Mininet [9] to provide an extensible, reproducible research environment for network communications, control systems, and physical-layer interactions in CPS. MiniCPS was also used to model the communication and control aspects of SWaT while illuminating example applications.

MiniCPS focuses on high-fidelity network emulation and being a framework for all fields of CPS. However, the full-fledged physical process simulation was considered to be out of the scope of MiniCPS since it does not aim to be a performance simulator. Furthermore, MiniCPS can only run on Linux operating systems due to its reliance on Mininet. MiniCPS also put very little emphasis on Visualization, such as Graphical User Interface (GUI).

The issues mentioned above could pose challenges to the reproducibility of research for researchers in specific domains.

3. Dam Cybersecurity Toolkit Architecture

For the physical part of our toolkit, we utilize the PLCs that had been used in a real dam field environment in Taiwan. We also obtained the dam history log which consists of related information during the period, such as water level, opening of the gates, gates inflow, gates outflow, etc.
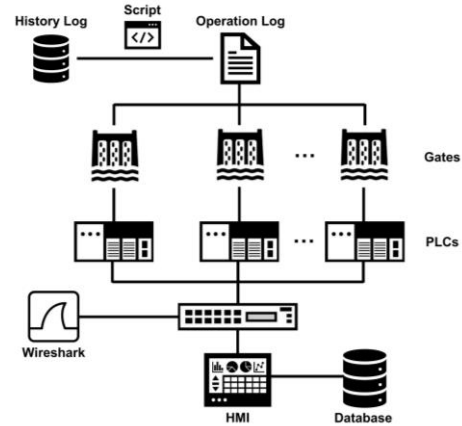
Fig. 1. Architecture of The Dam Cybersecurity Toolkit

The architecture of the dam cybersecurity toolkit is shown in Fig. 1. We further process the history log with the script based on the operational standards of the real dam to generate the operation log which defines how PLCs should work while facing the situation of the corresponding history log. As shown in Fig. 2, we also use our HMI to monitor the states of the PLCs, the overall operation is recorded in the database, and Modbus/TCP packets between HMI and PLCs are simultaneously

captured using Wireshark for CPS researchers in the dam field to conduct further analysis and experiments that need related dataset, such as machine learning [10].

Fig. 2. MiniDAM HMI

4.   Applications of The Dam Cybersecurity Toolkit
4.1. *Dam Environment Simulation*

In contrast to MiniCPS, MiniDAM places a heightened focus on Cyber-Physical Systems (CPS) within dam field environments. Leveraging pertinent Programmable Logic Controllers (PLCs) and historical logs derived from actual usage in dam fields, we meticulously construct a tailored Graphical User Interface (GUI) to facilitate operational tasks. This approach enables us to bridge the gap between theoretical models and real-world dam scenarios. As a result, we generate a diverse array of datasets corresponding to various operational scenarios. These datasets serve as valuable resources for dam domain researchers, offering them rich materials for in-depth analysis and exploration of the intricacies within dam environments.

4.2. *Dataset Generation*

In addition to recording related information in the database during operation. Modbus/TCP packets are also captured using Wireshark. As shown in Fig. 3, HMI and PLCs communicate with each other using Modbus/TCP protocol. We can generate datasets for different non-normal scenarios by modifying the logic within the script.

As illustrated in Fig. 4, due to significant variations in water flow in Taiwan during the summer season, a comparison of the history log and dataset regarding the water capacity difference over the period of the summer of 2017 was conducted within a normal scenario defined according to operational standards. We observed that the dataset and history log exhibit similar characteristics, demonstrating the validity of our operation log and the similarity between the dataset and the actual field.
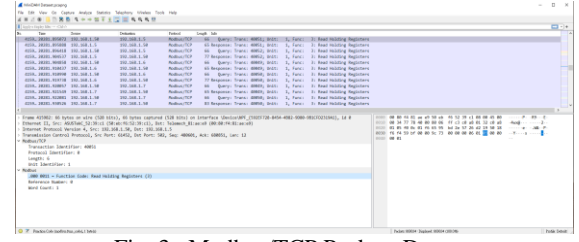
Fig. 3. Modbus/TCP Packets Dataset
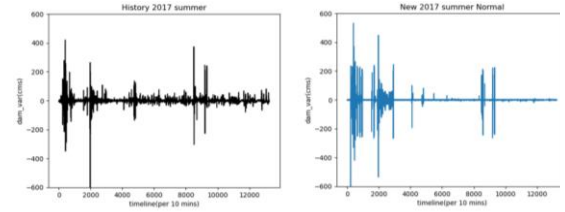
Fig. 4. Comparison of The History Log and Dataset regarding The Water Capacity Difference over The Period of 2017 Summer in a Normal Scenario

5. Conclusion

In this study, we introduce MiniDAM's physical settings based on the real dam operational standards and how our toolkit differs from MiniCPS. Furthermore, we illustrate the features of MiniDAM, including dataset generation.

In the future, we are heading to build a comprehensive CPS interface for the dam operations to extend the testbed's [11] functions. Thoroughly define how virtual and physical devices interface with our toolkit, taking into account hydrological information for broader watershed analysis.

Acknowledgment

References

1. W. Liu and Z. Song, "Review of studies on the resilience of urban critical infrastructure networks," *Reliability Engineering & System Safety,* vol. 193, 106617, 2020.
2. M. Ghobakhloo, "Industry 4.0, digitization, and opportunities for sustainability," *Journal of Cleaner Production,* vol. 252, 119869, 2020.
3. E. A. Lee, "CPS foundations," DAC '10: Proceedings of the 47th Design Automation Conference, Anaheim, California, USA, 12-18 Jun., 2010.
4. J. M. Taylor and H. R. Sharif, "Security challenges and methods for protecting critical infrastructure cyber-physical systems," 2017 International Conference on Selected Topics in

Mobile and Wireless Networking (MoWNeT), Avignon, France, 17-19 May, 2017.

5. A. P. Mathur and N. O. Tippenhauer, "SWaT: a water treatment testbed for research and training on ICS security," 2016 International Workshop on Cyber-physical Systems for Smart Water Networks (CySWater), Vienna, Austria, 11-11 Apr., 2016.

6. D. Antonioli and N. O. Tippenhauer, "MiniCPS: A Toolkit for Security Research on CPS Networks," CPS-SPC '15: Proceedings of the First ACM Workshop on Cyber-Physical Systems-Security and/or PrivaCy, Denver, Colorado, USA, 12-16 Oct., 2015.

7. R. Alguliyev, Y. Imamverdiyev and L. Sukhostat, "Cyber-physical systems and their security issues," *Computers in Industry,* vol. 100, pp. 212 - 223, 2018.

8. J. Shi, J. Wan, H. Yan and H. Suo, "A survey of Cyber-Physical Systems," 2011 International Conference on Wireless Communications and Signal Processing (WCSP), Nanjing, China, 09-11 Nov., 2011.

9. Mininet, "Mininet: An Instant Virtual Network on your Laptop (or other PC)," 2023. [Online]. Available: http://mininet.org/. [Accessed 10 12 2023].

10. O. Yavanoglu and M. Aydos, "A review on cyber security datasets for machine learning algorithms," 2017 IEEE International Conference on Big Data (Big Data), Boston, MA, USA, 11-14 Dec., 2017.

11. M.-W. Chang, J.-S. Li, I-H. Liu, 2023, "Cyber-Physical Security Testbed for Dam Control System", *Journal of Advances in Artificial Life Robotics*, Vol. 4, No. 2, pp. 63-66.

Authors Introduction

Mr. Tzu-En Peng

He is acquiring a master's degree in the Department of Electrical Engineering/Institute of Computer and Communication Engineering, National Cheng Kung University, Taiwan. He obtained his B.S. degree from the Department of Electrical Engineering, National Cheng Kung University, Taiwan in 2022. His interests are Cyber-Security and Industrial Control Systems.

Mr. Meng-Wei Chang

He was born in Pingtung, Taiwan in 1997. He is acquiring the master's degree in Department of Electrical Engineering/Institute of Computer and Communication Engineering, National Cheng Kung University in Taiwan. He received his B.S. degree from the Department of Physics, National Taiwan Normal University, Taiwan in 2021. His interests are Cyber-Security and ICS Security.

Mr. Yun-Hao Chang

He is acquiring a master's degree in the Department of Electrical Engineering /Institute of Computer and Communication Engineering, National Cheng Kung University, Taiwan. His interests are blockchain technology and Industrial Control Systems.

Prof. Jung-Shian Li

He is a full Professor in the Department of Electrical Engineering, National Cheng Kung University, Taiwan. He graduated from the National Taiwan University, Taiwan, with B.S. in 1990 and M.S. degrees in 1992 in Electrical Engineering. He obtained his PhD in 1999 in Computer Science from the Technical University of Berlin, Germany. He teaches communication courses and his research interests include cybersecurity, cloud computing and network management. He is currently involved in funded research projects dealing with cybersecurity and critical infrastructure protection. He is the director of Taiwan Information Security Center @ National Cheng Kung University.

Prof. I-Hsien Liu

He is an assistant professor in Department of Electrical Engineering, National Cheng Kung University, Taiwan. He obtained his Ph.D. in 2015 in Computer and Communication Engineering from the National Cheng Kung University. He teaches cybersecurity courses and his interests are Cyber-Security, Wireless Network, Group Communication, and Reliable Transmission. He is the deputy director of Taiwan Information Security Center @ National Cheng Kung University(TWISC@NCKU).

A Case Study of Network-Based Intrusion Detection System Deployment in Industrial Control Systems with Network Isolation

Nai-Yu Chen
Department of Electrical Engineering / Degree Program on Cyber-Security Intelligence,
National Cheng Kung University
No.1, University Rd., East Dist., Tainan City 701401, Taiwan

Pei-Wen Chou
Department of Electrical Engineering / Degree Program on Cyber-Security Intelligence,
National Cheng Kung University
No.1, University Rd., East Dist., Tainan City 701401, Taiwan

Jung-Shian Li
Department of Electrical Engineering / Institute of Computer and Communication Engineering,
National Cheng Kung University
No.1, University Rd., East Dist., Tainan City 701401, Taiwan

I-Hsien Liu
Department of Electrical Engineering / Institute of Computer and Communication Engineering,
National Cheng Kung University
No.1, University Rd., East Dist., Tainan City 701401, Taiwan
E- mail: nychen@cans.ee.ncku.edu.tw, pwchou@cans.ee.ncku.edu.tw, jsli@ cans.ee.ncku.edu.tw,
ihliu@cans.ee.ncku.edu.tw[*]
www.ncku.edu.tw

Abstract

Deploying intrusion detection systems is a common cybersecurity measure, and intrusion detection systems typically operate at the ports of gateways. In critical infrastructure, industrial control systems often employ network isolation strategies, lacking the role of gateways. This research primarily explores the deployment of the Snort intrusion detection system in such an environment, combined with specific OT rules. Validation is conducted using the cybersecurity testbed of the dam control system established by TWISC@NCKU in Taiwan. The results indicate that by employing our proposed approach, it is possible to effectively detect abnormal network traffic, addressing the common issue of inadequate monitoring in environments with network isolation.

Keywords: Industrial Control Systems, Critical infrastructure, Network Isolation, Network-Based Intrusion Detection System

1. Introduction

The cybersecurity landscape in critical infrastructure, particularly Industrial Control Systems (ICS), faces escalating threats [1], exemplified by the recent CISA AA22-103A alert. This alert, prompted by internal ICS intrusion cases, underscores operators' vulnerability to targeted attacks, emphasizing the need for robust cybersecurity measures [2]. Deploying Intrusion Detection Systems (IDS) is standard, but challenges arise in ICS scenarios with network isolation policies.

This study delves into deploying Network-Based Intrusion Detection System (NIDS), focusing on the Snort intrusion detection system within network-isolated ICS [3]. Departing from conventional approaches, it addresses challenges posed by network isolation, offering nuanced network security tailored to ICS requirements. Snort, typically installed at gateway, assumes a novel role within network-isolated ICS. The research is motivated by the recognition that existing IDSs may fall short in effectively monitoring ICS within network isolation, prompting a reevaluation of intrusion detection strategies.

To validate the approach, this study leverages the TWISC@NCKU dam control systems cybersecurity testbed in Taiwan, providing insights into real-world applicability. The research's significance lies in addressing insufficient monitoring in network-isolated

environments. Developing Snort rules to detect "Close PLC Gate Controller" commands enhances ICS's capacity to defend against potential threats, fortifying critical infrastructure against evolving cyber threats. This introduction establishes the groundwork for exploring NIDS deployment in network-isolated ICSs, aiming to provide effective and tailored cybersecurity solutions for safeguarding ICSs.

2. Background
2.1. Isolation of Industrial Control Networks

This study explores ICS cybersecurity through Purdue Enterprise Reference Architecture (PERA) [4], renowned for its comprehensive approach. PERA's distinctiveness lies in its lack of a gateway component, challenging conventional norms in critical infrastructure. This absence leads to a nuanced examination of network isolation strategies in ICS. While considering the Industrial Demilitarized Zone (IDMZ), its inapplicability due to the absence of a gateway is acknowledged. Consequently, our research underscores the need to enhance security by introducing NIDSs between the first and third layers of the PERA. Integrating PERA into intrusion detection improves understanding, optimizes deployment, and strengthens resilience against inappropriate network behaviors in industrial settings.

2.2. Network-Based Intrusion Detection System

Network-Based Intrusion Detection Systems (NIDS) are indispensable for cybersecurity, providing vigilant monitoring and timely alerts against potential intrusions in network traffic. The deployment of intrusion detection in ICSs, where network isolation is prevalent, poses a significant challenge. This study investigates the viability of integrating Snort, an open-source Intrusion Detection System, with tailored Operational Technology (OT) rules for such environments. Our methodology seamlessly incorporates Snort into ICSs, effectively addressing the complexities introduced by network isolation. Operating on signature and rule-based matching, Snort adeptly identifies abnormal network traffic and potential threats, solidifying its role as a reliable NIDS tool for ICSs in network-isolated environments. This integration significantly enhances cybersecurity by enabling the detection and response to potential intrusions within the constraints of network isolation.

2.3. Internal ICS Attack Scenarios

This study explores NIDSs deployment in Industrial Control Systems (ICS) within network-isolated environments. Our investigation identifies vulnerabilities, particularly in instances of command injection and IDE access attacks [5], posing risks to the confidentiality, integrity, and availability of ICS.

Examining NIDS deployment in ICS reveals internal attack methodologies. The TWISC@NCKU collaboration establishes a secure gateway control system testing platform at the third layer, emphasizing the critical role of sensor readings on Human-Machine Interface (HMI) in dam operations. Our research introduces cross-layer attack methodologies tailored to dam operations, disrupting sensor readings through Modbus/TCP packet injections to Programmable Logic Controllers (PLCs). Focusing on countering malicious shutdowns, specifically targeting PLC controllers, we develop security measures and IDSs to fortify ICS against evolving threats in complex network environments. This study offers practical solutions to enhance ICSs' security.

3. ICS Network Intrusion Detection Methods

To explore the deployment of NIDS in ICS, we developed a comprehensive framework encompassing SNORT deployment, PLC, NIDS, and HMI establishment. Initially, we utilized Twido to construct an HMI, offering monitoring and control for the PLC (TWDLCAE40DRF). To simulate an attack, we deliberately shutdown the controllers of PLC and employed Wireshark to capture Modbus packets. Subsequently, SNORT was implemented with a focus on configuring it as a NIDS.

NIDS is intended to be deployed at the third layer, specifically in the operational management segment, with the database at its core storing information displayed on the HMI. In this scenario, the objective of NIDS is to monitor the operational management layer of ICS to detect potential intrusion threats.
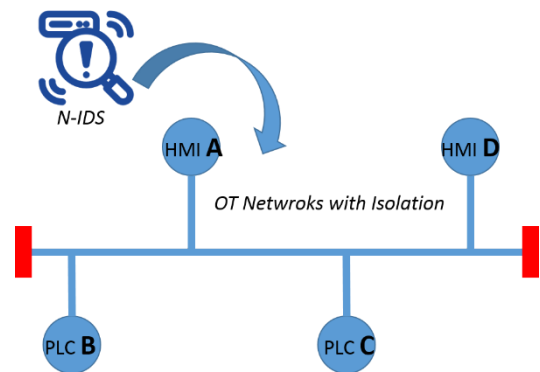
Fig. 1. NIDS, PLC, HMI Integration Architecture

During the integration of PLC and NIDS, we ensured seamless communication and monitored PLC communication to identify anomalies, with a particular emphasis on effectively detecting abnormal network behavior in an isolated network environment (Fig. 1). This comprehensive experimental procedure aims to assess the effectiveness of SNORT in ICS, contributing

to an improved understanding of cybersecurity enhancement in network-isolated environments.

In the pursuit of validating our proposed approach, we leverage the OT testbed provided by TWISC@NCKU. This cybersecurity testbed, specifically designed for dam control systems, serves as the foundation for our experiments. Its realistic representation of industrial control scenarios enables a practical examination of SNORT's deployment within a network-isolated environment.

Our investigation extends to the formulation of NIDS rules, particularly focusing on the integration of SNORT with tailored OT rules. The development of rules aims to detect attack techniques, with a special emphasis on countering malicious shutdowns, specifically targeting PLC controllers [6] (Fig. 2). This strategic development of rules contributes to fortifying ICS against evolving threats, thereby enhancing overall system security.

```
alert tcp any any -> any 502
(msg:"Detected Modbus/TCP Packet with
Data '41ff00'"; content:"|00 00 00 00
00 06 ff 90|"; depth: 8;
content:"41ff00"; sid:10000004;)
```
Fig. 2. Snort Rule for PLC Shutdown Detection

4. Experiment

In the experimental phase, we meticulously configured simulation parameters and established the experimental environment, focusing on realistic attack scenarios targeting Programmable Logic Controllers (PLCs). We simulated a shutdown attack on the controllers of PLC, capturing Modbus packets using Wireshark. Snort was configured as a NIDS with specific OT rules [7]. All experiments were conducted in the realistic dam control system testbed provided by TWISC@NCKU at National Cheng Kung University in Taiwan, ensuring authentic and reliable results. The configuration aimed to evaluate Snort's deployment effectiveness in industrial control systems with network isolation.

Utilizing the established experimental environment, a total of 30 experiments were conducted, each lasting 20 minutes and involving 0 to 20 instances of attacks. The attacks followed the methodologies outlined in Section 2.3. Background traffic, consisting of normal HMI [8] and PLC communication, was present in the environment (Fig. 3). The results indicated that, with an average of 4560 experimental packets, approximately 6.5 attacks were detected, resulting in a detection rate of 99.4%. The lowest detection rate, around 83%, occurred when interference devices were activated. In conclusion, there

is a 99.4% probability of detecting attack packets, demonstrating the robustness of the proposed approach.
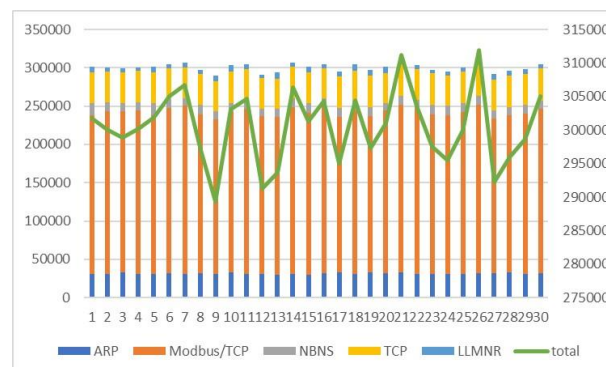
Fig. 3. Background Traffic in the Environment

5. Conclusion

This study delves into the deployment of NIDS, with a specific focus on Snort, in network-isolated ICS. Traditional intrusion detection systems often face challenges in ICS scenarios with network isolation, where they typically operate at gateway ports. Leveraging the cybersecurity testbed for dam control systems at TWISC@NCKU in Taiwan, our research validates Snort's efficacy in detecting abnormal network traffic, addressing the common issue of insufficient monitoring in network-isolated environments.

In the pursuit of enhancing the overall security of network-isolated ICS, this study emphasizes the necessity of adapting intrusion detection strategies to the unique challenges posed by critical infrastructure. Looking ahead, future research aims to broaden the scope by exploring additional attack scenarios. Due to current limitations in experimental settings, there are plans to incorporate Engineering Working Station (EWS) and introduce common IT attack scenarios to assess their effectiveness. This expansion will provide a more comprehensive understanding of the capabilities and limitations of intrusion detection systems in safeguarding network-isolated ICS against a diverse range of cyber threats.

Acknowledgements

References

1. E. Abdulova and A. Kalashnikov, "Categorization and Criticality Assessment of Facilities of Critical Infrastructure," 2022 15th International Conference

Management of large-scale system development (MLSD), Moscow, Russian, 26-28 Sep. 2022.

2. P. Hu, B. Yang, D. Wang, Q. Wang, K. Meng, Y. Wang and Z. Chen, "Research on Cybersecurity Strategy and Key Technology of the Wind Farms' Industrial Control System," 2021 IEEE International Conference on Electrical Engineering and Mechatronics Technology (ICEEMT), Qingdao, China, 02-04 July 2021.

3. D. Zhang and J. Wang, "Research on Security Protection Method of Industrial Control Boundary Network," 2021 IEEE Conference on Telecommunications, Optics and Computer Science (TOCS), Shenyang, China, 10-11 December 2021.

4. D. He, A. Lobov, L. E. G. Moctezumas and J. L. M. Lastra, "An approach to use PERA in Enterprise Modeling for industrial systems," IECON 2012 - 38th Annual Conference on IEEE Industrial Electronics Society, Montreal, QC, Canada, 25-28 October 2012.

5. I-H. Liu, K.-M. Su and J.-S. Li, 2021, "The Security Issue of ICS: The Use of IT Infrastructure," Journal of Robotics, Networking and Artificial Life, Vol. 8, No. 1, pp. 29-32.

6. E. R. Alphonsus and M. O. Abdullah, "A review on the applications of programmable logic controllers (PLCs)," Renewable and Sustainable Rnergy Reviews, Vol. 60, pp. 1185-1205, 2016.

7. J. Luswata, P. Zavarsky, B. Swar and D. Zvabva, "Analysis of SCADA Security Using Penetration Testing: A Case Study on Modbus TCP Protocol," 2018 29th Biennial Symposium on Communications (BSC), Toronto, ON, Canada, 06-07 June 2018.

8. P. Papcun, E. Kajati and J. Koziorek, "Human Machine Interface in Concept of Industry 4.0," 2018 World Symposium on Digital Intelligence for Systems and Machines (DISA), 23-25 August 2018.

Authors Introduction

Ms. Nai-Yu Chen

She is a postgraduate of Cloud and Network Security (CANS) Lab, Institute of Computer and Communication Engineering, National Cheng Kung University in Taiwan. She received her B.B.A. degree from the Bachelor of BioBusiness Management, National Chiayi University, Taiwan in 2021. Her interests are ICS Security and Network-Based Intrusion.

Ms. Pei-Wen Chou

She is a postgraduate of Cloud and Network Security (CANS) Lab, Institute of Computer and Communication Engineering, National Cheng Kung University in Taiwan. She received her B.B.A. degree from the Department of Healthcare Administration and Medical Informatics, Kaohsiung Medical University, Taiwan in 2022. Her interests encompass network security, blockchain, and industrial control systems.

Prof. Jung-Shian Li

He is a full Professor in the Department of Electrical Engineering, National Cheng Kung University, Taiwan. He graduated from the National Taiwan University, Taiwan, with B.S. in 1990 and M.S. degrees in 1992 in Electrical Engineering. He obtained his PhD in 1999 in Computer Science from the Technical University of Berlin, Germany. He teaches communication courses and his research interests include cybersecurity, cloud computing and network management. He is currently involved in funded research projects dealing with cybersecurity and critical infrastructure protection. He is the director of Taiwan Information Security Center @ National Cheng Kung University.

Prof. I-Hsien Liu

He is an assistant professor in Department of Electrical Engineering, National Cheng Kung University, Taiwan. He obtained his Ph.D. in 2015 in Computer and Communication Engineering from the National Cheng Kung University. He teaches cybersecurity courses and his interests are Cyber-Security, Wireless Network, Group Communication, and Reliable Transmission. He is the deputy director of Taiwan Information Security Center @ National Cheng Kung University(TWISC@NCKU).

The AI integration service innovation model of real estate industry in Taiwan

Li-Min Chuang

The Department of International Business, Chang Jung Christian University,No. 1, Changda Rd., Gueiren District, Tainan City, 711301, Taiwan

Chih-Hung Chen

The Ph.D. Program in Business and Operations Management, College of Management,Chang Jung Christian University, No. 1, Changda Rd., Gueiren District, Tainan City, 711301, Taiwan
Email: liming@gmail.com, macuro6793@gmail.com

Abstract

This paper aims to explore the innovative model of integrating AI (Artificial Intelligence) services in the Taiwanese real estate industry. The research employs literature analysis and the Fuzzy Analytic Hierarchy Process (FAHP) as the methodological approach. A FAHP questionnaire survey was conducted among real estate professionals in the Tainan region of Taiwan. Through the calculation of relative weights among various dimensions, the study identifies key factors related to the adoption of AI-based innovations by real estate agents in Taiwan. These findings serve as crucial references for the real estate industry in transactional and operational management.

Keywords: Real Estate Agency, Artificial Intelligence (AI), Service Innovation, Fuzzy Analytic Hierarchy Process (FAHP)

1. Introduction

Real estate transactions involve not only the buyers and sellers but also a crucial intermediary—the real estate agent. These agents play pivotal roles in property transactions, serving as key figures in advertising, marketing, and negotiation processes.

Quality real estate agents leverage effective communication and professional marketing strategies to facilitate successful transactions, guiding both parties towards the common goal of closing deals.

This paper asserts that real estate agents consistently improve overall marketing performance through various innovative services. However, the study suggests that by incorporating an AI-driven integration of innovative services, it is possible to identify key factors influencing various sales strategies employed by real estate agents. The research believes that uncovering these factors can significantly contribute to the operational performance of real estate agencies.

The objectives of this study are as follows:

(1)To examine relevant literature and identify key factors in service innovation, developing primary and secondary dimensions.

(2)To integrate AI technology into the reviewed dimensions and conduct a questionnaire study using the Fuzzy Analytic Hierarchy Process (FAHP) to obtain the relative weights of four main dimensions and twelve sub-dimensions.

(3)To draw conclusions and summarize the key factors in the integration of AI-based service innovations in the Taiwanese real estate industry. These findings serve as vital references for industry operational performance and business development.

2. Literature Review

Innovation has emerged as a pivotal force propelling the development of the service economy [1]. Within the real estate industry reports, there is a growing trend to incorporate elements such as digitization, new technologies, the sharing economy, tenant relationships, as well as new services and business models [2], [3], [4].

AI has the potential to drive market capabilities, utilizing various AI methods such as customer-oriented, industry-oriented, and cross-functional integration approaches to better understand consumer demands. These methods are then applied to the marketing behaviors of companies [5]. To streamline AI-driven innovation, service organizations need to develop an integrated Dynamic Capability Framework (DCF) that focuses on the ethical application of technology, talent, organizational culture, and structure to optimize the entire service journey [5].

This paper gathers insights from domestic and international scholars and experts to explore literature related to service innovation and the application of AI in

the field of real estate. Through this review, the study synthesizes and constructs the main dimensions of service innovation, including customer engagement, business models, technology integration, and process innovation. The subsequent sections provide detailed discussions on each of these dimensions.

At the "customer engagement" level, customers are paramount in the real estate industry. Real estate agents must fully comprehend customer needs to seize crucial transaction opportunities. Therefore, the concept of customer engagement necessitates the integration of various transaction methods, considering potential impacts on company behavior. To achieve this, companies need to understand customer demands, identify the value customers seek beyond normal transaction processes, recognize customer roles, and anticipate the roles customers and the company play in interacting with each other [6].

At the "business model" level, business models have become essential means for the commercialization of innovative ventures. They are perceived to provide a framework for companies to create and extract value from innovative ideas or technological developments, emphasizing that innovation concepts or technological developments alone do not represent any "single target value." The commercialization of such concepts occurs through business models [7], [8].

At the "technology integration" level, the rapid development and widespread deployment of technology integration form the foundation for many service innovations. The significance of technology integration for service-oriented enterprises and industries, as well as the importance of service innovation, has long been recognized [9]. Understanding the four dimensions of service innovation—service concept, customer interface, service delivery system, and technology—many service innovations involve some combination of these dimensions [10].

At the "process innovation" level, revolutionary approaches to enhance business performance must address how to perceive and build businesses and how to improve them. Businesses should not be viewed from the perspective of functions, departments, or products, but rather from the standpoint of key processes. Adopting a process-oriented approach means adopting the customer's viewpoint. Processes are the structures for performing operations required to create value for customers. Therefore, a crucial measure of processes is customer satisfaction with the process outputs, as they are the ultimate arbiters of process design and ongoing performance [11].

For real estate sales, the use of technology and digitization has become a fundamental marketing method. As information technology continues to advance, an increasing number of real estate agents are incorporating AI technology into their sales promotions. This involves integrating various service innovations and streamlining operational processes to address the dynamic demands of the real estate market, thereby strengthening business operations and achieving profitability goals.

3. Constructing an AI-Integrated Service Innovation Model in the Real Estate Industry

3.1. *Research Framework*

This study synthesizes the results of focus group discussions to establish a determined research framework for an AI-integrated service innovation model in the real estate industry. The developed framework comprises four main dimensions and twelve evaluation indicators, as illustrated in (Fig. 1).

Figure 1 Research Framework
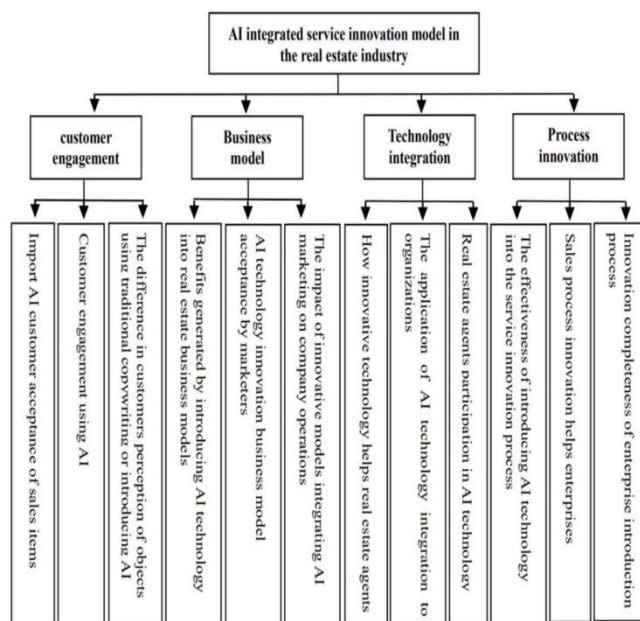
3.2. *Research Subjects*

To investigate the integration of AI technology in service innovation within the Taiwanese real estate industry, this paper focuses on real estate professionals in the Tainan region as the primary study participants. Employing purposive sampling, 12 scholars and industry experts specializing in AI and the real estate industry were selected to participate in the FAHP questionnaire. Prior to conducting the FAHP survey, the researchers provided an overview of the study's objectives to these participants and engaged willing real estate professionals in the FAHP questionnaire survey operations.

4. Empirical Evidence of AI-Integrated Service Innovation Models

In this paper, the multiplication of the four evaluation indicators by their respective dimensions results in composite weights.

As can be seen from Table 1, the most important evaluation indicators recognized by the 12 real estate industry players are "customer acceptance of AI for sales objects" (the comprehensive weight is 0.223), and "the benefit of sales process innovation to the enterprise" (the comprehensive weight is 0.173, ranking No. 2. Ranked third is "The difference in customer perceptions of objects using traditional copywriting or the introduction of AI" (the comprehensive weight is 0.101), ranking fourth is "The effectiveness of AI technology introduced into the service innovation process" (the comprehensive weight is 0.084), ranking The fifth is "The benefit of innovative technology to real estate agents" (the comprehensive weight is 0.083), and the evaluation indicators ranked sixth to twelfth are: "The acceptance of AI technology innovation business model by marketers" (the comprehensive weight is 0.067) , "Customer participation in using AI" (the comprehensive weight is 0.059), "The benefits generated by the introduction of AI technology into the real estate business model" (the comprehensive weight is 0.056), "The organizational application of AI technology integration" (the comprehensive weight is 0.054) , "Innovation completeness of enterprise introduction process" (the comprehensive weight is 0.041), "The impact of the innovative model integrating AI marketing on company operations" (the comprehensive weight is 0.030), and "The participation of real estate agents in AI technology" (the comprehensive weight is 0.029).

Table 1 The comprehensive weights for the AI-integrated service innovation model

Dimension	Weights	Evaluation indicators	Weights	Comprehensive weight
customer engagement	0.383	Import AI customer acceptance of sales items	0.582	0.223
		The difference in customers' perception of objects using traditional copywriting or introducing AI	0.263	0.101
		Customer engagement using AI	0.155	0.059
Process innovation	0.298	Sales process innovation helps enterprises	0.579	0.173
		The effectiveness of introducing AI technology into the service innovation process	0.283	0.084
		Innovation completeness of enterprise introduction process	0.138	0.041
Technology integration	0.166	How innovative technology helps real estate agents	0.502	0.083
		The application of AI technology integration to organizations	0.325	0.054
		Real estate agents' participation in AI technology	0.173	0.029
Business model	0.153	AI technology innovation business model acceptance by marketers	0.436	0.067
		Benefits generated by introducing AI technology into real estate business models	0.369	0.056
		The impact of innovative models integrating AI marketing on company operations	0.195	0.030

5. Conclusions

The four key findings of this study are listed below:

5.1. *The most important evaluation indicator for customer participation is "Customer acceptance of sales items imported into AI"*

"Customer participation" is the most important aspect considered by real estate industry players, among which "customer acceptance of AI for sales items" is the most important. For customers, no matter what kind of real estate property is needed or what sales method is used, the primary evaluation factor is whether the customer can accept and like it. It is also worth noting that no matter what kind of technology is used to promote the sale of objects, the key point is to let customers fully understand the characteristics, price, regional market and housing conditions of the object, which is the best way to achieve the final sale.

5.2. *The most important evaluation indicator of process innovation is "the benefit of sales process innovation to the enterprise"*

"Process innovation" is the second most important aspect considered by real estate industry players, among which "the benefit of sales process innovation to enterprises" is the most important. Research shows that the vast majority of real estate industry operators agree that continuous innovation in sales processes, in addition to the benefits to enterprises, It is obviously helpful, but it will also be a great help for marketers to implement business promotion. The possible reason is that with the changes of the times and in response to different customer needs, only innovating the sales process in a timely manner can effectively improve operational efficiency. In other words, real estate industry players cannot just focus on sales performance. Grasping the pulse of the market and establishing a process innovation mechanism are indispensable key factors.

5.3. *The most important evaluation indicator for technology integration is "the benefit of innovative technology to real estate agents"*

"Technology integration" is the third most important aspect considered by real estate industry players, among which "the benefit of innovative technology to real estate agents" is the most important. The main reason is that information technology is changing with each passing day, and AI is penetrating into every person. Today, the main purpose of promoting technology-based real estate brokerage is to create convenience, security, and innovation between operators and consumers. At this time, it must rely heavily on the integration and linkage of various technological

knowledge and information equipment. For example, through smart technology-related technologies, people can view houses remotely or use 360-degree panoramic views, so that they do not need to visit the house in person. Instead, they can use technology to filter suitable objects to save money. time and improve sales performance.

5.4. *The most important evaluation indicator of the business model is "marketing staff acceptance of AI technology innovation business model"*

"Business model" is an aspect that real estate industry players consider to be less important. Among which, "AI technology innovation business model marketers' acceptance" is the most important. The main reasons should be the age and acceptance of new technology applications by real estate agents and marketers. There is a large gap in levels, and not every colleague may be quick or familiar with the application mode of AI technology, and some consumers have not popularized the operation method of information technology for house viewing. There are still most consumers who can only discover the house through on-site viewing. The actual problem lies, but with the increasing popularity of smart systems and equipment, AI technology will gradually become more popular in the future, and the public will also have a certain degree of concept. The promotion of this business model will still be a key point in the future.

References

1. Rust, R. T., & Huang, M. H. (2014). The service revolution and the transformation of marketing science. Marketing Science, 33(2), 206–221.
2. Westergren, U.H., Saarikko, T. and Blomquist, T. (2017), "the IoT guide – a business guide to the internet of things", Industry Report, Department of Informatics, Umeå University.
3. Fastighetsägarna Stockholm. (2018), "Lysande utsikter för fastighetsbranschen", Industry Report,Fastighetsägarna Stockholm Publications.
4. FIBREE (2021), "Industry Report Blockchain Real Estate", industry report, FIBREE.
5. Akter S.et al.(2023), "A framework for AI-powered service innovation capability:review and agenda for future research.Technovation". Technovation 125, pp. 102768.
6. S.C. Ng, J.C. Sweeney, C. Plewa (2020) "Customer engagement: A systematic review and future research priorities" Australasian Marketing Journal, 28 (4), pp. 235-252..
7. Chesbrough, H and S Rosenbloom (2002). The role of the business model in capturing value from innovation: Evidence from Xerox Corporation's technology spinoff companies. Industrial and Corporate Change, 11(3), pp. 529–555.
8. Chesbrough, H (2010). Business model innovation: Opportunities and barriers. Long Range Planning, 43(2/3), pp. 354–363.
9. Barras, R. (1986). "Towards a Theory of Innovation in Services,"Research Policy (15), pp. 161-173.
10. Miles, I. (2008). "Patterns of Innovation in Service Industries," IBM Systems Journal, 47(1), pp. 115-128.
11. TH Davenport (1993). "Process innovation: reengineering work through information technology"

Authors Introduction

Dr. Li-Min Chuang

He received the Ph.D. degree from the National Cheng Kung University. He is currently an associate professor and department head at the Department of International Business at Chang Jung Christian University, Taiwan. His research interests focus on strategic management, innovation management, and service innovation.

Mr. Chih-Hung Chen

He received his Master's degree from the EMBA, Chang Jung Christian University, Taiwan in 2013. He is currently a Doctoral course student in Chang Jung Christian University, Taiwan.His research interests include the application of information technology to real estate industry.

A Study on Sales Patterns for Vegetable Products in Retail Stores

Yuhao Zhang[1]*, Shuangshuang Ma[2], Jiashuai Wang[3], Fengzhi Dai[4], Lijiang Zhang[5]

[1]College of Artificial Intelligence, Tianjin University of Science and Technology,

300222，China

[2]College of Light Industry Science and Engineering, Tianjin University of Science and Technology,

300222，China

[3]College of Artificial Intelligence, Tianjin University of Science and Technology,

300222，China

College of Electronic Information and Automation, Tianjin University of Science and Technology,

300222，China

[5]Xinjiang Shenhua Biotechnology Co., Ltd, China

E-mail: *a15839176622@163.com

www.tust.edu.cn

Abstract

In fresh produce supermarkets, the shelf life of vegetable products is typically short, necessitating daily restocking based on historical sales data and the formulation of a rational pricing strategy to maximize the store's profits. This paper, based on sales data for vegetable products in a particular store from July 2020 to June 2023, employs various analytical methods, including multidimensional analysis, clustering, and regression, to explore the interrelationships among different types of vegetables. Furthermore, it combines cost-plus pricing and price elasticity models to establish a pricing framework that optimizes revenue for the supermarket.

Keywords: vegetables for sale, vegetable pairing, cost plus pricing, pricing strategy

1. Introduction

Nowadays, with the continuous improvement of people's living standards, fresh merchants are opening in various places. In the sale of vegetable commodities, because the fresh-keeping period of vegetable commodities is relatively short, and the product phase becomes worse with the increase of sales time, it is of great significance for merchants to replenish according to the historical sales and demand of various types of vegetables in the past.

In the actual operation of the store, the store should consider the types of various vegetables, origin, shelf life, purchase time, season, store space and other factors to reasonably replenish different vegetables. In addition, the store should also consider the consumption time of customers, the number of types of dishes, market conditions and other factors to determine the price. The research on reasonable replenishment and pricing of vegetables has important guiding significance for the sale of vegetables in real life.

This paper will explore the distribution of different vegetable sales from different dimensions, and combined with the cost plus pricing [1] and price elasticity model [2] to explore and set a set of pricing relationships to maximize the profits of supermarkets.

Yuhao Zhang, Shuangshuang Ma, Jiashuai Wang, Fengzhi Dai, Lijiang Zhang

The rest of this paper is organized as follows. The second part studies the law of vegetable sales according to the data of previous years. Section III introduces the construction of the pricing model. In Section IV, the evaluation model and the obtained results are analyzed. Section 5 summarizes the main contents of this paper.

2. Rules of Vegetable Sales

The data statistics the flow details of vegetable products sold in a store from July 1, 2020, to June 30, 2023, which contains six vegetable categories and the corresponding single product sales volume and sales unit price. In the following, the sales rules of vegetables will be analyzed from three dimensions: quarterly, weekday and non-weekday, and freshness.

2.1. The quarterly factor

By summarizing the total sales volume of the six categories of goods in each quarter, the visualized data are presented in Fig.1.
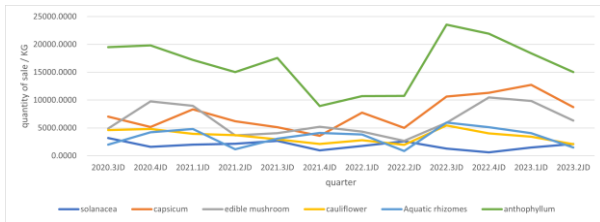
Fig.1. quarterly factor chart

As can be seen from the Fig.1, the sales volume of each vegetable category shows periodic changes over time. The sales volume of the third quarter of 2020, the third quarter of 2021 and the third quarter of 2022 to the third quarter of 2023 show the same trend. The trend between the third quarter of 2021-the third quarter of 2022 and the two cycles may be different because the COVID-19 pandemic has affected sales.

2.2. Working days and non-working days

In real life, people often have meals at work on weekdays and buy fewer vegetables every day. On non-working days, that is, Saturday and Sunday, people are more willing to buy more vegetables to eat. The average sales volume of six types of vegetables on weekdays and non-working days are shown in Fig.2.
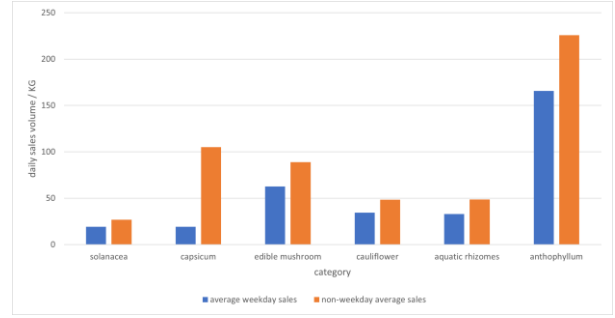
Fig.2. weekday and non-weekday factors

It can be seen from Fig.2 that each category sells well on Saturday and Sunday, and the sales volume is large, which reflects that the average sales volume of the six vegetable categories on non-working days is stronger than that on working days.

2.3. Freshness of vegetables

In a day, from morning to evening, the freshness of dishes gradually decreases. The average daily sales volume of the six vegetable categories is divided into three periods by time, and the sales volume of each vegetable category in the morning, middle and late periods is summed up, to analyze the distribution of the freshness of dishes in the store on the sales rule, The data are visualized in Fig.3.
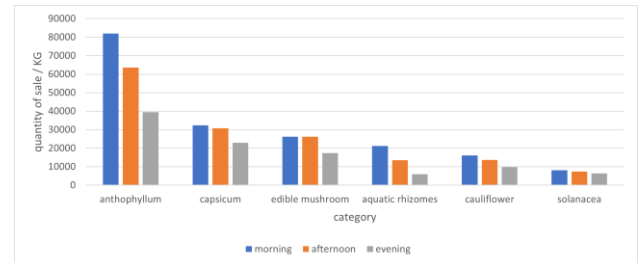
Fig.3. freshness factor map of vegetables

It can be seen from Fig.3 that the sales volume of all dishes is the largest in the morning, followed by the afternoon, and the least in the morning. This result is consistent with the actual life, and reminds the merchants to prepare all the goods in the morning, and carefully replenish the goods in the afternoon and evening.

2.4. The distribution law of sales volume of each single product

Each major category of vegetables contains a variety of single vegetable products. The single vegetable products are divided into 7 categories by K-means clustering

method [3], and the trend of the average sales volume of each category of single product over the quarter is drawn, which is shown in Fig.4.
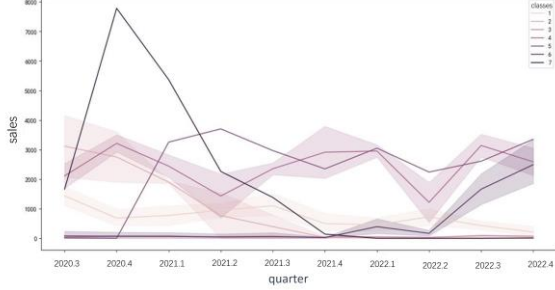
Fig.4. regular trend chart of single product sales

The average sales volume and sales volatility (standard deviation) of each category of single product are calculated, and combined with the quarterly sales trend of single product, the sales rule of single product is shown in Fig.5.

Category	Average Sales Volume/KG	Sales Volatility (Standard Deviation)	Quarterly Trend
First Category	About 663.58, at a medium level	About 529.82, relatively small	Sales volume is relatively stable, without obvious peaks or valleys
Second Category	About 754.04, also at a medium level	About 1211.53, high volatility	Sales volume has some fluctuations, with some seasonal effects
Third Category	About 54.89, significantly lower than other categories	About 88.17, lowest volatility	Sales volume is overall low and stable
Fourth Category	About 2278.61, significantly higher than other categories	About 858.75	There are several quarters with very high sales volume, possibly seasonal products
Fifth Category	About 2347.03, also a high sales volume category	About 1203.68, high volatility	Sales volume fluctuates greatly in different quarters
Sixth Category	About 739.19, at a medium level	About 1029.03, also high volatility	Sales volume has some fluctuations
Seventh Category	About 1598.93, high sales volume	About 2503.23, highest volatility	There are several quarters with very high sales volume, large seasonal impact

Fig.5. summary of single product sales rules

3. Building A Pricing Model

The purpose of establishing the pricing model is to maximize the profit obtained by the merchants [1]. Firstly, the purchase price and sales volume of each category of vegetables in the future should be predicted according to the previous data, and then the cost markup rate and elastic demand should be added for modeling.

3.1. Cost-plus pricing

Cost-plus pricing is a method to determine the price of a product based on the cost of the product, plus a certain profit margin [2]. This method is simple and easy to implement, which can ensure that the costs and profits of enterprises are compensated, and is also conducive to

maintaining the stability of prices. The calculation formula of the cost-plus pricing method is as follows.

$$\text{Price} = \text{unit cost} + \text{unit cost} \times \text{cost margin} \quad (1)$$

3.2. Price elastic demand

Price elastic demand refers to the response degree of quantity demanded to price changes, which is usually expressed by the ratio of the percentage change in quantity demanded to the percentage change in price, that is, the price elasticity coefficient of demand.

3.3. Buildup of model

The impact of establishing the selling price and cost markup rate on sales volume for each vegetable category is as follows.

$$Q_i = \alpha_i + \beta_{1i}P_i + \beta_{2i}R_i + \epsilon_i \quad (2)$$

Here, α_i is the intercept term. β_{1i} and β_{2i} are the coefficients representing the influence of selling price P_i and cost markup rate R_i on sales volume Q_i, and ϵ_i is the error term.

The price elasticity, denoted as $\beta_{2i}=E_i$, is defined as:

$$E_i = \frac{\Delta Q_i/Q_i}{\Delta P_i/P_i} \quad (3)$$

To maximize the profit for supermarkets, we establish a constraint equation for profit Π.

$$\Pi = \left(P_i - C_i \times (1 + L)\right) \times (E + 1) \times D \quad (4)$$

Here, P_i is the selling price, C_i is the cost, L is the loss rate, E_i is the price elasticity's impact on sales volume, and D is the demand quantity.

4. Interpretation of Result

According to the above model construction, six types of vegetable commodity data are brought into the model and calculated with Python to obtain the final answer.

4.1. Evaluation of the model

Advantages:

(1) By comprehensively considering the relationship between sales volume, cost markup and price elastic demand, merchants can maximize profits. This comprehensive consideration can help merchants more comprehensively consider the impact of various factors when formulating pricing strategies.

(2) The use of historical data for forecasting can provide certain reference basis to help merchants make a reasonable estimate of future sales volume and purchase price.

(3) The cost-plus pricing method is simple and easy to implement and can ensure that the costs and profits of enterprises are compensated, which is conducive to the stability of prices.

(4) Considering price elastic demand, that is, the sensitivity of demand to price, can help merchants better understand the changes in market demand and make corresponding adjustments when setting prices.

Disadvantages:

(1) The model excludes the influence of special factors such as weather and source of goods, which may have an important impact on the sales volume and purchase price. Therefore, in practical applications, it may be necessary to further consider the influence of these factors.

(2) The model does not consider the pricing strategies of competitors and the impact of market changes, which may limit the applicability of the model in the real market environment.

5. Conclusion

Through THE ANALYSIS OF THE flow records of vegetable commodities in a store, the sales relationship between different vegetable categories is explored, and the basic pricing model is obtained, which makes the store maximize the profit, which will help the store replenishment and pricing.

References

1. Cyril Tomkins, Making sense of cost-plus transfer prices where there are imperfect intermediate good markets by a 'pragmatic-analytical' perspective, *Management Accounting Research*, 1990, 1(3): 199-216.
2. Paik J, Baek S-J, Kim J-W, et al., Influence of Social Overhead Capital Facilities on Housing Prices Using Machine Learning, *Applied Sciences*, 2023, 13(19):10732.
3. Wang Qian, Wang Cheng, Feng Zhenyuan, et al., A Review of the K-means Clustering Algorithm, *Electronic Design Engineering*, 2012, 20(07): 21-24.

Authors Introduction

Mr. Yuhao Zhang

He is an undergraduate majoring in artificial intelligence at Tianjin University of Science and Technology. He learns about electronic information and artificial intelligence, and currently focuses on digital image processing, deep learning and pose estimation.

Ms. Shuangshuang Ma

She is an undergraduate in packaging engineering at Tianjin University of Science and Technology, and her current research direction is packaging machinery and intelligent packaging technology.

Mr. Xuran Wang

He passed The National College Entrance Examination in 2022 and studied applied chemistry at Tianjin University of Science and Technology. He is studying the molecular design and synthesis of porous inorganic-organic framework materials.

Dr. Fengzhi Dai

He received an M.E. and Doctor of Engineering (PhD) from the Beijing Institute of Technology, China in 1998 and Oita University, Japan in 2004 respectively. His main research interests are artificial intelligence, pattern recognition and robotics. He worked in National Institute of Technology, Matsue College, Japan from 2003 to 2009. Since October 2009, he has been the staff in College of Electronic Information and Automation, Tianjin University of Science and Technology, China.

Mr. Lijiang Zhang

He has been chairman of Xinjiang Shenhua Biotechnology Co., Ltd, China and Zhongnong Guoyou Technolegy (Tianjin) Co., Ltd, China since 2019. His reaearch interest is deep processing of agricultural products.

Research and Implementation of Cooperative Control for ROS Mobile Robot

Saijie Zhang, Huailin Zhao*

School of Electrical and Electronic Engineering, Shanghai Institute of Technology, Shanghai, 201418, China

*E-mail: *zhao_huailin@yahoo.com*
www.sit.edu.cn

Abstract

In this paper, the control of multiple mobile robots TurtleBot is studied for the robot operating system (ROS). A distributed control and communication platform is built to solve the problem of data transmission between multiple robots, and the configuration of robot workstations that can be monitored and remotely controlled is completed. On this basis, the SLAM mapping and autonomous navigation of the robot are completed. The SLAM mapping algorithm of Gampping is selected. With the help of Gazebo simulation platform and Rviz visualization tool, the local path planning of TEB is used to realize the autonomous navigation and multi-point navigation of the robot based on the known environment. On the basis of the constructed distributed control system and the control of a single robot, the research on the coordination of multiple robots is further carried out, and the synchronous control of the robot is realized by using a single node and multi-thread method. By changing the commands issued by the topic, the control of multiple robots for different attributes is released synchronously to achieve more accurate and more synchronous robot motion control. With the help of TF tools, multi-machine following and multi-machine formation are completed, and based on the path planning of a single robot, the navigation of multiple robots is realized.

Keywords: ROS, Multi-robot cooperation, Distributed control

1. Introduction

Robot Operating System (ROS) is an open source standard for the robotics industry. It integrates sensors, motors, and controllers into reusable modules through a distributed messaging architecture to achieve mobility, operation, navigation, and recognition tasks [1]. With the continuous development of the Robot Operating System (ROS), ROS has gradually entered various fields to complete various task requirements. However, when dealing with some complex tasks, it still seems to be somewhat inadequate, and sometimes it is necessary to cooperate with the staff to meet the requirements. Therefore, the research on multi-robot cooperative control is particularly important. The realization of multi-robot collaboration can not only cope with complex work scenarios and complete complex work requirements, but also improve work efficiency and enhance the intelligence of robots. Moreover, with the development of artificial intelligence, robots have ushered in new development opportunities. Based on the ROS multi-robot patrol algorithm simulation and benchmark test framework, the development of multi-robot systems, virtual reality and robot interaction has become popular. In the future, ROS will develop rapidly in the fields of navigation and positioning, 3D object recognition, motion planning, multi-joint manipulator motion control, machine learning, human-machine interaction, robot collaboration and so on. communication, because in the process of collaboration, task allocation, data transmission and information generalization are involved. Robots need to communicate frequently to obtain location information and working

2. Cooperative control of ROS mobile robot

The basis of collaboration between robots is two-way

status of other devices. Therefore, to realize the cooperative control of mobile robots based on ROS, we first need to build a communication platform, and all devices are connected to the platform, and use this as a benchmark for communication and wireless connection. At the same time, in order to facilitate the management and monitoring of all robot devices, a remote workstation device needs to be set up. The workstation can remotely connect any robot device for management, and can also monitor the working status and task progress of all robots.

2.1. *Design of distributed control system*

A distributed communication and control platform is built for robots to provide network support for robot collaboration.

A. *Static IP settings and master-slave IP configuration*

In order to ensure the normal communication function between robots and between robots and remote control terminals (workstations), each device needs to be connected to the same network, which can be provided by the wireless signal of the router or maintained by the hot spot service of the mobile phone. Because the general wireless network will assign different IP addresses to different users, although different robot devices are in the same network, because the IP address is different, and can not communicate, so the network provider is required to manually set up static IP and set up a multi-machine communication environment in the. bashrc file of each device. The setting of static IP is very simple. It only needs to change dynamic acquisition to manual acquisition in IPV4, and set fixed IP address, default routing, subnet mask and DNS.

ROS is a distributed computing environment. A running ROS system can contain multiple nodes distributed on multiple computers. According to the configuration of the robot operating system, multiple nodes between multiple computers, different nodes running on the same computer may need to communicate with each other at any time. Therefore, ROS requires a complete two-way connection between all robots on all ports for network configuration, and each computer must announce itself through a name that all other computers can parse.

B. *ssh Remote Connection*

In the distributed architecture of multi-processors, file transmission may be frequently involved between different robot systems. For example, ROS programs are written on

the workstation, and eventually need to run on the client. It is necessary to upload related directories and files from the workstation to the client. In fact, the design of the control system of multiple robots often adopts the mode of remote management of the workstation. After the network and function of each robot client are configured for the first time, in order to facilitate the update of the robot firmware, the processing when the operation is wrong and the iterative upgrade of the function package, it is also convenient to monitor the whole system. A remote workstation solves these problems. The problem is how to realize remote connection on the workstation and log in to the client for management. ROS gives the answer, which provides two remote control tools, ssh and vnc. Here only introduce ssh this way.

SSH (Secure Shell) is an encrypted network protocol for secure network services such as secure data communication, remote command line login, and remote command execution between two networked computers. It encrypts all data transmitted between two computers and provides secure and authenticated communication over an insecure network. It is commonly used by system administrators to manage remote servers and by developers to securely access and transfer files between machines. The whole process of data transmission is transparent and safe, using openssh tools will enhance your system security. The SSH implementation architecture is divided into two parts: client and server. The client is the requester and requests to upload files. The server is the receiver, receiving and downloading files from the client. The workstation belongs to the client and sends data to the robot. The robot belongs to the server. The implementation process is divided into four steps: install SSH client and server respectively ; the server starts the SSH service ; client remote login server ; realize data transmission.

C. *Optimization of ssh usage*

The client needs to enter the password every time it logs on to the server, and it also needs to enter the password when uploading and downloading files, which actually brings great inconvenience to the transmission of files. How to optimize the use of SSH, make the server trust the client, give the client a permission without password access, simplify the login steps with the help of the key, realize the secret-free login, and improve the efficiency of the operation. The implementation idea of optimization is to generate a pair of public key and private key. The private key is stored locally, and the public key is uploaded to the server. When each login, the local directly uploads the

private key to the server. The server can find the matching public key, and it is identified as a legitimate user. The ssh connection is directly created, and the client does not need to enter the password.

D. *Communication between robots*

The ROS kernel (roscore) is the basis of ROS operation, which has a parameter server. A running ROS has and only has one ROS kernel, and everything on ROS depends on this kernel [2]. ROS is a distributed framework based on the development and integration of many Nodes (nodes). Therefore, how to connect and manage many nodes is a problem that ROS must solve, which introduces a communication mechanism. In the communication mechanism, ROS Master (master node, node manager) manages and schedules the communication process between nodes in the network, and provides a service for configuring global parameters in the network. There are three basic communication modes of ROS, namely topic communication (publish-subscribe mode), service communication (request-response mode) and parameter server (parameter sharing mode).

2.2. *Synchronization control of mobile robot*

There are two control schemes to realize the synchronous control of robots. The first scheme is that the controller synchronously publishes multiple topics, and each topic corresponds to the ns attribute of different robots. The detailed process is shown in Fig. 1.
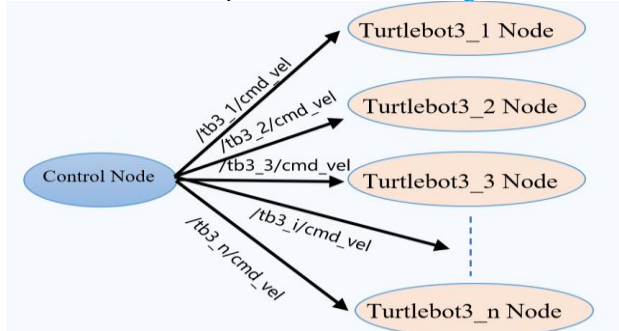
Fig. 1 The controller synchronously publishes multiple topics

The second scheme is to create an intermediate node, which receives the control instructions of the controller and processes them. After adding the ns attributes of different robots, it is released to the robot for control. Fig. 2 shows the implementation of scheme 2.
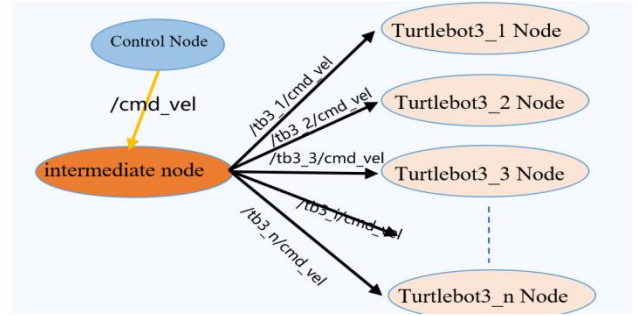
Fig. 2 Intermediate nodes assist in publishing

Both schemes allow the robot to obtain the same motion control. Scheme 1 needs to modify the code of the original given controller. After modification, it is impossible to control a single robot independently. Scheme 2 retains the original control node and still retains the control characteristics of a single robot. When multiple robots need to be controlled, only one more intermediate node needs to be run. Therefore, Scheme 2 is more practical.

2.3. *Mobile robot queuing to follow*

In order to realize the following of multiple robots, the following between two robots is considered first. If robot A follows robot B, the relationship between the position and attitude of the two robots must be known. Then through the coordinate transformation, the coordinate system transformation matrix between the two robots can be obtained, as well as the displacement vector randomly transformed to the pilot. The robot follows the simulation as shown in Fig. 3.
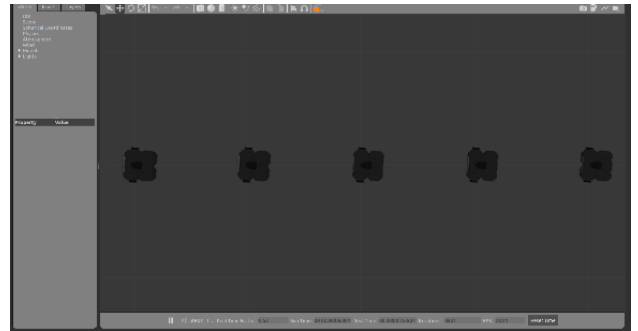
Fig. 3 Multi-robot following simulation

As long as the relationship between the coordinates is calculated in real time, ROS provides four methods to obtain the coordinate relationship of robot pose, which can be obtained by SLAM mapping, odometer, global camera, motion capture device and so on. After obtaining the coordinates, it can be sent to the TF tool in ROS by broadcast. Then subscribe to the relationship between the

two robots in the TF tool. The distance and attitude angle difference between the two rob real-time follow can be achieved through the TF tool. ots can be obtained by the pose relationship between the two robots. The speed and angular velocity instructions are issued by PID control to realize the following control of the robot.

The key to realize multi-machine following is to use TF (Transformations Frames) tool, which is a tool for monitoring and calculating robot pose information, providing pose monitoring, coordinate transformation and other functions. In the robot operating system, the TF toolkit includes Broadcaster, Listener, and TF conversion tools. The broadcaster broadcasts the position relationship between each robot and the world coordinate system to the TF tool through the topic; listener obtains the relative position of the two robot coordinate systems by viewing the TF toolkit.

2.4. *Multi-robot formation of mobile robot*

The principle of formation and following is that the TF tool is needed to obtain the relationship between the robot coordinate systems, but the following needs the transformation of the coordinate systems of the pilot and the random. The random moving target point is the coordinate origin of the pilot. The pilot moves, and the coordinate origin moves accordingly. TF will broadcast its pose in real time to control the random speed and angular velocity to achieve the purpose of real-time following. The coordinate system and coordinate transformation are shown in Fig. 4.
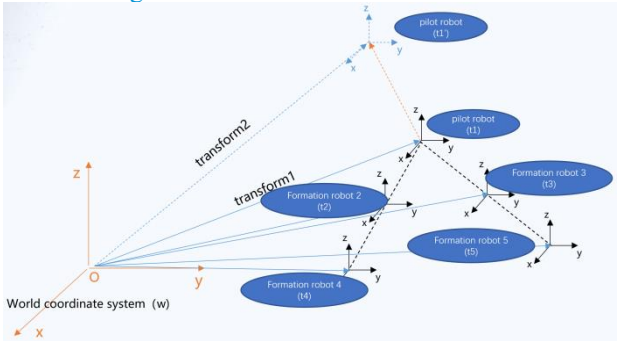
Fig. 4 Formation coordinate system and coordinate transformation

Compared with the following, the realization of the formation needs to establish a virtual coordinate system of the specified position around the leader, and obtain the change of the position and attitude of the virtual coordinate system following the leader coordinate system in real time

through the TF tool. The remaining formation aircraft track the position and attitude of the specified virtual coordinate system respectively, and the robot formation movement of various formations can be realized. By using the TF tool and the basic communication method to publish the virtual coordinate system of each specified position, the robot formation problem is transformed into the tracking problem of the target point. It is necessary to modify the callback function in the control of the pilot robot to realize the release of the specified virtual coordinates. Because the virtual coordinate is relative to the position of the pilot robot, it can be adjusted according to the required formation. In this paper, three formation modes are demonstrated, including horizontal line, triangle and square escort formation, and the position and transformation of virtual coordinates of triangle formation are given. They are shown in Fig. 5, 6 and 7.
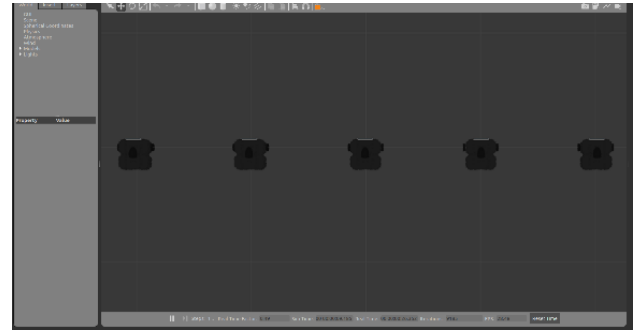
Fig. 5. Horizontal formation simulation
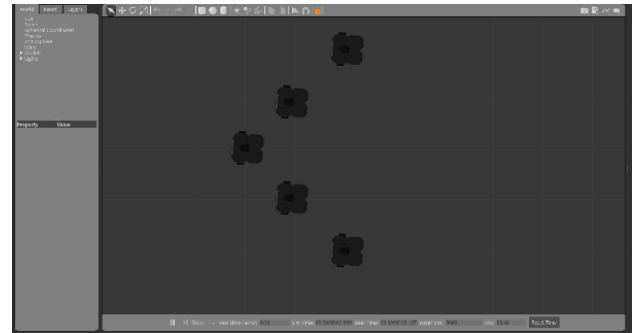
Fig. 6 Triangle formation simulation
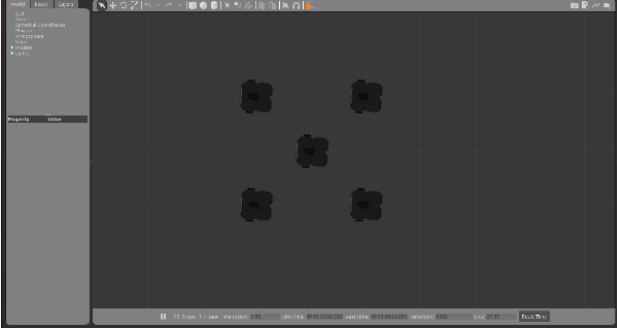
Fig. 7 Square escort formation simulation

2.5. *Cooperative navigation of mobile robot*
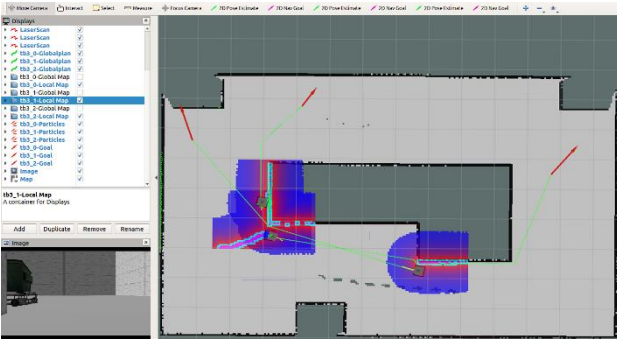
Fig. 8 Robot cooperative navigation

Fig. 8 shows the scene of three robots navigating at the same time in the same environment, and the three robots will determine their own movement routes through cooperation to avoid the risk of collision with each other, and move to the established target point in the shortest time and most efficiently.

A. *Robot path planning*

Before path planning, it is necessary to build a grid map of the robot 's working environment through SLAM mapping [3], because the robot 's navigation process needs to create target points on the map, and in the path planning process, the path planned by the robot will be displayed on the map in real time. Mapping algorithms are also diverse, including unscented Kalman filter UKF-SLAM, FAST-SLAM, Gmapping, etc [4], [5]. In this paper, the Gmapping mapping algorithm based on lidar data is selected. The basic idea of Gmapping algorithm is to obtain the information of the surrounding environment through lidar when the robot moves, and then construct a map and locate the position of the robot through filtering and matching algorithms.

In this paper, the TEB algorithm [6] is used to avoid obstacles. Firstly, several rough paths that can reach the target point are taken as the original paths. Then, various constraints are considered on each path, such as time, collision, speed, acceleration, etc., and these constraints are used to construct the optimization problem. Finally, the optimization problem of each initial path is solved, and a better path than initialization is obtained. Then, a better path is selected from the optimized path, and the result of local path planning can be obtained.

B. *Robot cooperative navigation*

ROS multi-robot cooperative navigation is to complete the cooperative work of multiple robots in the same environment to achieve navigation and exploration of the environment. In ROS, multi-robot cooperative navigation usually uses multiple nodes to achieve. Each robot has its own navigation nodes, which can exchange information with each other through ROS communication mechanism. In cooperative navigation, robots need to understand each other's position and state in order to coordinate their actions. When implementing ROS multi-robot cooperative navigation, the following aspects need to be considered:

Communication mechanism: ROS provides a variety of communication mechanisms, including Topic, Service and Parameter Server. In multi-robot cooperative navigation, topics and services are usually used to achieve communication between robots.

Position estimation: The robot needs to accurately estimate its position and posture. Laser radar, vision sensors and other sensors can be used to achieve position estimation.

Obstacle avoidance algorithm: In the navigation process, the robot needs to avoid obstacles and avoid collisions.

Cooperative planning algorithm: In multi-robot cooperative navigation, robots need to coordinate their actions to avoid mutual interference. Collaborative planning algorithms can be used to achieve collaborative action.

3. Conclusion

Since the birth of ROS in 2007, the research and development of this distributed framework system with nodes as processing units and functional units has flourished, and robot distributed control has great prospects in the development of service-oriented robots and warehouse logistics robots. As one of the popular development directions of robots in the future, multi-machine collaboration not only provides a solution to deal with complex working environments in the working process of robots, but also greatly improves the working efficiency of robots. The working requirements of service-oriented robots are not only the quality of service, but also the efficiency of their work. With the fast-paced operation of society, we are mostly in a ' fast-food ' attitude towards

Saijie Zhang, Huailin Zhao

life. The efficiency of a single robot is limited, and the management of multiple robots is complex. Therefore, a multi-machine cooperative control system that can be autonomously managed, information shared, divided and autonomous, and autonomously coordinated is developed, that is, swarm intelligence. It meets the requirements of high efficiency and simple management. Warehouse logistics robot is a system that

pursues time efficiency. How to ensure its work efficiency and achieve the required accuracy at the same time is more convenient for the management of multiple robots in the whole system. Multi-machine coordination of distributed control may give its answer. Let the robot coordinate autonomously, and at the same time share data in real time and allocate tasks, which is convenient for management and efficient.

References

1. Vincenzo DiLuoffo, William R Michalson & Berk Sunar. (2018).Robot Operating System 2. International Journal of Advanced Robotic Systems (3). doi:10.1177/1729881418770011.

2. Yahui Qi, Xiuzhen Wu, Qijie Chen, Shi Yan & Peng Zhang. (2022). Implementation of multi-UAV cooperative control based on ROS networking. (eds.) Proceedings of the Fifth National Conference on Cluster Intelligence and Cooperative Control in 2021 (pp.104-107).

3. Chengkan Huang. (2020). vision-based laser SLAM AGV initial positioning and path planning research (master 's thesis, Zhejiang University of Technology). https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD202101&filename=1020439617.nh

4. Xuebing Wang, Qingsheng Guo, Yong Wang, Qizhi Liu& Zhiwei Wei. (2019). Feature extraction and automatic recognition methods for map images. Surveying and mapping and spatial geographic information (09), 28-32.

5. José Carlos Rangel,Miguel Cazorla,Ismael García-Varea,Cristina Romero-González & Jesus Martínez-Gómez.(2019). Automatic semantic maps generation from lexical annotations. *Autonomous Robots* (3). doi:10.1007/s10514-018-9723-8.

6. C. Roesmann, W. Feiten, T. Woesch, F. Hoffmann and T. Bertram, "Trajectory modification considering dynamic constraints of autonomous robots," ROBOTIK 2012; 7th German Conference on Robotics, Munich, Germany, 2012, pp. 1-6.

Authors' Introduction

Mr. Saijie Zhang

He received the B.S. degree from Shanghai Institute of Technology, Shanghai, China. His main research interests is automation.

Dr. Huailin Zhao

He received his PhD from Oita University, Japan in 2008. He is a professor in the School of Electrical & Electronic Engineering, Shanghai Institute of Technology, China. His main research interests are robotics, multi-agent system and artificial intelligence. He is the members of both IEEE and Sigma Xi.

An OpenCV-based Method for Workpiece Residue Image Processing

Jiaxin Wang*, Hao He, Fangyv Liu, Fengzhi Dai
Tianjin University of Science and Technology, Tianjin, China

Email: *13293455078@163.com

Abstract

Workpiece residue refers to a thin film formed on the surface of the workpiece during the machining process, due to factors such as cutting fluid, chips, oil stains, etc., which affects the quality and performance of the workpiece. This paper proposes an OpenCV-based method for workpiece residue image processing, aiming to achieve automatic detection and analysis of workpiece residue. By building a workpiece image acquisition system, workpiece residue images of different types and degrees are collected, and the proposed method is verified and evaluated. The experimental results show that the proposed method can effectively detect and analyze workpiece residue, with high accuracy and robustness, providing an effective means for workpiece quality control.

Keywords: Workpiece residue, OpenCV-based, Image processing, Automatic detection and analysis

1. Introduction

With the continuous development of technology, artificial intelligence and machine learning have achieved remarkable results in many fields. Especially in the manufacturing industry, workpiece residue has a great impact on product quality and production efficiency. In order to solve this problem, many researchers have proposed a method of workpiece residue detection based on image processing. These methods process and analyze the image to identify the work parts and residues, so as to realize the automatic detection and classification of residues in the production process.

In the manufacturing industry, the strict requirements for product quality make the detection of workpiece residue an important task. Whether at home or abroad, many researchers are exploring how to effectively realize the automatic detection of workpiece residue. However, existing methods often have some problems, such as poor detection of workpiece and residue with complex shapes and sizes, or require a large number of known samples and expensive computational resources for training and learning.

This paper introduces a method of detecting workpieceresidue based on OpenCV image processing. OpenCV (Open Source Computer Vision Library) [1] is a widely used open source computer vision library. It contains rich image processing and analysis functions [2], and provides many efficient and practical tools and techniques, which are widely used in various image processing fields. Firstly, the image is preprocessed, including grayscale, binarization, noise removal and smoothing, to extract the features of the workpiece and residue. OpenCV's object detection algorithm and deep learning model are then used to identify areas where residue is likely to exist. Finally, geometric shape model and machine learning algorithm are used for post-processing to determine the location and size of the residue. This method can analyze the image in the production process and automatically identify the workpiece and residue. The experimental results show that the method can detect the workpiece residue effectively and improve the production efficiency and product quality.

2. Methodology

2.1. *Image preprocessing*

The image is read and preprocessed using OpenCV library. The pre-processing includes gray-scale, binarization, noise removal, smoothing, etc.[3], [4], in order to extract the features better. Grayscale is the conversion of color images to grayscale images in order to reduce computation and improve processing speed. Binarization is the conversion of a grayscale image into a binary image to highlight the differences in the object and background in the image. Noise removal and smoothing are designed to reduce noise and detail in images to improve the accuracy and stability of target detection. Some of the core code is shown in Fig.1.

Fig. 1 Preprocessing phase code

Jiaxin Wang, Hao He, Fangyv Liu, Fengzhi Dai

2.2. *Feature extraction*

In the pre-processed image, the features of the workpiece and residue are extracted. These features may include color, shape, texture, etc. OpenCV's feature extraction algorithms (such as SIFT, SURF, ORB, etc.) [5] are used to extract these features. These algorithms can automatically identify key points and feature descriptors in images, thus improving the accuracy and robustness of target detection. At the same time, edge detection is carried out to extract and draw the contour [4]. Some of the core code is shown in Fig. 2.

Fig.2 Feature extraction phase code

2.3. *Target detection*

Use OpenCV's object detection algorithms, such as Haar cascades [5] or deep learning models [6], to identify areas where residue may be present. Haar cascade is an object detection algorithm based on feature classification, which identifies objects in an image by combining a series of simple features together. Deep learning model is a kind of object detection algorithm based on neural network. By training and learning a large amount of data, a classifier can be used to classify and identify new samples. Some of the core code is shown in Fig. 3.

Fig.3 Target detection phase code

2.4. *Post-processing*

After possible residue areas are detected, some post-processing is performed to determine if these areas are indeed residue. For example, a geometric shape model is used to identify areas that do not match the shape of the workpiece, or a machine learning algorithm is used to classify and identify areas. Finally, the detected residue is marked out and the detection result is output. Some of the core code is shown in Fig. 4.

Fig.4 Post-processing phase code

2.5. *Result display and output*

Finally, the detected workpiece and residue are marked on the image, and the detection result is output. The qualified workpiece is marked with green "YES", and the image of the unqualified workpiece with residue is marked with red "NO".

3. Experimental Results and Analysis

3.1. *Related work*

In the field of image processing, many researchers have proposed various methods of workpiece residue Detection [7]. Among them, the method based on template matching is a simple and commonly used method. By matching the image to be detected with the standard template, the method finds the matching area with the template, and then identifies the workpiece and residue. However, this method is poor for the detection of workpiece and residue with complex shapes and sizes. To solve this problem, some researchers have proposed an approach based on machine learning [7], [8]. By training and learning a large number of known samples, a classifier is obtained, which can be used to classify and recognize new samples. However, this method requires a large number of known samples, and the training process of the classifier consumes a lot of time and computational resources.

In order to verify the validity of the proposed method, we conducted a series of experiments. The experimental results show that the method can detect the residue effectively and improve the production efficiency and product quality.

3.2. *Experiment settings*

We used a set of real artifact images as the experimental data set. The images include artifacts of various shapes and sizes as well as residues. The proposed method is verified by experiments, and the results are compared with traditional template matching methods and machine learning methods.

3.3. *Experimental result*

We use metrics such as accuracy, recall, and F1 scores to evaluate the performance of the approach presented in this paper. Accuracy refers to the proportion of correctly classified samples to the total number of samples; The recall rate refers to the proportion of the samples recalled

in the correctly classified samples to the total number of real samples. The F1 score is a harmonic average of accuracy and recall. Specific experimental results are shown in the following Table 1.

Table 1. Performance comparison of different methods

Method	accuracy	recall rate	F1 score
Template Matching	0.82	0.76	0.79
Machine Learning	0.85	0.82	

3.4. *Advantages*

1. Real-time: The method can detect the workpiece residue in the image in real time, so as to find the problems in the production process in time, improve production efficiency and product quality.

2. High efficiency: The method can automatically identify the workpiece and residue in the image, avoiding the tedious and time-consuming traditional manual detection, and improving the detection efficiency.

3. Accuracy: The method adopts advanced image processing and machine learning technology, which can accurately detect the workpiece and residue, avoiding the situation of missing and false detection.

4. Versatility: The method can be applied to different types and shapes of workpieces and residues, with strong versatility, can be widely used in various production fields.

5. Scalability: The method can improve the accuracy and efficiency of detection through continuous learning and training, and can also expand its function and application scope by adding new features and algorithms.

In short, the method of detecting workpiece residue based on OpenCV image processing can improve production efficiency and product quality, reduce production cost and risk, and has important practical significance and application value.

4. Conclusion

This paper presents an OpenCV based image processing method for detecting workpiece residue. Through feature extraction and object detection of the pre-processed images, the workpieces and residues can be accurately identified and classified. Compared with the traditional manual detection method, this method has higher detection efficiency and accuracy, and can effectively improve production efficiency and product quality. At the same time, the method has good universality and expansibility, and can be widely used in various production fields.

References

1. OpenCV: Introduction
2. Image Processing and Analysis_15_Image Registration : a survey of image registration techniques——1992...
3. Getting started: Image processing with OpenCV
4. OpenCV image preprocessing common functions and processes
5. OpenCV image feature extraction: corner feature, Harris and Shi-Tomas algorithm, SIFT/SURF algorithm, Fast and ORB algorithm
6. How to implement object detection algorithm in OpenCV?
7. Chen Xiao, ZHENG Xiaoyan, Zhao Lei, et al. Research on Surface defect image processing method of Workpiece based on OpenCV. Mechanical Engineering & Automation, 2018, (6): 142- 144.
8. Machine Vision based Defect Detection Approach using Image Processing

Authors Introduction

Ms. Jiaxin Wang

She is currently an undergraduate in Tianjin University of Science and Technology. She is majoring in robotics engineering.Her research area is about Image processing based on OpenCV.

Ms. Hao He

She is currently an undergraduate in Tianjin University of Science and Technology. She is majoring in robotics engineering.Her research area is about Image processing based on OpenCV.

Ms. Fangyv Liu

She is currently an undergraduate in Tianjin University of Science and Technology. Her research area is about intelligent robot and intelligent control.

Dr. Fengzhi Dai

He received an M.E. and Doctor of Engineering (PhD) from the Beijing Institute of Technology, China in 1998 and Oita University, Japan in 2004 respectively. His main research interests are artificial intelligence, pattern recognition and robotics. He worked in National Institute of Technology, Matsue College, Japan from 2003 to 2009. Since October 2009, he has been the staff in College of Electronic Information and Automation, Tianjin University of Science and Technology, China.

On Nonblockingness Verification and Enforcement of Controlled Nondeterministic Discrete-Event Systems

Xiang Ren*, Zipei Wang

College of Electronic Information and Automation, Tianjin University of Science and Technology, 300222，China

Email: *renxtust@163.com

Abstract

Discrete event systems, as an important kind of cyber-physical systems, have been widely used in engineering field. In this paper, we first express the dynamics of a controlled nondeterministic discrete-event system (acronym is DES) as an algebraic state-space representation using the semi-tensor product (STP) theory. And then, we discuss the problems of state-based nonblockingness verification and enforcement of nondeterministic DESs. Specifically, we obtain a criterion of verifying whether a given controlled nondeterministic DES is nonblocking. Further, we develop an efficient matrix-based approach to enforce state-based nonblockingness. We illustrate the applications of the proposed theoretical results using an example.

Keywords: Discrete event systems, nondeterministic, nonblockingness, semi-tensor product theory

1. Introduction

With the development of information and network technologies, discrete event systems (DESs) have received considerable attention in information physics systems for many years, synthesizing multidisciplinary methods such as automatic control and computer science, see, e.g., [1], [2]. Verification and synthesis are two main research issues in the field of DESs. We refer the readers to the recent survey papers [3] for more details.

In the above-mentioned literature, the DESs of interest, in general, are deterministic DESs. The case of nondeterministic DESs has not been considered widely so far since they have more complex dynamic evolution behaviors than deterministic DESs. It should be pointed out that, however, there are many realistic plants that need to be modeled as nondeterministic DESs. Therefore, how to model and analyze effectively the dynamics of controlled nondeterministic DESs are still an interesting topic for both computer and control researchers.

In classical DESs, the nonblockingness is a basic and important problem in the study of DESs. Also, some efficient approaches have been proposed. For instance, in [4], the authors studied the problem of blocking detection of DESs by means of language-based approach. In this paper, we develop a new methodology to investigate how to model controlled nondeterministic DESs by using the STP of matrices. Based the proposed new model, we further study the property verification and enforcement of nonblockingness.

The rest of this article is organized as follows. The second section presents some basic notations, concepts and a matrix-based expression needed in this paper. In the third part, we formulate the state-based nonblockingness property verification problem, and we show how the proposed approach can be to applied to enforce nonblockingness. In the fourth section, an example is presented to illustrate the application of the proposed approach. The fifth part summarizes the main content of this paper.

2. Preliminaries

2.1. *Notations*

In this subsection, we introduce some notations, which will be used in the sequel. $\square^n$ is the set of all vectors of dimension n ; $|X|$ is the cardinality of set X ; $M_{m \times n}$ is the set of $m \times n$ matrices; $M_{(i,j)}$ is the (i, j) element of matrix M ; $Col_j(M)$ is the j-th column of matrix M ; $Col(M)$ is the set of all columns of matrix M ; $0_n := \underbrace{[0,0,\cdots,0]}_{n}$;

$1_n := \underbrace{[1,1,\cdots,1]}_{n}$; $\delta_n^0 := \underbrace{[0,0,\cdots,0]}_{n}^T$; $\delta_n^k := Col_k(I_n)$,

$1 \le k \le n$; $\Delta_n := \left\{ \delta_n^1, \delta_n^2, \cdots, \delta_n^n \right\}$; $\tilde{\Delta}_n := \left\{ \delta_n^0, \delta_n^1, \cdots, \delta_n^n \right\}$; $L \in M_{m \times n}$ is a logical matrix (resp., generalised logical matrix) if $Col(L) \subseteq \Delta_m$ (resp., $Col(L) \subseteq \tilde{\Delta}_m$). We denote the set of $m \times n$ logical matrices (resp., generalised logical matrix) by $L_{m \times n}$ (resp., $\tilde{L}_{m \times n}$); If matrix $L \in \tilde{L}_{m \times n}$,

then it can be expressed as $L \in \left[\delta_m^{i_1}, \delta_m^{i_2}, \cdots, \delta_m^{i_n} \right]$ and it is briefly denoted as $L \in \delta_m \left[i_1, i_2, \cdots, i_n \right]$, where $i_k \in \{0, 1, \cdots, m\}$, $1 \le k \le n$.

2.2. *Semi-tensor product (STP) of matrices*

In this subsection, we recall the notions of STP and swap matrix. We refer the readers to [5] and/or [6] for more details on them.

Definition 2.1 ([5]): Let $A \in \mathrm{M}_{m \times n}$, $B \in \mathrm{M}_{p \times q}$. The STP of A and B is defined as
$$A \cdot B = (A \otimes I_{t/n})(B \otimes I_{t/p}), \qquad (1)$$
where t denotes the least common multiple of n and p, i.e., $t = lcm(n, p)$; $\otimes$ is the Kronecker product.

Definition 2.2 ([5]): A swap matrix $W_{[m,n]}$ is a $mn \times mn$ logical matrix, which is defined as
$$
\begin{aligned}
W_{[m,n]} = \delta_{mn}[&1, m+1, 2m+1, \cdots, (n-1)m+1, \\
&2, m+2, 2m+2, \cdots, (n-1)m+2, \qquad (2)\\
&\cdots, m, 2m, 3m, \cdots, nm].
\end{aligned}
$$

Lemma 2.1([5]): Let $X \in \square^m$ and $Y \in \square^n$ be two column vectors. Then
$$W_{[m,n]} XY = YX, \quad W_{[n,m]} YX = XY. \qquad (3)$$

2.3. *System model*

In this subsection we recall the formalism used in the paper. More details on DESs can be found in [2].

A DES is modeled as a nondeterministic finite automaton (NFA) $G = (X, \Sigma, \delta, X_0, X_m)$, where X is the finite set of states, Σ is the finite set of events called alphabet or input symbols, $X_0 \subseteq X$ is the set of initial states, $X_m \subseteq X$ is the set of marked states (or accepted states), $\delta : X \times \Sigma \to 2^X$ is the partial transition function(2^X denotes the power set of X), which describes the system dynamics: given states $x, y \in X$ and an event $\sigma \in \Sigma$, $y \in \delta(x, \sigma)$ means the execution of σ from state x takes the system to state y. Note that $\delta(x, \sigma)$ is undefined when the event σ cannot be executed from the state x. $\delta(x, \sigma)!$ denotes $\delta(x, \sigma)$ is well-defined. Obviously, the transition function can be extended to $\delta : X \times \Sigma^* \to 2^X$ in terms of $\delta(x, e) := \delta(\delta(\cdots \delta(\delta(x, e_{j_1}), e_{j_2}) \cdots), e_{j_t})$, where $e = e_{j_1} e_{j_2} \cdots e_{j_t} \in \Sigma^*$, Σ^* denotes the set of finite strings on the alphabet Σ, including the empty string $\bullet$. The objective of this paper is to investigate the controlled nondeterministic DESs. In this regard, the event set Σ can

be partitioned into two disjoint subsets, i.e., $\Sigma = \Sigma_c \bigcup \Sigma_{uc}$, where Σ_c denotes the set of controllable events, Σ_{uc} denotes the set of uncontrollable events. We here assume that all events in Σ are observable.

In general, we wish to adjoin a supervisor or a controller S to interact with G in a feedback manner. More precisely, the transition function of G can be controlled by S in the sense that the controllable events of G can be dynamically enabled or disabled by S after each transition. Formally, a *state-feedback supervisor*, denoted by S, is a function $S : X \to 2^{\Sigma_c}$ that determines the set of events $S(x) \subseteq \Sigma_c$ to be disabled at each state $x \in X$, while events not belonging to the set $S(x)$ remain enabled at state x. The *controlled system* (or called *supervised system*) consisting of G and S, denoted by S/G, is another nondeterministic finite automation given as
$$S/G = (X, \Sigma, \delta_S, X_0, X_m), \qquad (4)$$
where X, Σ, X_0 and X_m are as defined above, δ_S is the partial transition function of S/G, i.e.,
$$
\delta_S(x, \sigma) = \begin{cases} \delta(x, \sigma), \text{if } \delta(x, \sigma)! \text{ and } \sigma \notin S(x) \\ \text{undefined, otherwise.} \end{cases} \qquad (5)
$$

We use the notation $\mathrm{H}(x)$ to denote the set of feasible events of G at state x. Thus supervisor S is called permissible if for all $x \in X$, $S(x) \subseteq \mathrm{H}(x) \bigcap \Sigma_c$. Note that it is not difficult to see that nondeterministic DESs can be viewed as special case of controlled nondeterministic DESs with $S(x) = \varnothing$ for all $x \in X$. The controlled nondeterministic DES S/G is depicted in Fig.1.
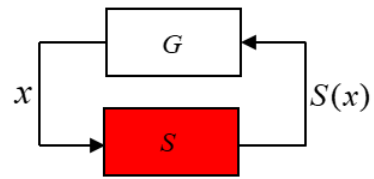
Fig. 1. The controlled nondeterministic DES S/G

3. Nonblockingness of Controlled Nondeterministic DESs

3.1. *Matrix-based expression of controlled nondeterministic DESs*

In order to obtain the matrix expression of the dynamics of controlled nondeterministic DESs, let us first give an equivalent description of the controlled system (4). For the nondeterministic DES G, we define a control pattern as a Boolean function: $\gamma : \Sigma_c \to \{0, 1\}$ and

we use the notation $\Gamma = \{0,1\}^{\Sigma_c}$ to denote the set of all Boolean functions on Σ_c. $\gamma \in \Gamma$ is interpreted as follows: for any $\sigma \in \Sigma_c$, $\gamma(\sigma) = 1$ means that the control pattern γ allows σ to happen, while $\gamma(\sigma) = 0$ means that the control pattern γ refuse σ to happen. Note that it is convenient to extend each $\gamma \in \Gamma$ to a function $\gamma : \Sigma \to \{0,1\}$ by defining $\gamma(\sigma) = 1$ for each uncontrollable event $\sigma \in \Sigma_{uc}$. Further, we define a partial function of the form $f : X \to \Gamma$, called the *state-feedback control function* or *state-feedback* for short, that maps each state x in X into control pattern γ, i.e. $\gamma(\sigma) = f(x)(\sigma)$. Thus the controlled nondeterministic DES consisting of the state-feedback f, control pattern γ and nondeterministic DES G is described as

$$G_\gamma^f = (X, \Sigma, \delta_\gamma^f, X_0, X_m), \qquad (6)$$

where X, Σ, X_0 and X_m are as defined above; δ_γ^f denotes the partial transition function of G_γ^f, which is defined as

$$\delta_\gamma^f(x,\sigma) = \begin{cases} \delta(x,\sigma), & \text{if } \delta(x,\sigma)! \text{ and } f(x)(\sigma) = 1 \\ \text{undefined}, & \text{otherwise.} \end{cases} \qquad (7)$$

For event $e_j (1 \le j \le m)$, we define a $n \times n$ matrix F_j as

$$F_{j(s,t)} = \begin{cases} 1, & \text{if } \delta_n^s \in \delta_\gamma^f(\delta_n^t, \delta_m^j) \wedge \delta_m^j \square e_j \in \Sigma_{uc} \\ r_{tj}, & \text{if } \delta_n^s \in \delta(\delta_n^t, \delta_m^j) \wedge \delta_m^j \square e_j \in \Sigma_c \\ 0, & \text{otherwise,} \end{cases} \qquad (8)$$

where F_j is called transition structure matrix *w.r.t.* event e_j. Thus, the *transition structure matrix* (TSM) of controlled nondeterministic DES (6) is defined as

$$F = [F_1, F_2, \cdots, F_m], \qquad (9)$$

where F is a $n \times mn$ symbol matrix.

Theorem 3.1: Given a controlled nondeterministic DES (6), the dynamics of (6) can be equivalently described as

$$x(t+1) = Fu(t)x(t), \qquad (10)$$

where F is the TSM of system (6), which is defined in (9); $x(t) = (x_1(t), x_2(t), \cdots, x_n(t))^T$ is the vector form of state at step t, $x_i(t)$ denotes the number of different paths from the set of initial states to the state $x_i = \delta_n^i$ with a feasible event string of length $t - 1$; $u(t) \in \Delta_m$ is vector form of event at step t.

Proof: We here omit this proof.

Remark 3.1: When the state-feedback f is known, the dynamics of (6) can be equivalently described by the following equation

$$x(t+1) = F_c u(t)x(t), \qquad (11)$$

where $x(t)$ and $u(t)$ have the same interpretation as in Theorem 3.1; F_c is the TSM of (6), which is represented as

$$F_c = \left[F_1^c, F_2^c, \cdots, F_m^c \right] \in \tilde{L}_{n \times mn}, \qquad (12)$$

where the $n \times n$ matrix F_j^c is defined as

$$F_{j(s,t)}^c = \begin{cases} 1, & \text{if } \delta_n^s \in \delta_\gamma^f(\delta_n^t, \delta_m^j) \\ 0, & \text{otherwise.} \end{cases} \qquad (13)$$

3.2. *Nonblockingness verification of controlled nondeterministic DESs*

In classical DESs, the problem of nonblockingness verification is a basic and important topic. A controlled deterministic DES, also denoted by S/G is said to be nonblocking if and only if $L(S/G) = \overline{L_m(S/G)}$, where $L(S/G)$ and $\overline{L_m(S/G)}$ stand for the language generated and language marked by S/G, respectively. See, e.g., [2] for more details. As a generalization of deterministic model, we now investigate the nonblockingness of controlled nondeterministic DESs. We for simplicity use a *state-based description* to present the notion of nonblockingness. We call it *state-based nonblockingness*, while the nonblockingness defined in [2] is called *the language-based nonblockingness*. Actually, these two concepts are equivalent.

Definition 3.1: Given a controlled nondeterministic DES (6), a state $x \in X$ is called a *nonblocking state* if it is reachable from initial state set X_0 and it can reach a marked state from X_m. (6) is called *nonblocking* if its each state is nonblocking.

Remark 3.2: 1) By Definition 3.1, we know that a state of (6) is a *blocking state* if and only if it is not reachable from the set of initial states, or it can not reach any marked state. 2) For the controlled nondeterministic DES (6), there exist two forms of blocking state, i.e., *deadlock state* and *livelock state*. More specifically, $x \in X$ is a deadlock state of (6) if it is an unmarked reachable state and $H(x) = \varnothing$; $x \in X$ is a livelock state of (6) if it can not reach any marked state and $H(x) \ne \varnothing$.

In what follows, we give a *matrix-based criterion* to verify the nonblockingness of controlled nondeterministic DES (6). Consider the controlled nondeterministic DES (6) with the matrix-based expression (11), from the definition of the TSM F_c presented in (13) and the viewpoint of graph theory, we readily know that the matrix $\sum_{j=1}^{m} F_j^c$ stands for the *adjacency matrix* of the directed graph representation of the controlled nondeterministic DES (6). Thus, state

$x_q = \delta_n^q$ is reachable from $x_p = \delta_n^p$ in t steps if and only if $(\sum_{j=1}^{m} F_j^c)^t_{(q,p)} > 0$. In this regard, we define a matrix $M =: \sum_{\alpha=1}^{n} (\sum_{j=1}^{m} F_j^c)^{\alpha}$, called *reachability matrix* of controlled nondeterministic DES (6), where $n = |X|$, $m = |\Sigma|$. Using it, we have the following result.

Proposition 3.1: Given a controlled nondeterministic DES (6) with the matrix-based expression (11), then

1) The set of states reachable from the set of initial states X_0 is

$$\mathrm{R}(X_0) = \{\delta_n^k \mid M_{(k,p)} > 0, \delta_n^p \in X_0, 1 \le k \le n\}, \quad (14)$$

where $M = \sum_{\alpha=1}^{n} (\sum_{j=1}^{m} F_j^c)^{\alpha}$ is the reachability matrix of (6).

2) The set of states that can reach X_m is

$$\tilde{\mathrm{R}}(X_m) = \{\delta_n^k \mid M_{(q,k)} > 0, \delta_n^q \in X_m, 1 \le k \le n\}, \quad (15)$$

where M is as defined above.

Proof: We here omit this proof.

Remark 3.3: By convention, we assume that $X_0 \subseteq \mathrm{R}(X_0)$ and $X_m \subseteq \tilde{\mathrm{R}}(X_m)$ hereinafter. Using Definition 3.1 and Proposition 3.1, the following main result on the verification of nonblockingness for the controlled nondeterministic DES (6) is very straightforward.

Theorem 3.2: Given a controlled nondeterministic DES (6) with the matrix-based expression (11), then (6) is nonblocking if and only if the following condition holds:

$$\mathrm{R}(X_0) = \tilde{\mathrm{R}}(X_m) = X, \quad (16)$$

where $\mathrm{R}(X_0)$ and $\tilde{\mathrm{R}}(X_m)$ have the same interpretation as in (14) and (15), respectively.

Notice that, from Theorem 3.2, it is not difficult to see that using our approach to verify nonblockingness of controlled nondeterministic DESs is very efficient and straightforward in comparison to the existing ones in the sense that it involves only some basic matrix manipulations that are completed via polynomial time.

3.3. *Nonblockingness enforcement of controlled nondeterministic DESs*

We know that, for uncontrolled DESs, the trimness property is closely related to the concept of nonblockingness. For instance, when a given original system does not satisfy the nonblockingness property, we can to enforce it via trim operation to restrict the system's behavior, see, e.g.,[2]. It is also meaningful for the case

of controlled nondeterministic DESs, since a controlled nondeterministic DES, denoted by S/G, has a generated language and a marked language associated it. In this subsection, we present a matrix-based approach to enforcing nonblockingness property called nonblockingness enforcement hereinafter.

Consider the controlled nondeterministic DES (6) with the matrix-based expression (11), and we assume that, without loss of generality, system (6) is blocking. Using Lemma 2.1, equation (11) can be rewritten as

$$x(t+1) = \bar{F}_c x(t) u(t), \quad (17)$$

where $\bar{F}_c = F_c W_{[n,m]} \in \mathrm{M}_{n \times mn}$, called the *dual* TSM of (6). Partition matrix $\bar{F}_c$ into $\bar{F}_c = [\bar{F}_1^c, \bar{F}_2^c, \cdots, \bar{F}_n^c]$, where $\bar{F}_i^c \in \mathrm{M}_{n \times m}, 1 \le i \le n$. Also, construct matrix:

$$\bar{F}^{trim} = [\bar{F}_1^{trim}, \bar{F}_2^{trim}, \cdots, \bar{F}_n^{trim}], \quad (18)$$

where

$$\bar{F}_i^{trim} = \begin{cases} \bar{F}_i^c, & \text{if } \delta_n^i \in \mathrm{R}(X_0) \bigcap \tilde{\mathrm{R}}(X_m) \\ 0_{n \times m}, & \text{otherwise.} \end{cases} \quad (19)$$

Based on the above-constructed matrix (18), we have the following result.

Theorem 3.3: Given a controlled nondeterministic DES (6) with the matrix-based expression (11), then the controlled nondeterministic DES described by the following equation is nonblocking.

$$x(t+1) = F^{trim} u(t) x(t), \quad (20)$$

where $F^{trim} = \bar{F}^{trim} W_{[m,n]}$, $\bar{F}^{trim}$ is defined in (18).

Proof: we here omit the proof of Theorem 3.3.

Remark 3.4: From Theorem 3.3, we can obtain that, when the DES (6) is blocking, the computational complexity of enforcing the nonblockingness property by using the proposed approach is also polynomial time, i.e., $o(n^2 m)$, where n and m are the numbers of state nodes and event nodes, respectively. An example concerning the supervisory control of a submarine guard system is presented (see Example 4.1 below) to illustrate the effectiveness and application of the proposed results.

4. Illustrative Example

In this section, we give an example to illustrate the proposed results.

Example 4.1: Let us consider the supervisory control of a submarine guard system. A sea region is divided into several sections in terms of the shape of the seabed and it is depicted in Fig. 2. Two submarines, namely, and, cruise and guard the sea region. has its own separate repair port and home harbor, i.e., section No. 5 and No. 2,

respectively. has repair and home harbor in section No. 4. Due to the seabed form and the sea underwater streams, as well as different size of the submarines, possible and controllable transitions of the submarines between these sections are also shown in Fig. 2.
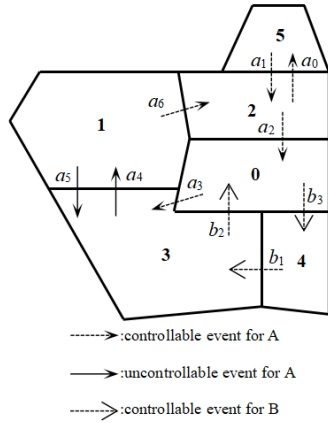
Fig. 2. Space guarded by the submarines.

At the beginning of inspection, A is in section No. 5 and B in section No. 4. It is required that each submarine can return its home harbor. Additionally, meeting of submarines A and B in any section of the sea region must be prohibited for safety reasons. The task of the supervisory control is how to synthesize the cruise trajectory of A and B to prevent the violation of the requirements described above.

Now we use the proposed approach to solve it.

According to Fig. 2, the dynamics of submarines A and B can be modeled by two finite-state automata shown in Fig. 3(a) and Fig 3(b), respectively. That is, $G_1 = (X_1, \Sigma_1, \delta_1, x_0^1, X_m^1)$ with the initial state $x_0^1 = 5$, and the marked state $X_m^1 = \{2\}$, while $G_2 = (X_2, \Sigma_2, \delta_2, x_0^2, X_m^2)$ with the initial state $x_0^2 = 4$ and the marked state $X_m^2 = \{4\}$.
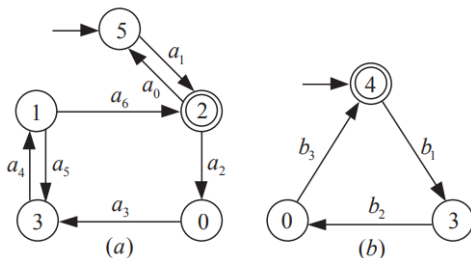
Fig. 3. (*a*) represents a DFA model of the submarine A cruise; (*b*) represents a DFA model of the submarine B cruise.

For the automation G_1, identifying $5 \square \delta_5^1$, $2 \square \delta_5^2$, $0 \square \delta_5^3$, $3 \square \delta_5^4$, $1 \square \delta_5^5$; $a_{j-1} \square \delta_{10}^j (1 \le j \le 6)$. While for G_2, identifying $4 \square \delta_3^1$, $3 \square \delta_3^2$, $0 \square \delta_3^3$; $b_k \square \delta_{10}^{7+k} (1 \le k \le 3)$.

According to lemma in reference [7], the dynamics of $G_1 \| G_2$ (called the parallel composition of G_1 and G_2) can be described by the following algebraic form

$$x(t+1) = Fu(t)x(t), \tag{21}$$

where $F = [F_1, F_2, \cdots, F_9, F_{10}] \in \tilde{L}_{15 \times 150}$, i.e.,

$F = \delta_{15}[0,0,0,1,2,3,0,0,0,0,0,0,0,0,0,4,5,6,0,0,0,0,$
$0,0,0,0,0,0,0,0,0,0,0,7,8,9,0,0,0,0,0,0,0,0,$
$0,0,0,0,0,0,0,10,11,12,0,0,0,0,0,0,0,0,0,0,$
$0,0,0,13,14,15,0,0,0,0,0,0,0,0,0,0,0,0,0,0,$
$0,10,11,12,0,0,0,0,0,0,0,0,0,0,0,4,5,6,2,0,$
$0,5,0,0,8,0,0,11,0,0,14,0,0,3,0,0,6,0,0,9,0,$
$0,12,0,0,15,0,0,1,0,0,4,0,0,7,0,0,10,0,0,13].$

Using Lemma 2.1, system (21) can be rewritten as

$$x(t+1) = \tilde{F}x(t)u(t), \tag{22}$$

where $\tilde{F} = FW_{[15,10]} \in \tilde{L}_{15 \times 150}$.

Partitioning $\tilde{F} = [\tilde{F}_1, \tilde{F}_2, \cdots, \tilde{F}_{15}]$, where $\tilde{F}_k \in \tilde{L}_{15 \times 10}$, $1 \le k \le 15$. Since states $00 \square \delta_5^3 \delta_3^3 = \delta_{15}^9$ and $33 \square \delta_5^4 \delta_3^2 = \delta_{15}^{11}$ are forbidden states², all events attached to δ_{15}^9 and δ_{15}^{11} should be deleted and each state that can reach δ_{15}^9 or δ_{15}^{11} in 1 step by an uncontrollable event is also forbidden state. To this end, by applying Proposition 3.1, the sets of states reachable from δ_{15}^9 and δ_{15}^{11} in 1 step are $\{\delta_{15}^7, \delta_{15}^{12}\}$ and $\{\delta_{15}^{12}, \delta_{15}^{14}\}$, respectively. The sets of states that can reach δ_{15}^9 and δ_{15}^{11} in 1 step are $\{\delta_{15}^6, \delta_{15}^8\}$ and $\{\delta_{15}^8, \delta_{15}^{10}, \delta_{15}^{14}\}$, respectively. According to the conclusion of reachability analysis in reference[8], we have $\delta_{15}^9 \xrightarrow{\delta_{10}^{10}} \delta_{15}^7$, $\delta_{15}^9 \xrightarrow{\delta_{10}^4} \delta_{15}^{12}$, $\delta_{15}^6 \xrightarrow{\delta_{10}^3} \delta_{15}^9$, $\delta_{15}^8 \xrightarrow{\delta_{10}^{10}} \delta_{15}^9$; $\delta_{15}^{11} \xrightarrow{\delta_{10}^{10}} \delta_{15}^{12}$, $\delta_{15}^{11} \xrightarrow{\delta_{10}^5} \delta_{15}^{14}$, $\delta_{15}^8 \xrightarrow{\delta_{10}^4} \delta_{15}^{11}$, $\delta_{15}^{10} \xrightarrow{\delta_{10}^8} \delta_{15}^{11}$, $\delta_{15}^{14} \xrightarrow{\delta_{10}^6} \delta_{15}^{11}$. Thus, state δ_{15}^{14} is also forbidden state. Analogous to δ_{15}^9 and δ_{15}^{11}, we have $\delta_{15}^{14} \xrightarrow{\delta_{10}^7} \delta_{15}^5$, $\delta_{15}^{14} \xrightarrow{\delta_{10}^9} \delta_{15}^{15}$, $\delta_{15}^{13} \xrightarrow{\delta_{10}^8} \delta_{15}^{14}$.

Based on the above analysis, we construct matrix

$$\tilde{F}_{c_1} = [\tilde{F}_1^{c_1}, \tilde{F}_2^{c_1}, \cdots, \tilde{F}_{15}^{c_1}], \tag{23}$$

where

$$\tilde{F}_i^{c_1} = \begin{cases} 0_{15 \times 10}, & i = 9,11,14 \\ \tilde{F}_i, & \text{otherwise.} \end{cases} \tag{24}$$

Consequently, system (22) becomes

$$x(t+1) = \tilde{F}_{c_1} x(t)u(t), \tag{25}$$

The next step to preserve nonblockingness of the controlled DES (25). Using Proposition 3.1, we can obtain easily that $R(\delta_{15}^1) = \{\delta_{15}^1, \delta_{15}^2, \delta_{15}^3, \delta_{15}^4, \delta_{15}^5, \delta_{15}^6, \delta_{15}^7, \delta_{15}^8, \delta_{15}^{10}, \delta_{15}^{13}\}$, $\tilde{R}(\delta_{15}^4) = \{\delta_{15}^1, \delta_{15}^2, \delta_{15}^3, \delta_{15}^4, \delta_{15}^5, \delta_{15}^6, \delta_{15}^7, \delta_{15}^{10}, \delta_{15}^{12}, \delta_{15}^{13}, \delta_{15}^{15}\}$. We know that states δ_{15}^8, δ_{15}^{12} and δ_{15}^{15} are forbidden states. Thus, let

$$\tilde{F}_{c_2} = [\tilde{F}_1^{c_2}, \tilde{F}_2^{c_2}, \cdots, \tilde{F}_{15}^{c_2}], \tag{26}$$

where

$$\tilde{F}_i^{c_2} = \begin{cases} 0_{15\times10}, & i = 8,12,15 \\ \tilde{F}_i^{c_1}, & \text{otherwise.} \end{cases} \quad (27)$$

Hence, the controlled system (25) becomes

$$x(t+1) = \tilde{F}_{c_2} x(t)u(t), \quad (28)$$

Again, using Lemma 2.1, we rewrite system (28) is

$$x(t+1) = \tilde{F}_{c_3} u(t)x(t), \quad (29)$$

where $\tilde{F}_{c_3} = \tilde{F}_{c_2} W_{[10,15]}$, specifically,

$\tilde{F}_{c_2} = \delta_{15}[0,0,0,1,2,3,0,0,0,0,0,0,0,0,0,4,5,6,0,0,0,0,$
$0,0,0,0,0,0,0,0,0,0,0,7,0,0,0,0,0,0,0,0,0,0,0,0,$
$0,0,0,0,0,0,0,0,10,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,$
$0,0,0,0,13,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,$
$0,10,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,4,0,0,2,0,$
$0,5,0,0,0,0,0,0,0,0,0,0,0,0,0,3,0,0,6,0,0,0,0,$
$0,0,0,0,0,0,0,0,1,0,0,4,0,0,0,0,0,0,0,0,0].$

It is readily verified that the controlled DES (29) satisfies the requirements described above. Consequently, the cruise trajectory of submarines A and B, by resorting to the algebraic representation (29), is depicted in Fig. 4.
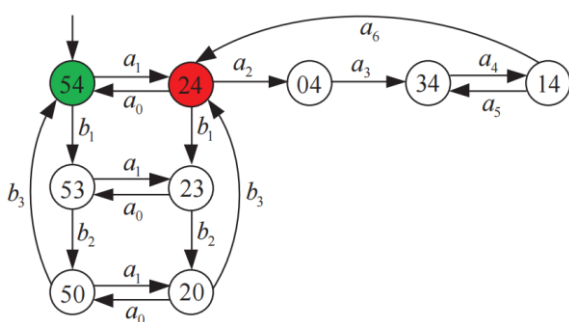
Fig. 4. The cruise trajectory of submarines *A* and *B* in the submarine system.

5. Conclusion

In this paper, we used matrix-based form to model the dynamics of controlled nondeterministic DESs. Using it, we investigated the verification of nonblockingness property of controlled nondeterministic DESs. Also, the criteria of verifying nonblockingness was presented in terms of the methodology described in this paper. Furthermore, the proposed approach can be applied to enforce nonblockingness of controlled nondeterministic DESs.

We hope that our work can establish the foundation for a new theory of supervisory control of controlled nondeterministic DESs. Future work will investigate the problem of synthesizing property-enforcing supervisor for controlled nondeterministic DESs by means of the proposed theoretical framework.

References

1. Y. F. Chen, Z. W. Li, K. Barkaoui, et al., "Compact supervisory control of discrete event systems by Petri nets with data inhibitor arcs," IEEE Trans. Syst., Man, Cybern., Syst., vol. 47, no. 2, pp. 364–379, Feb. 2017.
2. C. G. Cassandras and S. Lafortune, Introduction to Discrete Event Systems, 2nd Ed. New York: Springer, 2008.
3. A. Giua and M. Silva "Modeling, analysis and control of discrete event systems: A Petri net perspective," in Proc. the 20th IFAC World Congress, Toulouse, France, Jul. 2017, pp. 1808–1819.
4. S. Abdelwahed and W. M. Wonham, "Blocking detection in discrete event systems," American Control Conference, Denver, USA, Jun. 2003, pp. 1673–1678.
5. D. Z. Cheng, H. S. Qi, and Y. Zhao, An Introduction to Semi-tensor Product of Matrices and Its Applications. Singapore: World Scientific, 2012.
6. X. G. Han, Z. Q. Chen, Z. X. Liu, et al., "Calculation of siphons and minimal siphons in Petri nets based on semi-tensor product of matrices," IEEE Trans. Syst., Man, Cybern., Syst., vol. 47, no. 3, pp. 531–536, Mar. 2017.
7. X. R. Xu, Y. G. Hong, and H. Lin, "Matrix approach to simulation and bisimulation analysis of finite automata," in proc. 10th Word Congress on Intelligent Control and Automation, Beijing, China, Jul. 2012, pp. 2716–2721.
8. X. R. Xu and Y. G. Hong, "Matrix expression and reachability analysis of finite automata," J. Control Theory Appl., vol. 10, no. 2, pp. 210–215, 2012.

Authors Introduction

Mr. Xiang Ren

He received his B.S. degree from College of Electronic Information and Automation, Tianjin University of Science and Technology, China in 2021. He is currently a Master course student in Tianjin University of Science and Technology.

Ms. Zipei Wang

She received the B.S. degree from College of Electronic Information and Automation, Tianjin University of Science and Technology, China in 2023. She is currently a Master course student in Tianjin University of Science and Technology.

Modeling and Reachability Verification of Controlled Nondeterministic Finite-State Automata

Zipei Wang*, Xiang Ren

College of Electronic Information and Automation, Tianjin University of Science and Technology, 300222，China

E-mail: *wangzipei0301@163.com

Abstract

In this paper, we investigate the modeling and state reachability of controlled nondeterministic finite-state automata (NFA). The key feature of a controlled NFA is to admit a controller (also called a supervisor) to intervene the behavior of an original system. We first express the dynamics of a controlled NFA as an algebraic state-space representation in the framework of the semi-tensor product (STP) of matrices. Then, the necessary and sufficient condition for verifying state reachability of controlled NFA is presented. An explicit formula for calculating all paths of any two states is derived. Finally, we use an example to illustrate the application of the proposed theoretical results.

Keywords: Discrete event systems, nondeterministic finite-state automata, reachability, semi-tensor product of matrices

1. Introduction

Discrete event systems (DESs), also called plants, have received considerable attention within the automatic control and computer science communities for many years, see, e.g., [1], [2].

Nondeterministic plants can be viewed as a generalization of deterministic plants. It is useful when designing a system at a higher-level of abstraction so that lower-level details of system are omitted to obtain higher-level models that may be nondeterministic. It is well-known that controlled DESs modeled by controlled finite automata have more complex structures and dynamics than uncontrolled DESs. Therefore, how to model and analyze effectively the dynamics of controlled nondeterministic DESs are still an interesting topic.

In classical DESs, the reachability is a basic and important problem in the study of DESs. For instance, [3] studied the matrix expression and reachability verification of finite automata (including deterministic and nondeterministic). The approach in these two papers are based on algebraic state-space representation. In this paper, we develop a new methodology to investigate how to model controlled nondeterministic finite-state automata (NFA) using the STP of matrices.

The rest of this article is organized as follows. The second section presents some basic notations and concepts needed in this paper. In the third part, we present a matrix-based expression for the dynamics of controlled NFA. In the fourth section, we give a methodology of verifying state reachability of controlled NFA, and an explicit formula of finding all paths of any two states is also provided, if exists. In the fifth section, an example is presented to illustrate the application of the proposed approach. The sixth part summarizes the main content of this paper.

2. Preliminaries
2.1. *Notations*

In this subsection, we introduce some notations, which will be used in the sequel. $\square^n$ is the set of all vectors of dimension n; $|X|$ is the cardinality of set X; $\mathrm{M}_{m \times n}$ is the set of $m \times n$ matrices; $M_{(i,j)}$ is the (i,j) element of matrix M; $Col_j(M)$ is the j-th column of matrix M; $Col(M)$ is the set of all columns of matrix M;

$$0_n := \underbrace{[0,0,\cdots,0]}_{n} \; ; \quad 1_n := \underbrace{[1,1,\cdots,1]}_{n} \; ; \quad \delta_n^0 := \underbrace{[0,0,\cdots,0]}_{n}^T \; ;$$

$\delta_n^k := Col_k(I_n)$, $1 \le k \le n$; $\Delta_n := \left\{ \delta_n^1, \delta_n^2, \cdots, \delta_n^n \right\}$; $\tilde{\Delta}_n := \left\{ \delta_n^0, \delta_n^1, \cdots, \delta_n^n \right\}$; $L \in \mathrm{M}_{m \times n}$ is a logical matrix (resp., generalised logical matrix) if $Col(L) \subseteq \Delta_m$ (resp., $Col(L) \subseteq \tilde{\Delta}_m$). We denote the set of $m \times n$ logical matrices (resp., generalised logical matrix) by $\mathrm{L}_{m \times n}$ (resp., $\tilde{\mathrm{L}}_{m \times n}$); If matrix $L \in \tilde{\mathrm{L}}_{m \times n}$, then it can be expressed as $L \in \left[\delta_m^{i_1}, \delta_m^{i_2}, \cdots, \delta_m^{i_n} \right]$ and it is briefly denoted as $L \in \delta_m[i_1, i_2, \cdots, i_n]$, where $i_k \in \{0, 1, \cdots, m\}$, $1 \le k \le n$.

2.2. *Semi-tensor product (STP) of matrices*

In this subsection, we give some necessary basic knowledge of the STP of matrices used in the paper.

Definition 2.1 ([4]): Let $A \in \mathrm{M}_{m \times n}$, $B \in \mathrm{M}_{p \times q}$. The STP of A and B is defined as
$$A \cdot B = (A \otimes I_{t/n})(B \otimes I_{t/p}), \tag{1}$$

where t denotes the least common multiple of n and p, i.e., $t = lcm(n, p)$; $\otimes$ is the Kronecker product.

Remark 2.1: When $n = p$, $A \cdot B = AB$. Hence, the STP is a generalization of the standard matrix product. Throughout this paper the matrix product is assumed to be the STP. We mostly omit the symbol "$\cdot$" hereinafter.

Definition 2.2 ([4]): A swap matrix $W_{[m,n]}$ is a $mn \times mn$ logical matrix, which is defined as

$$W_{[m,n]} = \delta_{mn}[1, m+1, 2m+1, \cdots, (n-1)m+1,$$
$$2, m+2, 2m+2, \cdots, (n-1)m+2, \quad (2)$$
$$\cdots, m, 2m, 3m, \cdots, nm].$$

Lemma 2.1([4]): Let $X \in \square^m$ and $Y \in \square^n$ be two column vectors. Then

$$W_{[m,n]}XY = YX, \quad W_{[n,m]}YX = XY. \quad (3)$$

2.3. System model

In this subsection we recall the formalism used in the paper.

A nondeterministic DES is modeled as a NFA $G = (X, \Sigma, \delta, X_0, X_m)$, where X is the finite set of states, Σ is the finite set of events called alphabet or input symbols, $X_0 \subseteq X$ is the set of initial states, $X_m \subseteq X$ is the set of marked states (or accepted states), $\delta : X \times \Sigma \to 2^X$ is the partial transition function(2^X denotes the power set of X), which describes the system dynamics: given states $x, y \in X$ and an event $\sigma \in \Sigma$, $y \in \delta(x, \sigma)$ means the execution of σ from state x takes the system to state y. Note that $\delta(x, \sigma)$ is undefined when the event σ cannot be executed from the state x. $\delta(x, \sigma)!$ denotes $\delta(x, \sigma)$ is well-defined. Obviously, the transition function can be extended to $\delta : X \times \Sigma^* \to 2^X$ in terms of $\delta(x, e) := \delta(\delta(\cdots \delta(\delta(x, e_{j_1}), e_{j_2}) \cdots), e_{j_t})$, where $e = e_{j_1} e_{j_2} \cdots e_{j_t} \in \Sigma^*$, Σ^* denotes the set of finite strings on the alphabet Σ, including the empty string $\cdot$. The objective of this paper is to investigate the controlled NFA. In this regard, the event set Σ can be partitioned into two disjoint subsets, i.e., $\Sigma = \Sigma_c \bigcup \Sigma_{uc}$, where Σ_c denotes the set of controllable events, Σ_{uc} denotes the set of uncontrollable events. We here assume that all events in Σ are observable.

In general, we wish to adjoin a supervisor or a controller S to interact with G in a feedback manner. More precisely, the transition function of G can be controlled by S in the sense that the controllable events of G can be dynamically enabled or disabled by S after each transition. Formally, a *state-feedback supervisor*, denoted by S, is a function $S : X \to 2^{\Sigma_c}$ that determines the set of events $S(x) \subseteq \Sigma_c$ to be disabled at each state $x \in X$, while events not belonging to the set $S(x)$ remain enabled at state x. The *controlled system* (or called *supervised system*) consisting of G and S, denoted by S/G, is another nondeterministic finite automation given as

$$S/G = (X, \Sigma, \delta_S, X_0, X_m), \quad (4)$$

where X, Σ, X_0 and X_m are as defined above, δ_S is the partial transition function of S/G, i.e.,

$$\delta_S(x, \sigma) = \begin{cases} \delta(x, \sigma), if \ \delta(x, \sigma)! \ and \ \sigma \notin S(x) \\ undefined, \ otherwise \end{cases} \quad (5)$$

We use the notation $H(x)$ to denote the set of feasible events of G at state x. Thus supervisor S is called permissible if for all $x \in X$, $S(x) \subseteq H(x) \bigcap \Sigma_c$. Note that it is not difficult to see that NFA can be viewed as special case of controlled NFA with $S(x) = \varnothing$ for all $x \in X$. The controlled NFA S/G is depicted in Fig. 1.
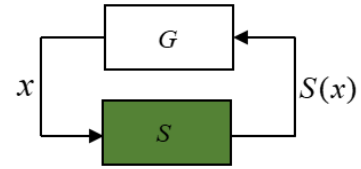
Fig. 1. The controlled NFA S/G

3. Matrix-Based Expression of Controlled NFA

In order to obtain the matrix expression of the dynamics of controlled NFA, let us first give an equivalent description of the controlled system (4). For the NFA G, we define a control pattern as a Boolean function: $\gamma : \Sigma_c \to \{0,1\}$ and we use the notation $\Gamma = \{0,1\}^{\Sigma_c}$ to denote the set of all Boolean functions on Σ_c. $\gamma \in \Gamma$ is interpreted as follows: for any $\sigma \in \Sigma_c$, $\gamma(\sigma) = 1$ means that the control pattern γ allows σ to happen, while $\gamma(\sigma) = 0$ means that the control pattern γ refuse σ to happen. Note that it is convenient to extend each $\gamma \in \Gamma$ to a function $\gamma : \Sigma \to \{0,1\}$ by defining $\gamma(\sigma) = 1$ for each uncontrollable event $\sigma \in \Sigma_{uc}$. Further, we define a partial function of the form $f : X \to \Gamma$, called the *state-feedback control function* or *state-feedback* for short, that maps each state x in X into control pattern γ, i.e., $\gamma(\sigma) = f(x)(\sigma)$. Thus the controlled NFA consisting of the state-feedback f, control pattern γ and NFA G is described as

$$G_\gamma^f = (X, \Sigma, \delta_\gamma^f, X_0, X_m), \quad (6)$$

where X, Σ, X_0 and X_m are as defined above; δ_γ^f denotes the partial transition function of G_γ^f, which is defined as

$$\delta_\gamma^f(x,\sigma) = \begin{cases} \delta(x,\sigma), & \text{if } \delta(x,\sigma)! \text{ and } f(x)(\sigma) = 1 \\ \text{undefined, otherwise.} \end{cases} \quad (7)$$

Similarly, the state-feedback f is called permissible if for all $x \in X$, $\sigma \in H(x) \bigcap \Sigma_{uc}$, we have $f(x)(\sigma) = 1$. There is a bijective correspondence between the state-feedback f and the state-feedback supervisor S, i.e., $\forall x \in X$ and $\forall \sigma \in H(x) \bigcap \Sigma_c$, $f(x)(\sigma) = 1$ (resp., $f(x)(\sigma) = 0$) if and only if $\sigma \notin S(x)$ (resp., $\sigma \in S(x)$). In this regard, we will use for simplicity the controlled system (6) instead of system (4) to present our results. Our objective, in this section, is to model the controlled NFA in the framework of the STP of matrices.

Let us consider the controlled NFA(6), we assume that the set of states is $X = \{x_1, x_2, \cdots, x_n\}$, and the set of events is $\Sigma = \{e_1, e_2, \cdots, e_m\}$. To obtain the dynamics of (6), identifying $x_i \Box \delta_n^i (1 \le i \le n)$, $e_j \Box \delta_m^j (1 \le j \le m)$, we call δ_n^i and δ_m^j the vector forms of x_i and e_j, respectively. Thus the state set X and event set Σ can be identified with Δ_n and Δ_m, respectively, where $\Delta_n = \{\delta_n^1, \delta_n^2, \cdots, \delta_n^n\}$ and $\Delta_m = \{\delta_m^1, \delta_m^2, \cdots, \delta_m^m\}$. Therefore, $x_s \in \delta_c^f(x_t, e_j)$ can be expressed equivalently as $\delta_n^s \in \delta_c^f(\delta_n^t, \delta_m^j)$, where δ_c^f represents the partial transition function of G_γ^f defined in (7). For brevity, we use the notation r_{ij} to denote the state-feedback $f(x_i)(e_j)$, i.e., $r_{ij} = f(x_i)(e_j)$ where $x_i \in X$ and $e_j \in \Sigma_c$ are the state and controllable event of controlled NFA(6), respectively.

For event $e_j (1 \le j \le m)$, we define a $n \times n$ matrix F_j as

$$F_{j(s,t)} = \begin{cases} 1, & \text{if } \delta_n^s \in \delta_\gamma^f(\delta_n^t, \delta_m^j) \wedge \delta_m^j \Box e_j \in \Sigma_{uc} \\ r_{ij}, & \text{if } \delta_n^s \in \delta(\delta_n^t, \delta_m^j) \wedge \delta_m^j \Box e_j \in \Sigma_c \\ 0, & \text{otherwise,} \end{cases} \quad (8)$$

where F_j is called transition structure matrix *w.r.t.* event e_j. Thus, the *transition structure matrix* (TSM) of controlled NFA (6) is defined as

$$F = [F_1, F_2, \cdots, F_m], \quad (9)$$

where F is a $n \times mn$ symbol matrix.

Based on the above representations and the STP of matrices, we can obtain the following result on the matrix expression of dynamics of controlled NFA (6).

Theorem 3.1: Given a controlled NFA (6), the dynamics of (6) can be equivalently described as

$$x(t+1) = Fu(t)x(t), \quad (10)$$

where F is the TSM of system (6), which is defined in (9); $x(t) = (x_1(t), x_2(t), \cdots, x_n(t))^T$ is the vector form of state at step t, $x_i(t)$ denotes the number of different paths from the set of initial states to the state $x_i = \delta_n^i$ with a feasible event string of length $t-1$; $u(t) \in \Delta_m$ is vector form of event at step t.

Let us assume (10) holds when $t = p-1$, i.e., $x(p) = Fu(p-1)x(p-1)$. And let $u(p) = \delta_m^k$ for any $1 \le k \le m$. Then

$$Fu(p)x(p) = F_k(x_1(p), \cdots, x_n(p))^T$$
$$= (\sum_{i=1}^{n} F_{k(1,i)} x_i(p), \cdots, \sum_{i=1}^{n} F_{k(n,i)} x_i(p))^T.$$

Using (8), one has

$$\sum_{i=1}^{n} F_{k(j,i)} x_i(p) = \sum_{i \in \Lambda_j} F_{k(j,i)} x_i(p), 1 \le j \le n,$$

where $\Lambda_j = \{i \mid 1 \le i \le n, \delta_n^j \in \delta_\gamma^f(\delta_n^i, \delta_m^k)\}$. From the above assumption, we know that $x_i(p)$ stands for the number of different paths from the initial state $x(1) = \delta_n^{i_0}$ to state δ_n^i with a feasible string of length $p-1$, then $\sum_{i \in \Lambda_j} F_{k(j,i)} x_i(p)$ represents the sum of the number of different paths by which the initial state set $x(1) = \sum_{\lambda=1}^{l} \delta_n^{i_\lambda}$ can reach state δ_n^i with a feasible string of length $p-1$, and δ_n^i can reach δ_n^j with the event $u(p) = \delta_m^k$. Consequently, we define $x_j(p+1) := \sum_{i \in \Lambda_j} F_{k(j,i)} x_i(p)$, which implies $x(p+1) = Fu(p)x(p)$. By mathematical induction, the proof is completed.

By Theorem 3.1, we know readily that verifying whether or not any two states of controlled NFA (6) are reachable can be determined by both the transition function δ defined in G and the state-feedback f. In particular, when the state-feedback f is known, the matrix F becomes a constant matrix. More concretely, F is a generalised logical matrix. In this case, we replace F with F_c. Therefore, by Theorem 3.1, we have the following result.

Corollary 3.1: Given a controlled NFA (6) in which the state-feedback f is known, the dynamics of (6) can be

equivalently described by the following equation

$$x(t+1) = F_c u(t)x(t), \quad (11)$$

where $x(t)$ and $u(t)$ have the same interpretation as in Theorem 3.1; F_c is the TSM of (6), which is represented as

$$F_c = \left[F_1^c, F_2^c, \cdots, F_m^c \right] \in \tilde{L}_{n \times mn}, \qquad (12)$$

where the $n \times n$ matrix F_j^c is defined as

$$F_{j(s,t)}^c = \begin{cases} 1, & \text{if } \delta_n^s \in \delta_\gamma^f(\delta_n^t, \delta_m^j) \\ 0, & \text{otherwise.} \end{cases} \qquad (13)$$

Remark 3.1: It should be pointed out that our approach are also suitable for the controlled deterministic DESs in the sense that deterministic DESs can be viewed as special case of nondeterministic DESs. In particular, when the controlled DES (6) is deterministic, i.e., $|X_0| = 1$ and $|\delta(x, \sigma)| \leq 1$ for all $x \in X$ and $\sigma \in \Sigma$, the state that is reachable from initial state with a feasible event string $e = \sum_{k=1}^{t-1} e_{jk} \in \Sigma^*$ is unique, which means that there exists only one $1 \leq i \leq n$ such that $x_i(t) = 1$ in (11). Namely, $x(t) \in \Delta_n$. In contrast, if there is no a state that is reachable from initial state with $e = \sum_{k=1}^{t-1} e_{jk} \in \Sigma^*$, then we have $x(t) = \delta_n^0$.

4. Reachability of Controlled NFA

4.1. *Reachability verification*

We, in this subsection, verify the reachability property of controlled NFA (6) by means of equation (11). We first give the following lemma, which is crucial to find all paths from the set of initial states to any target state for the controlled NFA (6).

Lemma 4.1 ([5]): Let $\bullet_{i=1}^t \delta_{m_i}^{j_i} = \delta_{m_1 \times m_2 \cdots m_t}^l$, then the formulas $\delta_{m_i}^{j_i} = S_i \bullet \delta_{m_1 \times m_2 \cdots m_t}^l$, $i = 1, 2, \cdots, t$ hold, where

$$\begin{cases} S_1 = I_{m_1} \otimes 1_{m_2 \times \cdots \times m_t}, \\ S_2 = [\underbrace{I_{m_2} \otimes 1_{m_3 \times \cdots \times m_t}, \cdots, I_{m_2} \otimes 1_{m_3 \times \cdots \times m_t}}_{m_2}], \\ \vdots \\ S_{t-1} = [\underbrace{I_{m_{t-1}} \otimes 1_{m_t}, \cdots, I_{m_{t-1}} \otimes 1_{m_t}}_{m_2 \times \cdots \times m_{t-1}}], \\ S_t = [\underbrace{I_{m_t}, \cdots, I_{m_t}}_{m_2 \times \cdots \times m_t}]. \end{cases} \qquad (14)$$

To present the main results on reachability of the controlled NFA (6), we for brevity need to introduce the following notation: let $\alpha = (a_1, a_2, \cdots, a_n)^T$ be a nonnegative column vector of dimension n, we define $\Xi(\alpha) := \{\delta_n^k \mid a_k \neq 0, 1 \leq k \leq n\}$. For instance, let $\alpha = (1, 2, 0, 3)^T$, then $\Xi(\alpha) = \{\delta_4^1, \delta_4^2, \delta_4^4\}$.

Theorem 4.1: Given a controlled NFA (6) with its dynamics (11), let $X_0 = \{\delta_n^{i_1}, \delta_n^{i_2}, \cdots, \delta_n^{i_l}\}$ and $x^* = \delta_n^q$ be the set of initial states and target state of (6), respectively. Then

1) $x^* = \delta_n^q$ is reachable from X_0 in t steps if and only if there exists a positive integer $k (1 \leq k \leq m^t)$ such that

$$\delta_n^q \in \Xi(Col_k((F_c W_{[n,m]})^t x(1))), \qquad (15)$$

where $x(1)$ represents the vector form of the initial state set X_0, i.e., $x(1) = \sum_{\lambda=1}^l \delta_n^{i_\lambda}$.

2) Assume that $L_t^q(X_0)$ denotes the set consisting of all event strings of length t by which the initial state set X_0 can reach $x^* = \delta_n^q$ in t steps, then

$$L_t^q(X_0) = \{e = e_{k_1} e_{k_1} \cdots e_{k_t} \in \Sigma^* \mid \text{there is } k \text{ of satisfying} \\ (15) \text{ such that } \delta_{m^t}^k = e_{k_1} e_{k_1} \cdots e_{k_t}\}, \qquad (16)$$

where the feasible event string of $e = e_{k_1} e_{k_1} \cdots e_{k_t}$ can be easily obtained from the formula $\delta_{m^t}^k = e_{k_1} e_{k_1} \cdots e_{k_t}$ in terms of Lemma 4.1.

Proof: we here omit the proof of Lemma 4.1.

From (11), we have

$$\begin{aligned} x(t+1) &= F_c u(t) x(t) \\ &= F_c W_{[n,m]} x(t) u(t) \\ &= (F_c W_{[n,m]})^2 x(t-1) u(t-1) u(t) \\ &\vdots \\ &= (F_c W_{[n,m]})^t x(1) u(1) u(2) \cdots u(t) \\ &= ((F_c W_{[n,m]})^t x(1)) \bullet_{j=1}^t u(j). \end{aligned} \qquad (17)$$

Note that the following corollary gives a criterion to verify whether any two states of a controlled NFA are reachable or not. Also, an effective approach of finding all paths of any two reachable states is provided.

Corollary 4.1: Given a controlled NFA (6) with its dynamics (11), let $x_p = \delta_n^p$ and $x^q = \delta_n^q$ be any two states of (6). Then

1) $x^q = \delta_n^q$ is reachable from $x_p = \delta_n^p$ in t steps if and only if there exists a positive integer $k (1 \leq k \leq m^t)$ such that

$$\delta_n^q \in \Xi(Col_k((F_c W_{[n,m]})^t \delta_n^p)). \qquad (18)$$

2) Let $L_t^q(p)$ be the set consisting of all event strings of length t by which $x_p = \delta_n^p$ can reach $x^q = \delta_n^q$ in t steps, then

$L_t^q(p) = \{e = e_{k_1}e_{k_1}\cdots e_{k_t} \in \Sigma^* \mid$ there exist k of satisfying

(18) such that $\delta_{m^t}^k = e_{k_1}e_{k_1}\cdots e_{k_t}\}$,　　　(19)

where the event string $e = e_{k_1}e_{k_1}\cdots e_{k_t}$ can be obtained from $\delta_{m^t}^k = e_{k_1}e_{k_1}\cdots e_{k_t}$ in terms of Lemma 4.1.

Proof: Obviously, here we omit its proof.

Remark 4.1: From Corollary 4.1, we readily know that if the controlled NFA (6) is deterministic, then the number of different paths from state $x_p = \delta_n^p$ to $x^q = \delta_n^q$ in t steps coincide with the number of k satisfying (18). In contrast, when the controlled NFA (6) is nondeterministic, we for convenience assume that

$$\Upsilon = (\gamma_1, \gamma_2, \cdots, \gamma_n)^T := \sum_{j=1}^{m^t} Col_j((F_c W_{[n,m]})^t \delta_n^p) \ .$$ Then the

number of different paths from $x_p = \delta_n^p$ to $x^q = \delta_n^q$ in t steps is equal to γ_q.

4.2. *Comparison with the existing approaches*

There are several papers addressing modeling and reachability analysis in DESs in terms of the algebraic state-space approaches, see, e.g., [3] and [6]. Among them, [3] investigated matrix expression and reachability verification of finite automata in which all events are controllable and the cardinality of the set of initial states equals 1. We know that the controlled DESs have more complex structures and dynamics than the uncontrolled DESs as an external control input is allowed to intervene the behaviors of original systems. Again, they did not consider to adjoin a supervisor to interact with an original system. These restrictions significantly simplify the behavior of the system. Obviously, their approaches are different completely from the approach using in this paper since we do not make the above-mentioned these restrictions.

[6] is most closely related to our work, but there are some fatal errors therein. Concretely, the authors asserted that they presented a sufficient and necessary condition to verify whether or not any two states of a controlled DES are reachable (see, e.g., Theorem 2 in [6]). Also, an algorithm was designed to find all paths from one state to another one if they are reachable (see, e.g., Algorithm 1). These results were obtained from Corollary 1 in [6]. It should be pointed out that, however, Corollary 1 addressed the problem of which states are reachable from a given state and a given input string of length t for a controlled DES, which means a contradiction with the previous assertion. Furthermore, the state-feedback control specification was defined as $f : X \to \Gamma$ in [6]. In this regard, $r_j = f(\tilde{F}^t \ltimes \delta_n^i \quad u(j))(\sigma_{j+1})$ is undefined in (11) of [6] since $\tilde{F}^t \ltimes \delta_n^i \quad u(j)$ is not necessarily the vector form of a state for a controlled nondeterministic system, which means that equation (11) in [6] is incorrect. We refer the reader to [6] for the interpretations of these notations. As a part of our work, these errors presented in [6] were corrected and the interesting issues mentioned in [6] were discussed systematically in a more generalized DES.

5. Illustrative Example

In this section, we use an example to illustrate the proposed results.

Example 5.1 : Let us consider the uncontrolled NFA, $G = (X, \Sigma, \delta, X_0, X_m)$ depicted in Fig. 2, where $\Sigma_c = \{e_2, e_3\}$, $\Sigma_{uc} = \{e_1, e_4\}$, $X_0 = \{e_1, e_4\}$, $X_m = \{x_5\}$. The state-feedback supervisor S is given as $S(x_3) = \{e_2, e_3\}$.
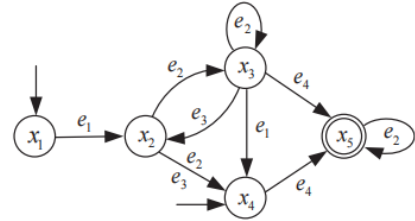
Fig. 2. An uncontrolled NFA.

We now verify the state reachability of the controlled system S/G consisting of the uncontrolled NFA G shown in Fig.2 and the state-feedback supervisor S.

By Corollary 3.1, the dynamics of the uncontrolled system S/G can be expressed as

$$x(t+1) = F_c u(t)x(t), \tag{20}$$

where $x(1) = \sum_{j=1,4} \delta_n^j$,

$$F_c = \begin{bmatrix} 0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&1&0&0&0&1&0&0&0&0&1&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&1&1&0 \end{bmatrix}.$$

When $t = 1$, we have

$$Col_1(F_c W_{[5,4]}x(1)) = (0,1,0,0,0)^T,$$
$$Col_4(F_c W_{[5,4]}x(1)) = (0,0,0,0,1)^T,$$
$$Col_j(F_c W_{[5,4]}x(1)) = (0,0,0,0,0)^T, j = 2,3.$$

By applying Theorem 4.1, states x_2 and x_5 are reachable from the set of initial states X_0 in 1 step, and the corresponding event strings of length 1 are $e = e_1$ and $e = e_4$, respectively. Specifically, we have $x_1 \xrightarrow{e_1} x_2$; $x_4 \xrightarrow{e_4} x_5$.

When $t = 2$, we have

$$Col_2((F_c W_{[5,4]})^2 x(1)) = (0,0,1,1,0)^T,$$
$$Col_3((F_c W_{[5,4]})^2 x(1)) = (0,0,0,1,0)^T,$$
$$Col_{14}((F_c W_{[5,4]})^2 x(1)) = (0,0,0,0,1)^T.$$

with all the other columns equal to δ_5^0. By Theorem 4.1, states x_3, x_4 and x_5 are reachable from X_0 in 2 steps, and the corresponding event string of length 2 are $e = \delta_{16}^2 = \delta_4^1 \delta_4^2$, $e = \delta_{16}^2 = \delta_4^1 \delta_4^2$ or $e = \delta_{16}^3 = \delta_4^1 \delta_4^3$, and $e = \delta_{16}^{14} = \delta_4^4 \delta_4^2$, respectively. Specifically, we have

$$x_1 \xrightarrow{e_1} x_2 \xrightarrow{e_2} x_3 \ ; \quad x_1 \xrightarrow{e_1} x_2 \xrightarrow{e_2} x_4 \quad \text{or} \quad x_1 \xrightarrow{e_1} x_2 \xrightarrow{e_3} x_4 \ ; \quad \text{and}$$
$$x_4 \xrightarrow{e_4} x_5 \xrightarrow{e_2} x_5 \ .$$

When $t = 3$, we have

$$Col_5((F_c W_{[5,4]})^3 x(1)) = (0,0,0,1,0)^T,$$
$$Col_8((F_c W_{[5,4]})^3 x(1)) = (0,0,0,0,2)^T,$$
$$Col_{12}((F_c W_{[5,4]})^3 x(1)) = (0,0,0,0,1)^T,$$
$$Col_{54}((F_c W_{[5,4]})^3 x(1)) = (0,0,0,0,1)^T$$

with all the other columns equal to δ_5^0. Using Theorem 4.1, states x_4 and x_5 are reachable from X_0 in 3 steps, and the corresponding event string of length 3 are $e = \delta_{64}^5 = \delta_4^1 \delta_4^2 \delta_4^1$, $e = \delta_{64}^8 = \delta_4^1 \delta_4^2 \delta_4^4$ or $e = \delta_{64}^{12} = \delta_4^1 \delta_4^3 \delta_4^4$, and $e = \delta_{64}^{54} = \delta_4^4 \delta_4^2 \delta_4^2$, respectively. Specifically, we have

$$x_1 \xrightarrow{e_1} x_2 \xrightarrow{e_2} x_3 \xrightarrow{e_1} x_4; \quad x_1 \xrightarrow{e_1} x_2 \xrightarrow{e_2} x_3 \xrightarrow{e_1} x_5;$$
$$x_1 \xrightarrow{e_1} x_2 \xrightarrow{e_2} x_4 \xrightarrow{e_4} x_5, \quad x_1 \xrightarrow{e_1} x_2 \xrightarrow{e_3} x_4 \xrightarrow{e_4} x_5;$$
$$x_4 \xrightarrow{e_4} x_5 \xrightarrow{e_2} x_5 \xrightarrow{e_2} x_5 \ .$$

The cases of $t \geq 4$ are similar to $t = 1, 2, 3$, the details are omitted here for space limitations.

6. Conclusion

In this paper, we proposed a new framework that is matrix-based form to model the dynamics of controlled NFA. Using it, we investigated the verifications of reachability property of controlled NFA. Also, the criteria of verifying this property was presented in terms of the methodology described in this paper.

Future work will concentrate on investigating the problem of synthesizing property-enforcing supervisor for controlled nondeterministic DESs by means of the proposed theoretical framework.

References

1. R. Jacob, J.J. Lesage, and J.M. Faure. Overview of discrete event systems opacity: Models, validation, and quantification. Annual Reviews in Control, 41:135–146, 2016.
2. C.G. Cassandras and S. Lafortune. Introduction to disctete event systems. Springer, Switzerland AG, 3rd Edition, 2021.
3. X. R. Xu and Y. G. Hong, "Matrix expression and reachability analysis of finite automata," J. Control Theory Appl., vol. 10, no. 2, pp. 210–215, 2012.
4. D. Z. Cheng, H. S. Qi, and Y. Zhao, An Introduction to Semi-tensor Product of Matrices and Its Applications. Singapore: World Scientific, 2012.
5. J. J. Wang, X. G. Han, Z. Q. Chen, et al, "Model matching of input-output asynchronous sequential machines based on the semi-tensor product of matrices," Future Generation Computer Systems: The International Journal of eScience, vol. 83, pp. 468-475, 2018.
6. Y. Y. Yan, Z. Q. Chen, J. Y. Yue, et al, "STP approach to model controlled automata with application to reachability analysis of DEDS," Asian J. Control, vol. 18, no. 6, pp. 1–10, Nov. 2016.

Authors Introduction

Ms. Zipei Wang

She received the B.S. degree from College of Electronic Information and Automation, Tianjin University of Science and Technology, China in 2023. She is currently a Master course student in Tianjin University of Science and Technology.

Mr. Xiang Ren

He received his B.S. degree from College of Electronic Information and Automation, Tianjin University of Science and Technology, China in 2021. He is currently a Master course student in Tianjin University of Science and Technology.

Detection and Identification of Daylily Maturity Based on YOLOv8

Fangyan Li

College of Electronic Information and Automation, Tianjin University of Science and Technology, 300222，China

E-mail: 3162129540@qq.com

www.tust.edu.cn

Abstract

To better apply object detection and identification techniques from deep learning to the field of agricultural automation, this paper focuses on the growth process of daylilies. It employs the state-of-the-art YOLOv8 model to achieve accurate assessment of daylily maturity. The backbone network of YOLOv8 draws inspiration from the CSPDarkNet network structure to extract image features. The introduction of cross-stage connections in the middle layers of the network enhances the efficiency of feature propagation. The neck section adopts the PAFPN bidirectional channel network to further process the features extracted by the backbone network, facilitating the smooth transmission of information from the bottom to the top layers. The Head section utilizes the Decoupled-Head structure to generate the output for object detection. The trained model had a mAP which reached up to 94%, an accuracy rate up to 95.6%, a recall rate up to 90.5%. Moreover, the identification speed is significantly improved.

Keywords: Deep Learning, Object Detection, YOLOv8, Daylily

1. Introduction

With the gradual maturity of AI technology, the focus of agricultural harvesting is gradually shifting to automation. The agricultural sector has been continuously exploring how to leverage modern computer vision technology to enhance crop production efficiency and quality. In the production and management of agricultural products, understanding the maturity of plants is crucial for achieving optimal harvesting and resource utilization. As an important edible plant, the timely monitoring and accurate assessment of the maturity of daylilies are essential for agricultural production [1]. Traditional harvesting of daylilies typically relies on manual inspection and experiential judgment, which poses challenges in terms of high labor costs and low efficiency. With the rapid development of computer vision and deep learning technologies, object detection techniques have gradually become an effective means to address this issue. For instance, Song [2] constructed and trained a Faster R-

CNN model implemented by VGG16, which demonstrated good detection performance for fruit images under different time and lighting conditions. Liang [3] proposed improvements on the SSD network based on VGG16 and ZFNet for real-time mango detection. The enhanced algorithm achieved higher accuracy compared to Faster R-CNN, although the deep network structure led to slower computation speed. Zhen [4] introduced a vegetable image classification method, utilizing the Caffe open-source deep learning framework. They improved the VGG network by adding a Batch Normalization layer to the VGG-M network, enhancing the convergence speed and accuracy of the network (VGG-M-BN) with a vegetable classification recognition accuracy of up to 96.5%. Xue [5] aimed to improve the accuracy of immature mango detection in orchard scenes. They modified the YOLO V2 algorithm, developing a tiny-yolo network structure by combining multiple feature layers. This addressed the challenging issue of detecting overlapping and occluded parts of mangoes, achieving a detection accuracy of

97.02%. Zhao [6] proposed an apple localization method based on the YOLOv3 deep convolutional neural network. This method ensures the detection of apples in complex environments while balancing efficiency and accuracy.

The main contributions of this paper include the construction of a large-scale daylily image dataset, detailed annotation of daylily features at different growth stages, and the design of an architecture and training methodology based on the YOLOv8 model. Through experiments, the proposed approach has been demonstrated to exhibit high accuracy and practicality in the automatic detection and identification of daylily maturity.

2. Data Acquisition and Annotation

2.1. Data acquisition

The datasets for this experiment are all derived from manual field shooting with an f/1.6 wide-angle camera. Image categories encompass immature daylilies, mature daylilies, and overly mature daylilies. The dataset includes single and multiple targets, unobstructed and clustered daylily samples, captured under different weather conditions (sunny and cloudy) and at various time intervals. After acquiring a sufficient amount of video footage, the videos were decomposed frame by frame and converted into static image data. To curate the dataset, images with minimal variations in daylily features within the same second or in close proximity were excluded, as well as those that lacked specific images of daylily maturity flowers. The initial dataset was then refined by retaining images with distinct characteristics and clearly identifiable daylily buds. In order to expand the diversity of data, this paper uses the ACE image enhancement algorithm and histogram equalization algorithm to enhance the daylily dataset. In the end, we obtained a dataset of 2060 daylily images.

2.2. Images annotation and dataset production

To ensure the accuracy of training data, prior to experimentation, the daylily dataset was meticulously annotated, defining precise labels for optimal training outcomes. The Labelme annotation software was uniformly used for annotating daylily images. Based on characteristics such as color and shape of daylily flowers, the images were categorized into three classes: immature daylily, mature daylily, and overmature daylily. Immature daylilies, characterized by unopened flower buds, exhibit deep green or green coloration. These buds are relatively small and not ready for harvesting. Harvestable daylilies have bright yellow-green buds, larger in size, and fully developed, ready for picking. Overly mature daylilies have either opened or wilted buds, displaying bright yellow or dark yellow coloration with petal-like structures. The annotation process is shown in Fig. 1. The marked data is a JSON file, which is converted to a txt file . The dataset was split into 80% for training and 20% for testing and validation collectively. The testing set was utilized to assess the model's generalization capabilities post-training, while the validation set was employed for fine-tuning hyperparameters and conducting preliminary assessments of the model's performance.

Fig. 1 Dataset labels

3. Methodology

This study selected YOLOv8n as the foundational framework for research. YOLOv8, being at the forefront of object detection technology, is recognized for its efficient detection speed and accuracy across various object detection tasks. The algorithm takes the input image and resizes it to a fixed size, dividing the resized image into a grid of cells. Each cell is responsible for predicting objects that fall within its boundaries. For each cell, the algorithm predicts B bounding boxes and the probability C for each category, each containing 5 values: (x, y, w, h, c). These (x, y) values represent the center of the bounding box relative to the cell boundary, and (w, h) represent the width and height of the bounding box. This C value indicates the confidence of the algorithm in the prediction. Non-maximal suppression (NMS) is applied to remove redundant bounding box predictions. NMS compares the overlap of the predicted bounding box, leaving only the

one with the highest confidence, and IoU evaluates the accuracy of the bounding box prediction from 0 to 1, with higher values indicating better overlap between the predicted bounding box and the true bounding box. The

offering an innovative and intelligent solution for agriculture.

The YOLOv8 model comprises three main components: the backbone, neck, and head network.CSPDarknet53 serves as the backbone network for YOLOv8, primarily responsible for extracting high-level features from input images. It extracts features at multiple levels and outputs feature maps at different scales. These feature maps possess varying receptive fields and semantic information, aiding the model in better handling targets of different sizes and complexities. In contrast to earlier versions of YOLOv5, YOLOv8 eliminates the less user-friendly Focus module, and the initial network layers are directly accomplished by straightforward and conventional convolutions. CSPDarknet53 is an improvement upon Darknet53, introducing the CSPNet (Cross Stage Partial Network) structure. This structure involves cross-stage partial connections, enhancing information propagation efficiency. It divides feature maps into two parts, with one part directly passed to the next stage, and the other part subjected to further processing. This contributes to strengthening information propagation and feature fusion within the network.

The neck still employs the PaFPN (Path Aggregation Feature Pyramid Network) structure to construct the YOLO feature pyramid, facilitating comprehensive fusion of multi-scale information. In comparison to YOLOv5, the only difference lies in the reduction of a 1x1 convolutional layer in the upsampling operation during the top-down process. Additionally, the C3 module has been replaced with the C2f module. Notably, the channel numbers returned from the three scales are kept equal to the channel numbers output by the backbone for the three scales.

The Head section adopts a decoupled structure, separating the classification and detection heads. Two parallel branches independently extract category and position features, followed by a 1×1 convolution layer for each to perform classification and localization tasks. Simultaneously, the use of anchor boxes is discarded.

The loss calculation process comprises two main components: positive and negative sample assignment strategy and loss computation. Considering the superior performance of dynamic assignment strategies, YOLOv8

final output is the detection result, including the category label, bounding box coordinates, and confidence score.

Leveraging its capabilities, the study applies YOLOv8 to the detection of daylily maturity, algorithm directly incorporates TOOD's Task Aligned Assigner. It selects positive samples based on the weighted scores of classification and regression. Loss computation includes classification and regression. Binary Cross Entropy (BCE) Loss is still employed for classification. For the regression branch, it is bound to the integral representation introduced in Distribution Focal Loss, thus utilizing Distribution Focal Loss and CIoU Loss. The three losses are weighted using a certain proportion. The final training results are generated accordingly.

4. Experiments

4.1. Experimental environment

To ensure the rigor of the experiments, both the training and testing processes in this paper were conducted on the same experimental platform. The experimental environment was built and configured based on Ubuntu 18.04, with a NVIDIA GeForce GTX 1070 graphics card, Intel Core i5-9400F CPU, and programming platform Anaconda 2022.05. The CUDA version used was 11.2, and the deep learning framework employed was PyTorch 1.10 +CUDA11.1 support. The programming language used was Python 3.7.

4.2. Model training and evaluation

Through multiple experiments in the early stage, the maximum convergence of daylily target detection model training was obtained. The optimal number of iterations was 400 times, and 400 iterations were used for model training to ensure the consistency of experimental parameters. The basic parameters of the model are set to epoch=400 and batch size=16, and the initial learning rate is set to 0.01 using the SGD optimizer.

In this paper, an objective criterion was used to evaluate the daylily identification line. we use precision (P), recall (R), and mAP (mean average precision) to evaluate the model trained in this paper. As shown in Equation (1).

$$P = \frac{TP}{TP + FP}$$

$$R = \frac{TP}{TP + FN} \quad (1)$$

$$mAP = \frac{1}{C}\sum_{k=i}^{N} P(k)\Delta R(k)$$

where TP is the true number of positive samples and FP is the false number of positive samples, FN is the number
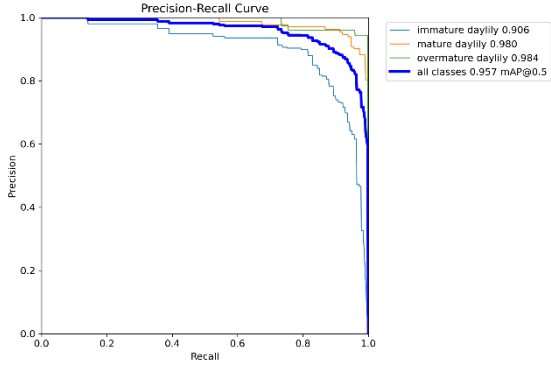
of values, k is the threshold, P(k) is the accuracy, and R(k) is the recall.

5. Result

The results of different species of daylily tested on this model are shown in Fig. 2. The P-R curve and R curve of the model are shown in Fig. 3.

of spurious negative samples, C is the number of categories, and N is the reference threshold. The number
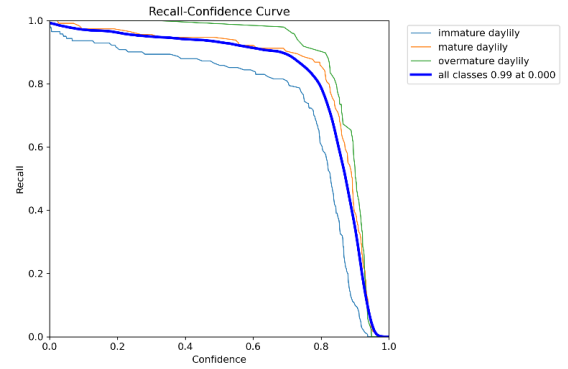
(a) P-R Curve

(b) R-Curve

Fig. 2 Detection result of the algorithm
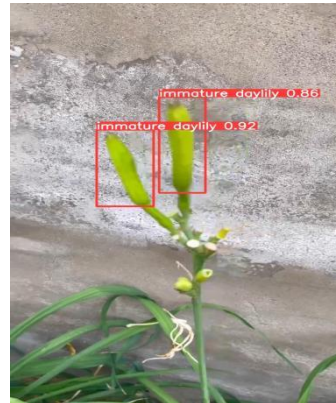
Fig. 3 Performance parameters

6. Conclusion

After completing the model training, we conducted validation on daylilies at different growth stages. Experimental results indicate that the proposed method exhibits high accuracy and robustness in detecting and

identifying the maturity of daylilies. The accuracy achieved is above 95.6%, with a recall rate of 90.5% and a mean Average Precision (mAP) of 0.94. Furthermore, our model demonstrates stability under varying lighting conditions and significant detection performance across daylilies with different shapes and colors.

Our model not only provides a reliable solution for the intelligent cultivation of daylilies but also offers effective support for agricultural decision-making. The stability of performance in diverse lighting conditions and its ability to detect daylilies of varying shapes and colors make it a valuable tool for farmers. This research lays the foundation for agricultural intelligence, enhancing the productivity and decision-making capabilities of farmers.

References

1. Zhang X, Ma H, Wei S, et al. Design of Day-Lily Robot Recognition System Based on ZYNQ. *International Conference on Networking, Communications and Information Technology (NetCIT).* IEEE, 2021: 154-157.
2. SONG ZH ZH，FU L SH，WU J ZH，et al． Kiwifruit detection in field images using Faster Ｒ-CNN with VGG16． IFAC-PapersOnLine，2019，52（30）: 76-81.
3. LIANG Q K，ZHU W，LONG J Y，et al． A real-time detection framework for on-tree mango based on SSD network. *International Conference on Intelligent Ｒobotics and Applications*．*Springer*，*Cham*，2018: 423-436.
4. Zhenbo Li, Fei Li, Ling Zhu, et al. Vegetable Recognition and Classification Based on Improved VGG Deep Learning Network Model. International Journal of Computational Intelligence Systems, 2020, 13(1) : 559-564.
5. XUE Y J，HUANG N，TU SH Q，et al． Immature mango detection based on improved YOLOv2. Transactions of the Chinese Society of Agricultural Engineering，2018，34(7): 173-179.
6. ZHAO D AN，WU Ｒ D，LIU X Y，et al． Apple positioning based on YOLO deep convolutional neural network for picking robot in complex background. Transactions of the Chinese Society of Agricultural Engineering，2019，35(3) : 164-173.

Authors Introduction

Ms. Fangyan Li

She is the scond-year graduate student of Tianjin University of Science and Technology. Her research direction is machine Learning.

Chaos Synchronization and Circuit Design of Chen System and Lü System with Different Structures

Haozhe Sun

College of Electronic Information and Automation, Tianjin University of Science and Technology, 300222, China

E-mail: 2921407938@qq.com

www.tust.edu.cn

Abstract

In this paper, by using nonlinear feedback control, chaos synchronization is achieved between the Chen system and the Lü system with different initial values, and the error curves and state synchronization curves of the corresponding states in the response Lü system and the drive Chen system are plotted. Finally, the simulation circuit model of the synchronization system of the drive Chen system and the response Lü system is designed by Multisim circuit simulation software. Comparing the output curves with the curves obtained by MATLAB simulation software, it can be found that the two curves achieve a good qualitative agreement. It prove that the circuit model of the synchronization system is correct and the synchronization of the drive Chen system and the response Lü system is accomplished.

Keywords: Chaos; Chaos synchronization of different systems; Nonlinear feedback control; Chaotic circuit

1. Introduction

In 1999, Professor Guanrong Chen of the University of Houston discovered a new chaotic attractor, the Chen system, using engineering feedback control [1]. which is similar to the Lorenz system but not topologically equivalent, with a more complex structure and more diverse dynamic behavior.In 2001, Jinhu Lü and Guanrong Chen discovered a new chaotic system, the Lü system, that system links the Lorenz system and the Chen system and represents a continuous evolution between the two systems [2].

In this paper, the above two systems are taken as the research objects to analyze their chaotic properties, and based on this, the two systems are synchronized to chaos by using the method of nonlinear feedback [3], and then numerical simulation is carried out by using the MATLAB software, and the simulated circuits are designed, and the circuit simulation is carried out in Multisim, which is designed to provide sufficient theoretical support for the future application of the heterostructured chaos synchronization of the Chen system and the Lü system to be successfully applied in the field of confidential communication.

2. Chaos Synchronization and Circuit Design of Chen System and Lü System with Different Structures

2.1. Modeling of the Drive Chen system

$$\begin{cases} \dot{x}_1 = a_1(y_1 - x_1) \\ \dot{y}_1 = (c_1 - a_1)x_1 - x_1 z_1 + c_1 y_1 \\ \dot{z}_1 = x_1 y_1 - b_1 z_1 \end{cases}$$

$$(1)$$

Where, x_1, y_1, $z_1 \in R$ is the state variable of the system and $a_1 = 35$, $b_1 = 3$, $c_1 = 28$ is the typical parameter of Lü system.

2.2. Modeling of Response Lü Systems

$$\begin{cases} \dot{x}_2 = a_2(y_2 - x_2) + u_{c1} \\ \dot{y}_2 = -x_2 z_2 + c_2 y_2 + u_{c2} \\ \dot{z}_2 = x_2 y_2 - b_2 z_2 + u_{c3} \end{cases}$$

(2)

Where, x_2, y_2, $z_2 \in R$ is the state variable of the system and $a_2 = 36$, $b_2 = 3$, $c_2 = 20$ is the typical parameter of Lü system. $u_c = [u_{c1}\, u_{c2}\, u_{c3}]^T$ is the required synchronization controller, through which the synchronization of two chaotic attractors with very different structures can be realized.

2.3. Design of Synchronization Controller

Subtracting the corresponding terms in the mathematical model Eq. (1) of the driving Chen system from the terms in the mathematical model Eq. (2) of the response Lü system, the error system model is shown as Eq. (3) below:

$$\begin{cases} \dot{e}_1 = \dot{x}_2 - \dot{x}_1 \\ \quad = a_2(y_2 - x_2) - a_1(y_1 - x_1) + u_{c1} \\ \dot{e}_2 = \dot{y}_2 - \dot{y}_1 \\ \quad = -x_2 z_2 + c_2 y_2 \\ \quad -[(c_1 - a_1)x_1 - x_1 z_1 + c_1 y_1] + u_{c2} \\ \dot{e}_3 = \dot{z}_2 - \dot{z}_1 \\ \quad = x_2 y_2 - b_2 z_2 \\ \quad -(x_1 y_1 - b_1 z_1) + u_{c3} \end{cases}$$

(3)

Where $e = [e_1 \quad e_2 \quad e_3]^T = [x_2 - x_1 \quad y_2 - y_1 \quad z_2 - z_1]^T$, e_1, e_2, e_3 is the state variable of the error system. The synchronization controller is designed as:

$$\begin{cases} u_{c1} = (a_1 - a_2)(y_1 - x_1) - k_1 e_1 \\ u_{c2} = x_2 z_2 - x_1 z_1 + (c_1 - c_2)y_1 \\ \quad +(c_1 - a_1)x_2 - k_2 e_2 \\ u_{c3} = -x_2 y_2 + x_1 y_1 + (b_2 - b_1)z_1 \\ \quad -k_3 e_3 \end{cases}$$

(4)

Where k_1, k_2, $k_3 \geq 0$, Eq. (4) is substituted into Eq. (3), and the mathematical model of the error system is obtained as shown in Eq. (5).

$$\begin{cases} \dot{e}_1 = \dot{x}_2 - \dot{x}_1 \\ \quad = a_2(e_2 - e_1) - k_1 e_1 \\ \dot{e}_2 = \dot{y}_2 - \dot{y}_1 \\ \quad = c_2 e_2 + (c_1 - a_1)e_1 - k_2 e_2 \\ \dot{e}_3 = \dot{z}_2 - \dot{z}_1 = -b_2 e_3 - k_3 e_3 \end{cases}$$

(5)

The error system is then written in the form of $\dot{e} = f(e, t)$. If the equilibrium state of error system at the origin is uniformly asymptotic and stable in a large range, it indicates that the chaos synchronization of different structure has been realized between Chen system and Lü system.

The value of the synchronization controller parameter $k = [k_1 \quad k_2 \quad k_3]^T$ is determined by the Lyapunov second method.

According to Lyapunov's second method, firstly take the positive definite Lyapunov function $V(\tilde{e}) = 1/2[(a_1 - c_1)e_1^2/a_2 + e_2^2 + e_3^2]$, and then take its derivative:

$$\begin{aligned} V(\tilde{e}) &= \frac{a_1 - c_1}{a_2}e_1 \dot{e}_1 + e_2 \dot{e}_2 + e_3 \dot{e}_3 \\ &= (a_1 - c_1)e_1 e_2 - (a_1 - c_1)e_1^2 \\ &\quad -\frac{k_1(a_1 - c_1)}{a_2}e_1^2 + c_2 e_2^2 \\ &\quad +(c_1 - a_1)e_1 e_2 - k_2 e_2^2 \\ &\quad -b_2 e_3^2 - k_3 e_3^2 \\ &= -\left[\frac{k_1(a_1 - c_1)}{a_2} + (a_1 - c_1)\right]e_1^2 \\ &\quad -(k_2 - c_2)e_2^2 - (k_3 + b_2)e_3^2 \end{aligned}$$

(6)

According to Lyapunov's second method, if $V(\tilde{e})$ is positively definite and $\dot{V}(\tilde{e})$ is negatively definite, then the error system is uniformly asymptotically stable in a large range at the origin. At this time, the variable coefficients of the error system should satisfy:

$$\begin{cases} -[\frac{k_1(a_1 - c_1)}{a_2} + (a_1 - c_1)] < 0 \\ -(k_2 - c_2) < 0 \\ -(k_3 + b_2) < 0 \end{cases}$$

(7)

Solved:

$$\begin{cases} k_1 > -a_2 = -36 \\ k_2 > c_2 = 20 \\ k_3 > -b_2 = -3 \end{cases}$$

(8)

If $(k_1, k_2, k_3) = (0, 25, 0)$, the synchronous controller u_c is as follows:

$$\begin{cases} u_{c1} = -(y_1 - x_1) \\ u_{c2} = x_2 z_2 - x_1 z_1 + 33 y_1 \\ \quad -7x_2 - 25 y_2 \\ u_{c3} = -x_2 y_2 + x_1 y_1 \end{cases}$$

(9)

So far the synchronization controller has been designed.

The numerical simulation diagram of the chaotic synchronization controller of Chen system and Lü system with different structures built in Simulink is shown in Fig. 1.

After setting up the simulation environment, click the run button, and the corresponding state error curve is shown in Fig. 2.

The corresponding state synchronization curve is shown in Fig. 3.

Through careful observation of the images, it can be seen that the controller makes the corresponding state error curve gradually converge to zero with the passage of time, and the corresponding state synchronization curve follows the running trajectory of Chen driving system. This is enough to indicate that all the above derivations of all equations on the different structure chaos synchronization of Chen system and Lü system are correct, and Chen system and Lü system can achieve different structure chaos synchronization.
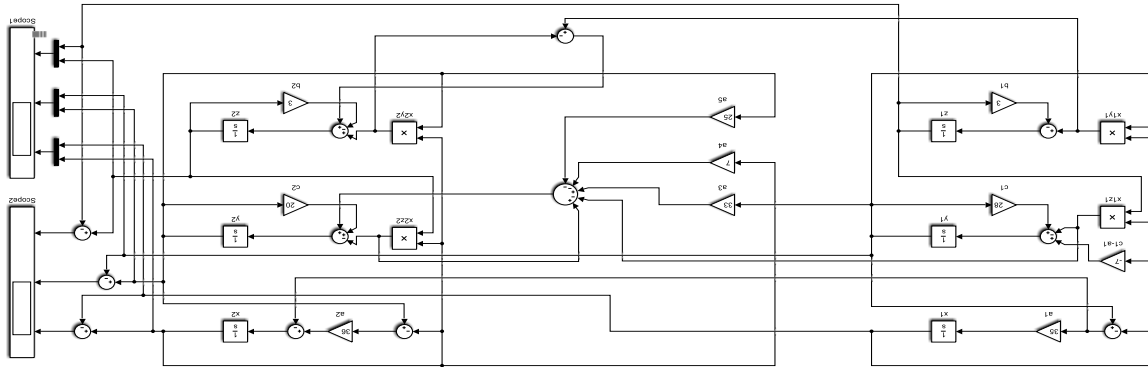
Fig. 1 Numerical simulation of chaos synchronization of Chen system and Lü system with different structure

Fig. 2 State error curve: (a) e_1; (b) e_2; (c) e_3
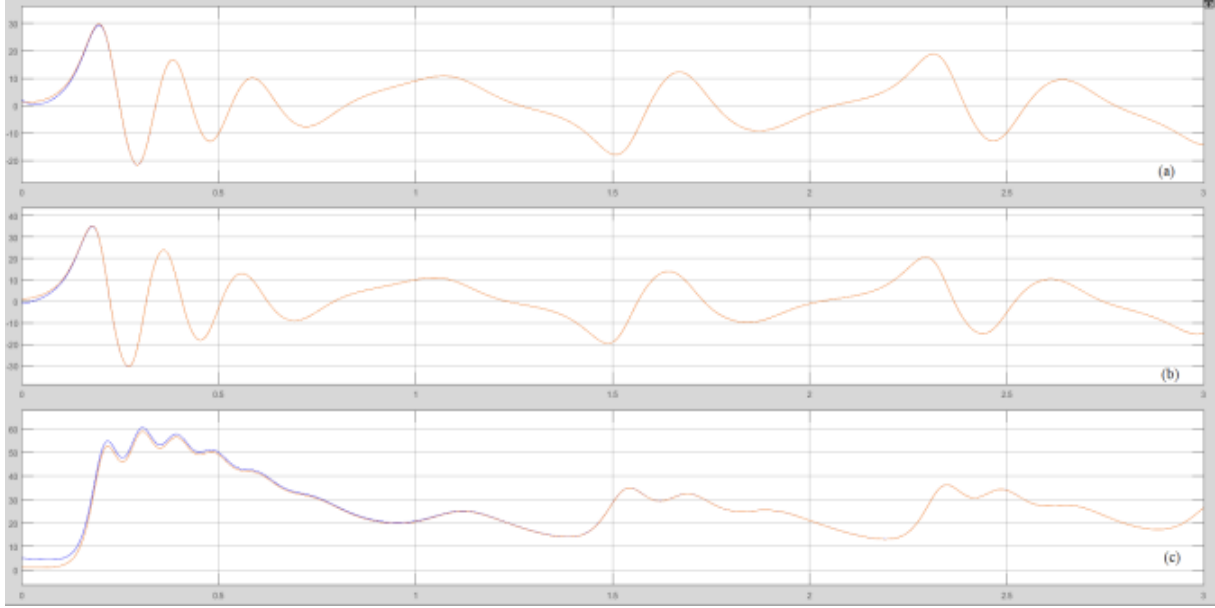
Fig. 3 State synchronization curve: （a）x1-x2；（b）y1-y2；（c）z1-z2

2.4. Design and construction of an analog circuit model for heterostructure chaotic synchronous control

The circuit model of the synchronous controller is designed and constructed, and the circuit of the synchronous controller is designed by the method of improved modular circuit design.

The state expression of the control circuit is shown in Eq. (10).

$$\begin{aligned}
\tau_0 u_{c1} &= \tau_0[-(y_1 - x_1)] \\
&= -\tau_0 y_1 - \tau_0(-x_1) \\
&= -\frac{1}{R_{14}C_4}y_1 - \frac{1}{R_{15}C_4}(-x_1) \\
\tau_0 u_{c2} &= \tau_0[10x_2z_2 - 10x_1z_1 \\
&\quad +33y_1 - 7x_2 - 25y_2] \\
&= -10\tau_0(-x_2z_2) - 10\tau_0 x_1 z_1 \\
&\quad -33\tau_0(-y_1) - 7\tau_0 x_2 - 25\tau_0 y_2 \\
&= -\frac{1}{10R_{16}C_5}(-x_2z_2) \\
&\quad -\frac{1}{10R_{17}C_5}x_1 z_1 - \frac{1}{R_{13}C_5}(-y_1) \\
&\quad -\frac{1}{R_{19}C_5}x_2 - \frac{1}{R_{20}C_5}y_2 \\
\tau_0 u_{c3} &= \tau_0(-10x_2y_2 + 10x_1y_1) \\
&= -10\tau_0 x_2 y_2 - 10\tau_0(-x_1y_1) \\
&= -\frac{1}{10R_{21}C_6}x_2 y_2 \\
&\quad -\frac{1}{10R_{22}C_6}(-x_1y_1)
\end{aligned}$$
(10)

Substitute $\tau_0 = 100$ into Eq. (10) and take $C_4 = C_5 = C_6 = 10\text{nF}$, and the calculation is as follows:

$$\begin{aligned}
R_{14} = R_{15} &= \frac{1}{\tau_0 C_4} = \frac{1}{100 \times 10 \times 10^{-9}} \\
&= 1\text{M}\Omega \\
R_{18} &= \frac{1}{33\tau_0 C_5} = \frac{1}{33 \times 100 \times 10 \times 10^{-9}} \\
&= 30.3\text{k}\Omega \\
R_{19} &= \frac{1}{7\tau_0 C_5} = \frac{1}{7 \times 100 \times 10 \times 10^{-9}} \\
&= 142.8\text{k}\Omega \\
R_{20} &= \frac{1}{25\tau_0 C_5} = \frac{1}{25 \times 100 \times 10 \times 10^{-9}} \\
&= 40\text{k}\Omega \\
R_{16} = R_{17} = R_{21} = R_{22} &= \frac{1}{100\tau_0 C_5} \\
&= \frac{1}{100 \times 100 \times 10 \times 10^{-9}} = 10\text{k}\Omega
\end{aligned}$$
(11)

The analog circuit for the synchronous control of Chen system and Lü system is shown in Fig. 4.

Build the analog circuit shown in the above figure in Multisim software. Wait for the oscilloscope to display the waveform, observe the corresponding waveform, and obtain the corresponding synchronization curve shown in Fig. 5 and the corresponding error curve shown in Fig. 6.
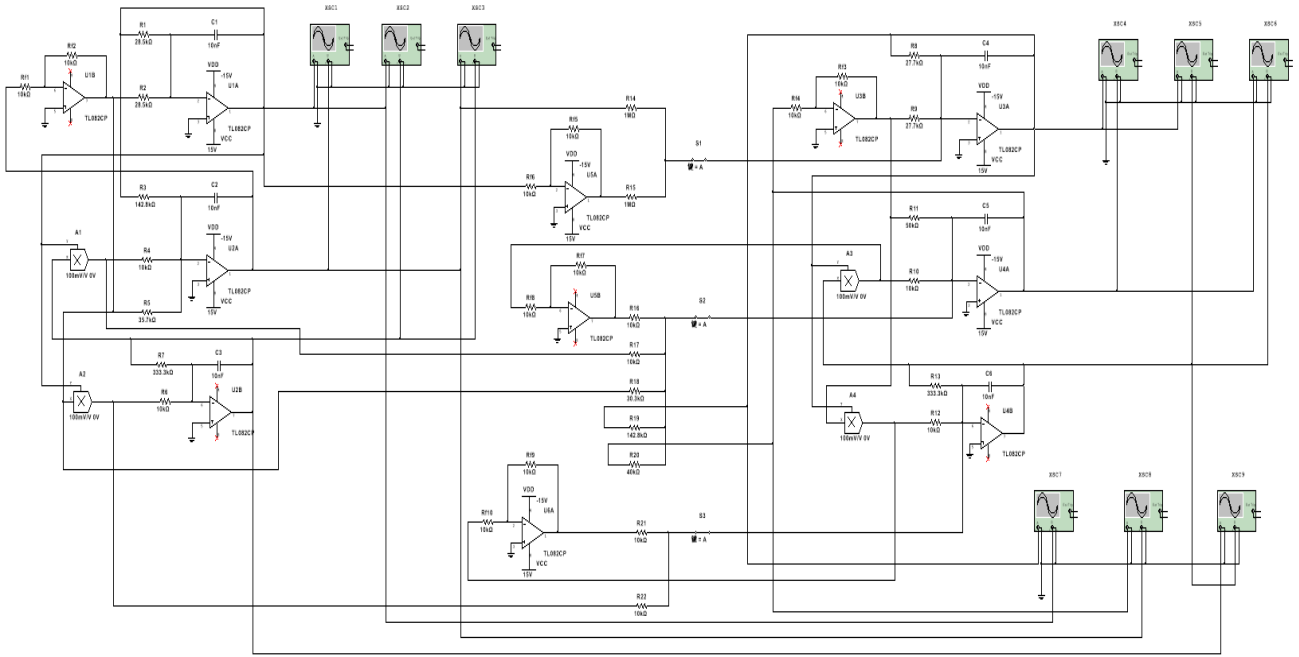
Fig. 4 Analog Circuit for Synchronous Control of Chen System and Lü System
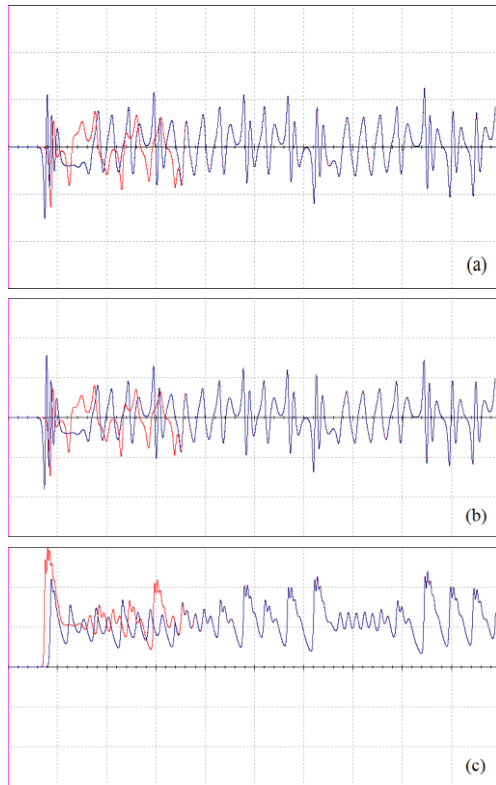
Fig. 5 Synchronization state curve:（a）x_1-x_2；（b）y_1-y_2；（c）z_1-z_2
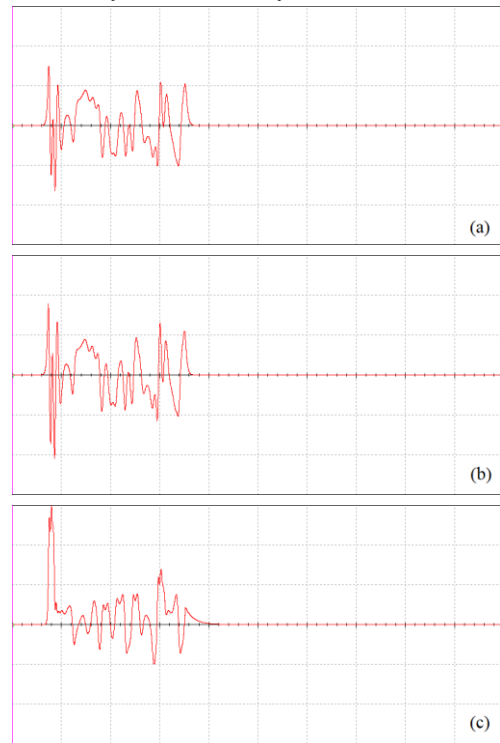
Fig. 6 Error curve:（a）e_1；（b）e_2；（c）e_3

3. Conclusion

In this paper, the driving Chen system and the responding Lü system are first mathematically modeled separately,

and then the synchronous controllers are designed to achieve the purpose of removing all nonlinear terms. The parameters of the controller are solved according to Lyapunov's second method. The controller allows the numerical simulation of the system in MATLAB to show that the corresponding state error curves gradually converge to zero over time and the corresponding state synchronization curves follow the trajectory of the Chen drive system.

The simulation circuit model of the Chen chaotic system and the Lü chaotic system with heterostructure chaotic synchronization is designed and constructed with the help of Multisim software using a modified modular approach. The installation of separate switches for the structural compensator and the controller to ensure individual control is a critical step in the design of the system, and the observation of the oscilloscope showing the corresponding error curves reveals that the curves quickly converge to 0 after the switches are closed, and the whole experiment shows that the designed synchronization controllers perform the control role. The heterostructured chaotic synchronization control of Chen system with Lü system and its circuit design are implemented.

References

1. Guopeng Zhou, Xiaoxin Liao, Bingji Xu, Pei Yu, Guanrong Chen. Simple algebraic necessary and sufficient conditions for Lyapunov stability of a Chen system and their applications. Transactions of the Institute of Measurement and Control, 2018, 40(7).
2. Jiakun Zhao, Di Zhou, Yexin Li. A new impulsive synchronization of Chen hyper-chaotic system and Lü hyper-chaotic system. Journal of Vibration and Control, 2013, 19(12).
3. W. Kinzel, A. Englert, I. Kanter. On chaos synchronization and secure Communication. Philosophical Transactions Of The Royal Society A, 2010, 368, 379–389.

Authors Introduction

Mr. Haozhe Sun

In 2023, he received his Bachelor of Engineering degree from the School of Electronic Information and Automation, Tianjin University of Science and Technology, China. He is pursuing a master's degree in engineering from Tianjin University of Science and Technology.

A Parking Space Recognition Method Based on Digital Image Technology

Hao He*, Fangyv Liu, Jiaxin Wang
Tianjin University of Science and Technology, Tianjin,300222, China
E-Mail: *hehao0310252022@163.com

Abstract

In recent years, the number of cars in the city has been increasing, leading to an increasingly prominent issue of urban parking space. Consequently, the automatic identification method for parking spaces has emerged as a crucial research direction. This paper presents a design and implementation scheme for recognizing the status of parking spaces in urban areas based on digital image processing and other technologies. In this paper, the real-time pictures of multiple parking Spaces are collected to identify the parking space, so as to determine the use of the parking space. The experimental results demonstrate the feasibility and effectiveness of the proposed method, which holds practical significance in addressing parking space detection problems through digital image processing.

Keywords: OpenCV, parking spot detection, convolutional neural network, intelligent transportation

1. Introduction

In recent years, with the popularity of family cars, urban road traffic and urban parking space management problems become increasingly prominent.

In the new intelligent parking management system, the method of identification of parking space status is also improved with the progress of technology. At first, all kinds of sensors are used to recognize and judge the status of parking Spaces, but these methods have disadvantages: ultrasonic recognition method makes the error of the sensor larger because there are multiple echoes. The method of infrared identification of parking space status is affected by the change of ambient light source and heat source, resulting in high misjudgment rate. Geomagnetic detection is extremely inconvenient to install.

In response to the above problems, experts in related fields began to look for new solutions. A new research hotspot has been formed. According to the research findings, it has certain advantages to use digital image technology for real-time monitoring of parking lots and to analyze and judge the use of parking Spaces within the scope of video surveillance through image algorithms. Therefore, this paper studies the parking space recognition method based on digital image technology.

This paper is divided into four parts. The first part is the background. The second part is an introduction to the development environment, briefly introducing Python, OpenCV and CNN. The third part introduces the system process, including pre-processing, parking space division, CNN model construction and parking space detection. The fourth part summarizes the main content of this paper.

2. Introduction to development environment

2.1. Introduction to Python

Python was designed in the early 1990s by Guido van Rosum of the National Center for Mathematical and Computer Science Research in the Netherlands as a replacement for a language called ABC. Python provides efficient high-level data structures, as well as easy and effective object-oriented programming. Python's syntax and dynamic typing, as well as the nature of the interpreted language, make it a programming language for scripting and rapid application development on most platforms, as versions are updated and new features of the language are added, it is gradually used for independent, large-scale project development [1]. Python is one of the most popular programming languages for beginners, and the Python interpreter is easy to extend with new functions and data types in C, C++, or other languages that can be called through C. Python can also be used as an extension language in customizable software. Python's rich standard library provides source code or machine code for every major system platform.

2.2. Introduction to OpenCV

OpenCV is an open-source library for machine vision and machine learning, based on Apache2.0. It boasts excellent compatibility and can be deployed across multiple operating systems including Linux, Windows, Android, and Mac OS. Comprising both C and C++ languages as a whole, it also provides interfaces for Python, Ruby, MATLAB, and other programming languages to facilitate development work in various language preferences. OpenCV incorporates a vast array of general algorithms in the fields of image processing

and computer vision. Its functionalities encompass object recognition, face recognition, action recognition, motion tracking among others. As such, it stands as a pivotal software library within the realm of image processing and machine vision [2], [3].

2.3. *Introduction to CNN*

The human visual system is processed hierarchically in the visual cortex of the brain, and the working process of the brain's processing of visual information is a continuous iterative and abstract process. Starting with the original image of pupil intake, some cells in the cerebral cortex find the edge and direction of the object. Then, the brain determines the general shape of the object in front of the eyes and abstracts. Then the big brain further determines the class of the object. Finally, through the analysis of the information, the conclusion is drawn. CNN network, full name Convolutional neural network, is a deep learning model inspired by the human visual nervous system. It can be divided into convolution layer, pooling layer and fully connected layer. Convolution layer is used to extract local features in image. Pooling layer is used to greatly reduce the magnitude of parameters. The fully connected layer is used to output the result. Convolutional neural network (CNN) is a classic algorithm of deep learning, which is widely used in computer vision, natural language processing and other fields, and can better process data information with stable effects such as pixels and audio and video without additional feature engineering requirements.

3. Methodology

The system design process can be roughly divided into pre-processing, parking space division, CNN model construction and training, parking space recognition and other stages. In the pre-processing stage, the information will be processed and only the effective information will be retained. Parking Spaces are extracted and stored as samples in the stage of parking space division. In the model training stage, the two-classification model will be trained to recognize whether there is a car in the parking space. Finally, the images are monitored in real time and the gaps are marked.

3.1. *pretreatment*

Due to the influence of various natural factors, the difference between the target object in the picture and the backscene image is not high, the available information is weak, and the parking space can not be recognized directly. Therefore, a variety of processing of the original image is required. For example, removing noise and enhancing image detail to distinguish the target object from the background image.

In order to facilitate post-processing, background filtering should first be carried out to filter out irrelevant background information in the image and facilitate subsequent recognition. The method is to define the mask, set the maximum and minimum thresholds, and convert information outside the threshold interval into zero value information.

The second step is binary processing, which converts the image into grayscale. This is because the grayscale image only retains the brightness information and omits a lot of irrelevant information, which effectively reduces the amount of calculation in the later stage.

The third step is Canny edge detection, which is divided into four steps: Gaussian filtering, calculating gradient value and gradient direction, filtering non-maximum value, double threshold detection method. It can effectively determine real and potential edges, and can better detect most of the required boundaries [4].

The original image is shown in Fig. 1. and the pre-processed image is shown in Fig. 2.

Fig. 1 Original image
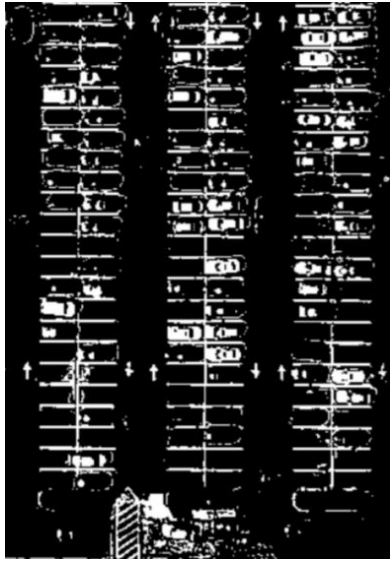
Fig. 2 Pretreatment result

3.2. Parking space division

Through observation, it can be found that the parking lot is composed of different straight lines, which can be further divided by detecting these straight lines. The method is to use Hough transform, a feature detection algorithm widely used in computer vision, to find out some features in graphics. Hough transform obtains more accurate results by giving the type of shapes to be identified, and then determining the shape of the object through voting mechanism in the parameter space, and filtering out objects with too large or too small intervals. Some cracks or other line segments in the image may be mistaken for parking space lines, and the parking space lines are basically vertical or horizontal. The irregular line segments can be further eliminated by calculating the slope of the line segments to improve the accuracy.

3.3. CNN network training

After the parking space image is obtained, it is divided into training set and test set according to the proportion of 70% and 30%, and then divided into occupied set and non-occupied set according to whether or not the vehicle occupies the parking space, which is used for subsequent model training.

This model training uses VGG16 network, VGGNet is a

neural network jointly developed by Oxford University and Google DeepMind in 2014, which is composed of 13 convolutional layers +3 fully connected layers. The convolutional layer is the core of the convolutional neural network. The essence of the convolutional process is based on a small moment matrix, that is, the convolutional kernel. The pixel matrix of each layer of the image is continuously scanned by step size, and the scanned value is multiplied by the number of the convolution check position, and then added and summed. Convolution kernel is equivalent to a filter, used to extract image features, convolution kernel generally has 3*3 and 5*5 sizes, usually used is 3*3 size, training effect is better. The fully connected layer refers to the connection between each node of the N-1 layer and the n layer, that is, the activation function input of the node of the n layer is the weight of all the nodes of the N-1 layer, which is used to synthesize the features extracted from the previous side. It plays the role of "classifier" in the whole convolutional neural network.

Due to the small size of the data set, in order to get a better model, the third step is to carry out data enhancement operations. Data enhancement can be divided into online enhancement and offline enhancement according to the use method. Offline enhancement is suitable for small data sets and can effectively increase their size. The multiplier of enhancement depends on the number of images converted, such as rotating all images to double the data set. There are two types of commonly used data enhancement, supervised and unsupervised. The commonly used one is supervised data enhancement, which is based on the existing data set and expands the existing data by analyzing the completeness of the data and adopting certain rules. Supervised data enhancement can be subdivided into single sample data enhancement and multi-sample data enhancement. In practice, single sample data enhancement is more used.

Finally, the data of the training set is imported into the model for training. The accuracy of the test set obtained after the completion of all programs is 91.4%, and the accuracy rate is consistent with the prediction period. Later, VCC16 can be changed to ResNet50 to further improve the accuracy, but the training speed of ResNet50 is slower than that of VCC16, so from the perspective of efficiency, VCC16 is finally chosen.

The results of parking space detection are shown in Fig. 3.

Fig. 3 Parking test results

4. Conclusion

The program will process and detect the picture in real time, and the vacant parking Spaces in the picture are marked green, and the occupied parking Spaces are not marked. Some areas in some pictures are incorrectly marked, such as the guiding line area of the parking lot is marked as empty parking space, which can be further improved by modifying the CNN model. The detection and statistics of the remaining areas are basically in line with expectations, and the processing and detection of parking space information is realized.

Through OpenCV and convolutional neural network, real-time monitoring of parking space occupancy is realized, which can be applied to urban parking lots for convenient management. In addition, the system can be further optimized to increase speed and adapt to more complex scenarios, thus providing more efficient and timely data for parking lot management.

References

1. Haitao Huang, Python 3 Ice-Breaking Artificial Intelligence from Entry to Actual Combat (in Chinese).
2. Yaling Zhou, Traffic flow detection system based on OpenCV. Beijing: Beijing University of Posts and Telecommunications, 2015.
3. Minglu Zhang, He Xiaodong, Research on Traffic flow statistics Algorithm based on OpenCV. Computer Knowledge and Technology, 2016,12 (19) : 169-170.
4. Mingrui Zhao, Research on parking status detection method based on machine vision. Shenyang: Shenyang Jianzhu University, 2021.

Authors Introduction

Ms. Hao He

She is currently an undergraduate in Tianjin University of Science and Technology. She is majoring in robotics engineering.Her research area is about Image processing based on OpenCV.

Ms. Fangyv Liu

She is currently an undergraduate in Tianjin University of Science and Technology. Her research area is about intelligent robot and intelligent control.

Ms. Jiaxin Wang

She is currently an undergraduate in Tianjin University of Science and Technology. She is majoring in robotics engineering.Her research area is about Image processing based on OpenCV.

A Design of Intelligent Handling Robot Based on AT89C52

Fangyv Liu*, Jiaxin Wang, Hao He

College of Electronic Information and Automation, Tianjin University of Science and Technology, 300222,China

Email:*himawari0216@163.com

Abstract

At present, manual logistics handling is inefficient and limited, manual operation is not only difficult to manage and maintain, but also has certain security risks in the handling process. This paper presents a logistics handling robot based on AT89C52 single chip microcomputer. The integration of the power module, sensor, and drive motor module enables automatic obstacle avoidance and information collection. The incorporation of an ultrasonic obstacle avoidance module and an infrared tracking module enhances the capability for obstacle avoidance and path searching, and can easily cope with different workplace. Handling robot arm structure reference Yaskawa MPL manipulator,the Yaskawa MPL manipulator is highly suitable for high-speed and high-precision palletizing, picking, packaging, and other industries.

Keywords: logistics handling; robot; MCU; modularization; obstacle avoidance

1. Introduction

With the rapid development and application of technology in the 21st century, the logistics transportation and handling industry has also ushered in great changes and development. Manual handling in the logistics production line has many years of history, however, the traditional manual handling has many shortcomings, such as large physical consumption, low degree of automation, low efficiency, environmental impact and so on. Nowadays, many domestic logistics enterprises still use a lot of physical handling and manual operation in the daily logistics handling operation. This situation in the current logistics industry can no longer adapt to the development needs of the market.

Although manual logistics handling still plays an important role in reality, it also faces some status quo and drawbacks, including human resource dependence, the influence of human factors, low efficiency and rising labor costs. Therefore, the use of handling robots can improve the efficiency of logistics handling, reduce costs, improve safety, and provide greater flexibility and adaptability. With the continuous development of technology, handling robots will play an increasingly important role in the logistics industry.

However, the current handling robot still has some shortcomings and still faces certain difficulties in the complex and dynamic environment. For example, objects of different shapes and sizes, different types of ground, and the presence of people and other robots can have an impact on the robot's operation and navigation. The accuracy and flexibility of handling robots still need to be improved. This paper proposes the design of logistics handling robot based on AT89C52 single chip microcomputer, and optimizes the current handling robot on the system to realize the autonomous tracking of the handling robot, avoid obstacles and complete the automatic picking of objects by the robot arm, so as to achieve the purpose of more efficient and safe logistics handling.

The rest of this article is organized as follows.The second part of the design of the robot, first introduces the system design of the robot, and then introduces the design of the drive module and the robot arm according to the system design. The third part is hardware design, which introduces the main hardware used by the robot. The fourth part of the module test, mainly ultrasonic obstacle avoidance module and infrared tracking module for testing, data analysis. The fifth part summarizes the main content of this paper.

2. Design of Handling Robot

2.1. Robot system design

The design of this paper is based on the modular design of AT89C52 microcontroller. The functions of the whole design system are divided into power module, motor drive module, ultrasonic obstacle avoidance module, infrared tracking module and robot arm module. The control of the robot arm is achieved through a vision sensor, through which information about the desired object is obtained [1]. The System composition diagram is shown in Fig. 1.
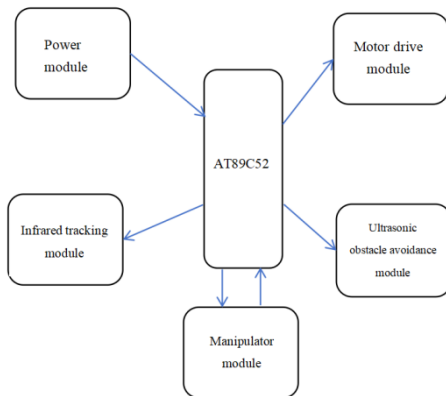
Fig.1 The system composition diagram

2.2. Drive module design

The whole work flow of the handling robot starts from the designated starting work area and gets the route information through the road condition feedback, so as to reach the designated area in accordance with the set route to complete the corresponding work. Therefore, the design of the driving module of the robot in this paper needs an automatic tracking system to realize. The automatic tracking system flow is shown in Fig. 2.
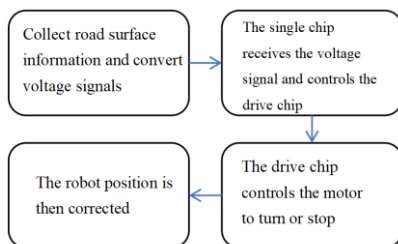
Fig. 2 The automatic tracking system flow

The infrared tracking module uses the TCRT5000 infrared sensor as the path detection sensor to realize the detection of the road track route [2].

2.3. Robot arm design

Mechanical arm structure reference Yaskawa MPL series mechanical arm, Yaskawa MPL series mechanical arm structure is simple, fewer parts, parts of low failure rate, reliable performance, simple maintenance and maintenance, less inventory parts required. The MPL structure has high control performance and precision, can meet the complex motion control needs, and can be flexibly configured and expanded to adapt to different application scenarios and requirements.

The Yaskawa MPL manipulator has a modular design, and individual joints and components can be installed and replaced independently. This design makes repair and maintenance easier and faster, reducing downtime and repair costs. The Yaskawa MPL0100 robot arm model and fixture model is shown in Fig. 3.
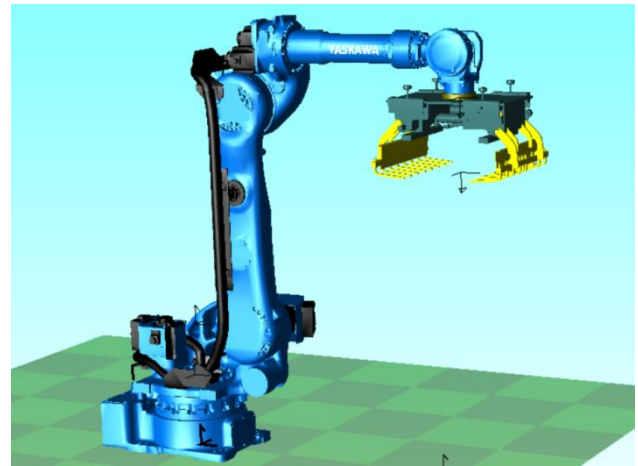
Fig.3 Yaskawa MPL0100 robot arm model and fixture model

3. Hardware Design

3.1. Main control chip

The main controller of the control system uses AT89C52 microcontroller, AT89C52 microcontroller has powerful functions and flexibility, suitable for a variety of embedded systems and control applications, belongs to CMOS8-bit microcontroller. Its advantages are mainly reflected in low power consumption, high performance, and flash memory [3], [4]. Receive the input signal collected by the sensor, and output the control signal to the driver chip, so as to control the robot to walk along the specified route and stop to the specified destination.

3.2. Power supply

In the whole system design, AT89C52 MCU is the control

center, and it is powered by 2 3.7V power modules. Information acquisition through the sensor, its operating voltage requirement is 5V, the use of AMS1117 regulator to achieve voltage regulation, AMS1117 is a cost-effective regulator, relatively low price, stable and reliable performance, At the same time, the AMS1117 regulator has the advantages of simple and easy to use, stable and reliable, low power consumption, overload protection, multiple output voltages and high cost efficiency. The operating voltage of the motor drive module in the device is 5 V, if the voltage is greater than 5 V, it is directly supplied by the power module.

3.3. Drive motor

The robot is usually driven by DC reducer motor, which can output large torque through the deceleration of the gear box [5]. After comparison and selection, the model JGB37-550 DC gear motor is selected. The JGB37-550 DC gear motor has the advantages of high torque, good stability, wide speed regulation range, high energy efficiency, high reliability and easy installation, which makes it widely used in many industrial and commercial applications.

3.4. Controls the robotic arm chip

In order to identify the grasping material workpiece, the color and shape of the material workpiece are used to identify the target. The recognition sensor adopts OpenMV3 vision module with MCU interface. There is a camera on the OpenMV3 module for taking images of objects, and the images taken are processed by AT89C52 microcontroller.

4. Module Testing

4.1. Ultrasonic obstacle avoidance module test

According to the actual working environment of the handling robot, the indoor environment is selected for ultrasonic obstacle avoidance test, obstacles are set, and ECH0 of the HC-SR04 ultrasonic module is connected to the needle J9-P2.0, TRIG is connected to J9-2.1, VCC is connected to J9-VCC, and GND is connected to J9-GND. The LCD1602 liquid crystal was installed in the LCD1602 socket of the single chip microcomputer, and the contrast of the liquid crystal was adjusted. Ultrasonic obstacle avoidance module test data is shown in Table 1.

Table 1. Ultrasonic obstacle avoidance module test data

Ultrasonic ranging(mm)	Ruler ranging(mm)	Error(%)
266.8	266.0	0.15
350.6	350.0	0.17
439.8	439.0	0.18
658.8	658.0	0.12
772.1	771.0	0.14

The error between the ultrasonic obstacle avoidance module's ranging value and the actual ruler's ranging value is less than 2%, which can realize automatic obstacle avoidance.

4.2. Infrared tracking module test

Natural light has a certain influence on the receiving ability of the infrared receiving tube, so the indoor environment test matching the actual use environment is used. Before the test, connect P3.2~3.5 of the pin J11 and P3.2~3.5 of the AT89C52. Place the robot on the white test bench, turn on the power, and fine-tune the potentiometer RW3 and RW4 in a clockwise manner until LED2 and LED light up. Infrared tracking module test data is shown in Table 2.

Table 2. Infrared tracking module test data

Transfer robot	Left infrared tube	Right infrared tube
Go forward	1	1
Back up	0	0
Turn left	1	0
Turn right	0	1

5. Conclusion

This paper takes the handling robot as the research object, uses the AT89C52 single chip microcomputer as the main control chip, uses the infrared tracking module, ultrasonic obstacle avoidance module, power module, motor module and robot arm module to design the handling robot, optimizes the shortcomings of some handling robots in the general working environment, improves the efficiency of logistics handling and reduces the risk of manual logistics handling.

Intelligent handling robot has broad development prospects. With the continuous progress of technology and the expansion of application fields, intelligent handling robots will play an important role in various industries, improve work efficiency, reduce costs, and

promote the intelligent transformation of the industry.

References

1. Luo Xiangxi, Yuan Fengwei, Liu Zhiwei.Design and research of manipulator based on vision orientation. Mechanical Engineer, 2016(12): pp.54-56.
2. Zhang Meng, Shi Baohua.PC motherboard automatic assembly based on machine vision system.China Instrumentation, 2016(9): pp.32-35.
3. LI Y K，WEI S L.Research on automatic control of cooperative operation trajectory of unmanned warehouse multi handling robot.Computer Measurement and Control:1-8[2023-01-05DOI]: 10.16526/j.cnki.11-4762/tp.2023.02.018.
4. WU Y Q,YU T. Design of dual robot coordinated controlsystem. Manufacturing Automation, 2022, 44(10):pp.1-5.
5. Liu Zhenyu, Li Zhongsheng, Zhao Xue,et al.,Research on sorting technology of industrial robot based on machine vision.Manufacturing Automation, 2013, 32(9):pp.25-30.

Authors Introduction

Ms. Fangyv Liu

She is currently an undergraduate in Tianjin University of Science and Technology. Her research area is about intelligent robot and intelligent control.

MS. Jiaxin Wang

She is currently an undergraduate in Tianjin University of Science and Technology. She is majoring in robotics engineering.Her research area is about Image processing based on OpenCV.

MS. Hao He

She is currently an undergraduate in Tianjin University of Science and Technology. She is majoring in robotics engineering.Her research area is about Image processing based on OpenCV.

Application and Differences of Robotic Arms, Traditional Machines, and Manual Work in Production

Xue Yang*, Ying Su, Yuping Mei, Haiquan Wang

College of Electronic Information and Automation, Tianjin University of Science and Technology, 300222，China

E-mail: *wybzdwzdbzd@126.com

Abstract

This article aims to explore the differences between robotic arms and traditional machines and humans. Firstly, robotic arms are a kind of automated equipment with high flexibility and accuracy, which can perform a variety of complex tasks. In contrast, traditional machines lack flexibility and accuracy, while humans have problems such as low work efficiency and high error rates. Secondly, the emergence of robotic arms can solve many problems in traditional machines and humans, and improve production efficiency and quality. Finally, there is no contradiction between robotic arms, traditional machines and humans, but they can complement each other. By using robotic arms and humans reasonably, we can give full play to their respective advantages and improve the overall production efficiency and economic benefits.

Keywords: robotic arm, traditional machine, human, production efficiency, economic benefit

1. Introduction

Robotic arms are a kind of automated equipment with high flexibility and accuracy, which can perform a variety of complex tasks. In contrast, traditional machines lack flexibility and accuracy, while humans have problems such as low work efficiency and high error rates.

The emergence of robotic arms can solve many problems in traditional machines and humans, and improve production efficiency and quality.

There is no contradiction between robotic arms, traditional machines and humans, but they can complement each other. By using robotic arms and humans reasonably, we can give full play to their respective advantages and improve the overall production efficiency and economic benefits.

The rest of this article is organized as follows. The second section introduces the appearance differences between robotic arms and traditional machines. In the third part, the application of robotic arms in manufacturing can solve many problems. In the fourth part, robotic arms and human labor can complement each other. At the same time,

robotic arms also need human supervision and management to ensure their normal operation and production safety. In the future, with the continuous advancement and development of technology, robotic arms will be applied and developed in more fields.

The picture of the robotic arm board is shown in Fig. 1.
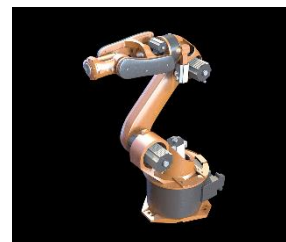
Fig. 1 The picture of the robotic arm

2. The Different Appearances

Xue Yang, Ying Su, Yuping Mei, Haiquan Wang

Machine arms and traditional machines have distinct differences in appearance.

2.1. Traditional machines

Traditional machines often have a rigid mechanical structure, consisting of various metal and plastic components and parts. Their shape and size are relatively fixed, and they can only be produced and processed on a large scale in factories.

The picture of the traditional machines is shown in Fig. 2.
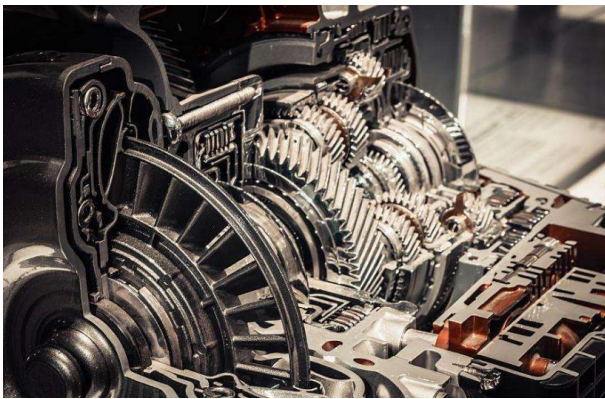
Fig. 2 The picture of the traditional machines

2.2. Robotic arms

The design of robotic arms typically adopt a more flexible mechanical structure, with multiple movable joints and arms, as well as various sensors and actuators at the end. These joints and arms can move independently, allowing the machine arm to adapt to various work environments and tasks. In addition, machine arms are usually equipped with various sensors, such as visual sensors, tactile sensors, and force sensors, to perceive the surrounding environment and perform precise operations.

The picture of the is robotic arms shown in Fig. 3.
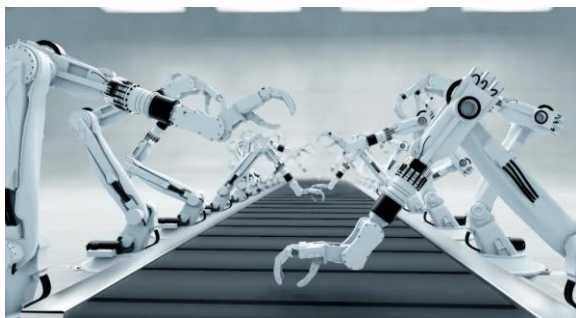
Fig. 3 The picture of the robotic arms

Traditional human production has been unable to keep up with basic standards, which can easily lead to a decline in economic efficiency [1]. Therefore, in appearance, machine arms are more flexible, complex, and intelligent compared to traditional machines

3. The Application of Robotic Arms

The application of robotic arms in manufacturing does provide solutions to many problems. Here is a detailed explanation of some specific applications and advantages of robotic arms in manufacturing.

3.1. Substitution for dangerous and heavy work

In the manufacturing industry, there are many dangerous, heavy, and repetitive tasks that can have negative effects on workers' physical and mental health. Robotic arms can take on these tasks, avoiding direct contact with hazardous environments while reducing their workloads. For example, in the automotive manufacturing process, robotic arms can be used for heavy lifting, welding, and painting, which are more dangerous tasks, thus protecting workers from work-related injuries.

3.2. Improving production quality

The precise control and repeatability of robotic arms enables them to perform production tasks more accurately than humans. Through pre-set programs or precise sensor technology, robotic arms can ensure consistency and product quality during the production process. On assembly lines, robotic arms can precisely assemble and tighten components, ensuring that each product meets specifications and quality requirements.

3.3. 24-hour uninterrupted work

Robotic arms can work continuously without needing to rest or sleep. They can work for 24 hours straight, which improves production efficiency and production cycles. In contrast, humans need to rest and recover and cannot work continuously for long periods of time. Therefore, robotic arms are very useful in situations where high production volume and rapid production are required.

In short, the application of robotic arms in manufacturing can solve many problems, including the replacement of dangerous work, improving production quality, 24-hour uninterrupted work, and increasing production efficiency and economic efficiency. With the

continuous progress and development of technology, the application of robotic arms will become more and more extensive, bringing more innovation and value to the manufacturing industry.

4. The collaboration

4.1. Collaboration between robotic arms and traditional machines

Robots are highly automated and precise machines that can perform a series of repetitive physical tasks. Traditional machines are more versatile machines that can perform a wider range of tasks. By combining robots with traditional machines, we can create a more efficient and flexible production line.

For example, in the manufacturing industry, traditional machines can complete a series of complex processing tasks, while robotic arms can be responsible for repetitive tasks such as handling, assembly, and inspection. This collaborative approach allows traditional machines and robotic arms to each leverage their strengths, thereby improving overall production efficiency. At the same time, this collaboration can also reduce production costs, as robotic arms can replace humans in performing dangerous or high-intensity tasks, reducing the reliance on human labor.

4.2. Cooperation between robotic arm and human

The collaboration between robotic arms and human beings can be divided into two aspects: one is human-computer interaction, and the other is machine-assisted human.

Human-robot interaction refers to the execution of tasks by a robot arm under the guidance of a human. In this mode, humans can control the actions of the robot arm by entering commands or operating it. This collaborative approach is often used for dangerous or high-intensity work, such as deep-sea operations or space exploration. Through human-robot interaction, humans can fully utilize their creativity and flexibility while avoiding some dangerous tasks.

Machine-assisted human labor refers to the use of robotic arms as an assistive tool to help humans complete complex tasks. For example, in the medical field, robotic arms can assist doctors in surgical operations, thereby improving the accuracy and efficiency of the procedure. This collaborative approach can free humans from tedious

or high-intensity work, thereby improving work efficiency and quality.

The picture of the cooperation between robotic arm and human is shown in Fig. 4.

Fig. 4 The picture of the cooperation between robotic arm and human

In general, there is no contradiction between robotic arms, traditional machines, and human labor, but rather a collaborative relationship. By using robotic arms and human labor in a reasonable manner, we can fully leverage their respective strengths and improve overall production efficiency and economic benefits. In the future, as technology continues to advance and innovate, this collaborative approach will become increasingly common and important.

5. Conclusion

Through in-depth research, it is found that there is no contradiction between robotic arms, traditional machines, and human labor, but rather they can collaborate with each other to improve overall production efficiency and economic benefits.

In summary, robotic arms, traditional machines, and human labor each have their own advantages and disadvantages, and are suitable for different production scenarios. In the future, with the continuous advancement and innovation of technology, we can fully leverage their respective strengths to improve overall production efficiency and economic benefits. At the same time, we also need to focus on how to better achieve human-machine collaboration, improve production safety, and product quality.

The picture of the human-machine collaboration is shown in Fig. 5.

Xue Yang, Ying Su, Yuping Mei, Haiquan Wang

Fig. 5 The picture of the human-machine collaboration

Acknowledgements

References

[1] Qi Li. Application analysis of robotic arms in the field of mechanical engineering technology. Technological Innovation and Application, 2021, 14(3): 182-184.

Authors Introduction

Ms. Xue Yang

She is studying in College of Electronic Information and Automation, Tianjin University of Science and Technology, China. Her research interest is Robotics Arm.

Ms. Ying Su

She is studying at Tianjin University of Science and Technology, majoring in Robotics Engineering. Passionate about drone machines.

Ms. Yuming Mei

She is studying at Tianjin University of Science and Technology, majoring in Robotics Engineering. She is very interested in drones.

Mr. Haiquan Wang

He is currently an undergraduate inTianjin University of Science andTechnology. He is majoring in Artificial intelligence. His field of research is deep learning.

A Deep Exploration of the Mounting Issues Related to Six Rotor UVA

Yuping Mei*, **Ying Su**, **Xue Yang**
Tianjin University of Science and Technology, Tianjin, China

E-mail: **2070909305@qq.com*

Abstract

In recent years, the hexacopter UAV has developed rapidly, the technological breakthrough in the field of automatic driving, which is of great significance in both military and civilian fields. The six-rotor UAV uses six rotors as the power source and adjusts the attitude by changing the rotor speed to further achieve position control. It has excellent hovering ability and sensitivity, and is equipped with a precise positioning system and advanced sensors. However, there are still common problems such as weak mounting capacity and single mounting mode, in view of this problem, this paper will test the fuselage structure, avionics system and power system one by one from two aspects: changing the fuselage structure and installing the motor position.

Keywords：the hexacopter UAV, weak mounting capacity, changing the fuselage structure, installing the motor position

1. Introduction

With the continuous development of science and technology, drones have become an important tool in the modern military and civilian fields.

As a new type of UAV, hexagoning UAV has the characteristics of high flight altitude and long endurance, so it has a wide range of application prospects in military reconnaissance, geological exploration, agricultural plant protection and other fields.

However, in actual use, hexagoning UAVs often encounter some technical and performance problems in the process of mounting, such as weak mount capacity and single mount mode, which limits its wider application. Therefore, it is of great significance to conduct in-depth discussion and research on the mounting problem of hexagoned UAV.

2. The Development Process and Characteristics of Hexagoning UAVs

2.1 Development history

The development history of hexagoning UAVs can be traced back to the earliest quadrotors. With the continuous innovation of technology and the continuous expansion of demand, hexarotor UAVs, as a multi-rotor aircraft, have gradually emerged in various fields. Early six-rotor UAVs were mainly used in scientific research and military applications to perform tasks that required high maneuverability and flexibility. With the rise of aerial photography technology, hexacopter UAVs have begun to be widely used in the civilian field for aerial photography, monitoring and survey work. In recent years, the application of hexacopter UAVs in emergency rescue, agriculture, logistics and other fields has also increased.

2.2 Features

Just as shown in Fig. 1, the hexacopter UAVs have the following characteristics. (1) The multi-rotor design: Compared to quadcopter drones, six-rotors employ more

rotors, making them more stable during vertical take-off, landing and hovering. This design improves mobility and adaptability to complex environments, making it ideal for performing specific tasks. (2) High maneuverability: Hexacopter UAVs have excellent maneuverability due to more rotors and are able to flexibly cope with a variety of flight missions. This mobility makes it particularly advantageous when performing search and rescue missions in emergency situations. (3) Load capacity: Hexacopter drones typically have a higher load capacity and are able to carry more sensors, cameras, or other equipment.

Fig. 1 A hexacopter UAV

This makes it stand out for tasks that require large equipment, such as scientific research, monitoring, and surveys. Diversification: Due to its flexibility and stability, the application of hexacopter UAV in various fields is constantly expanding, including but not limited to military reconnaissance, aerial photography, agricultural spraying, cargo transportation, etc., showing diversified application scenarios.

3. The Function of the Six-wing UAV is Realized

3.1 Positioning system

Global Positioning System (GPS): GPS is one of the most common and widely used positioning systems that determines the location, speed, and altitude of a drone through satellite signals. It provides global location information for flight, but in some environments, such as urban canyons or dense woodlands, there may be signal occlusion issues.Inertial

Navigation System (INS): An inertial navigation system uses sensors such as accelerometers and gyroscopes to measure the linear acceleration and angular velocity of an aircraft to estimate its position and orientation. However, errors that accumulate over time can lead to navigation drift and are therefore often used in conjunction with other positioning systems.

Visual positioning system: Hexacopter drones can achieve visual positioning through cameras or multiple cameras. Computer vision algorithms can identify landmarks, landmarks, or environmental structures to

determine the drone's position relative to these reference objects.

3.2. Sensor system

Inertial Measurement Unit (IMU): The IMU includes accelerometers and gyroscopes to measure the linear acceleration and angular velocity of the aircraft, supporting flight control and navigation. Magnetometer: Magnetometers are used to detect the geomagnetic field to help determine the direction and heading of an aircraft.

Barometer: A barometer is used to measure atmospheric pressure, which provides altitude information. It is often used in combination with other sensors to improve the accuracy of altitude measurements.

Millimeter-wave radar: Provides high-precision obstacle detection and ranging capabilities, suitable for flying in complex environments.

Infrared sensor: Used to detect thermal radiation for target detection and identification at night or in low-light conditions. Optical Flow Sensor: Used to measure motion relative to the ground to help achieve hovering and stable flight.

Temperature sensors: monitor the ambient temperature and play an important role in some special applications such as meteorological research, cold chain logistics, etc.

4. Analysis of the Current Situation of the Mounting Problem of Hexagophoid UAV

4.1 Mountability issues

Load capacity limitations: The structure and power system of a hexacopter drone may not be designed to withstand large or heavy equipment, resulting in reduced performance or unstable flight when loaded.

Reduced flight time: Large loads often cause the drone to consume more power, resulting in a lower flight time. The weak payload capacity may limit the durability of the UAV to perform tasks in the air.

Height impacted: Hexacopter drones with heavy loads can lead to increased sensitivity to altitude changes and difficulty in stable flight at different altitudes, especially at low altitudes or in complex environments.

Reduced handling performance: Large or heavy loads may increase the inertia of the drone, making it more sluggish or inflexible when maneuvering, affecting the maneuverability of the flight.

4.2 Single mount type

The reason for the single type of hexacopter UAV mount may involve the following factors:

Technical limitations: When manufacturing hexacopter drones, technical and engineering constraints may result in limited payload types. For example, flight controls, powertrains, or structural design may make it more complex or impractical to support multifunctional mounts.

Cost considerations: Versatile mount systems can add to manufacturing costs and complexity. In order to keep costs under control or to offer a more competitive price, manufacturers may choose to adopt a design with a single mount type.

Regulatory and compliance: Regulatory requirements in specific industries or regions may impose restrictions on the design and use of drones, leading manufacturers to select specific payload types to comply with regulatory requirements.

5. The Solution to the Mounting Problem of the Six-wing UAV

In response to the above problems, we put forward some possible solutions and suggestions:

5.1. Performance optimization

Structural and powertrain optimization：The structure and power system of the drone were redesigned to enhance its load capacity. This could include the use of lighter but stronger materials, optimized motor and propeller configurations, and improved battery technology to provide greater power output. A fixed-time control scheme based on integral terminal sliding mode can be designed to realize the trajectory tracking control of UAV, and the fixed-time convergence performance of the closed-loop control system is analyzed through the Lyapunov theory [1].

Multi-Functional Mounting System: Design a system that supports multi-functional mounting, allowing for quick replacement and integration of different types of devices or sensors. Sensors such as gyroscope, accelerometer and magnetometer can be used to obtain the current attitude information of the aircraft in real time, the whole flight control system can be designed as a whole by using the modular idea, and the complementary filtering algorithm based on the quaternion and the cascade PID control algorithm can be used to realize the current attitude solution and control output of the aircraft.

Intelligent flight control system: The intelligent flight control system is introduced to achieve more accurate flight control through advanced algorithm and sensor data fusion, and improve stability and maneuverability when loading loads. Fault-state controllers can be designed using a backstepping method that sacrifices yaw control to achieve hovering and simple trajectory tracking.

Efficient Power Management: Optimize power management systems to increase battery capacity and performance to support sustained flight time when mounting large equipment. The use of advanced energy storage technology and power management systems is a key factor.

System integration and collaborative optimization: Implement comprehensive optimization in the entire system, including the collaborative work of the mounting system, flight control system, power system, etc. By considering the individual components together, the overall performance can be improved. Technological Improvements

5.2. Function optimization

Design a universal mount system: Design a system that supports multi-functional mounts, allowing users to swap out different types of devices or sensors. Such a system can increase the flexibility of the drone, making it suitable for a variety of application scenarios. Simulink was used to model the UAV with six degrees of freedom, and the motion attitude characteristics of the UAV during the gliding process, the influence of ailerons on the motion of the UAV, and the movement attitude characteristics of the UAV in the environment of wind interference were analyzed [3].

Modular design: The modular design allows the mount system to be easily integrated and replaced. This allows users to quickly adjust the mount to the needs of the specific task, thus achieving versatility.

Open interface standards: Develop or adopt open interface standards to ensure that a wide range of devices and sensors are compatible with the drone's payload system. This helps drive industry standards and improve device interoperability.

Technology Upgrades and Updates: Regular technology upgrades and updates are carried out to support the integration of new types of devices and sensors. This helps to keep the drone's technology competitive and adapt to the development of emerging technologies. Combined with the sliding mode observer to estimate the fault information matrix, a fault-tolerant controller for a six-rotor actuator based on the switching system was designed by using the handover control algorithm. The simulation analysis shows that the designed fault-tolerant control system has good control performance and robustness, which can effectively reduce the impact of actuator monowing damage on flight [2].

6. Future Research Directions and Development Trends

With the continuous development of science and technology, the research on the mounting problem of hexapopter UAV will become an important direction in the future. In future research, we can try to apply advanced materials and technologies to the mounted equipment to improve the performance and stability of the equipment, and at the same time, we can further study the co-design method of the aircraft and the mounted equipment to achieve better coordination and overall performance. These research results will provide stronger support for the application of hexagophoid UAVs and promote the continuous progress of UAV technology.

7. Conclusion

In this paper, the mounting problem of hexacopter UAV is studied in depth, and the types of mounting equipment and their influence in different application scenarios are discussed. Through the analysis of camera mounting, multi-rotor design, high maneuverability, etc., the development process and characteristics of six-rotor UAV are comprehensively demonstrated.

The paper mentions the methods to solve the problem of weak payload capacity, including structural optimization, multi-functional payload system and intelligent flight control system.

Finally, the importance of improving the adaptability and flexibility of UAVs to meet the needs of different missions was emphasized. Through the discussion of the single problem of sensor and mount type, it provides useful theoretical support for the field of UAV research and points out the direction for future technology development.

Reference

1. Xiong Hang, Zhang Haichao, Qin Ke. Integrated control method of UAV fixed-time position and pose based on jamming observer. Manufacturing automation, 2023, 45(01): 149-155.
2. Wang Siming, Wang Tianyu, Wang Yihu. Fault-tolerant control of six-rotor UAVs based on switching system. 2020, (06): 35-41+48. DOI:10.13645/j.cnki.f.d.20200528.001.
3. Li Wenqiang, Peng Xuefeng, Zheng Zhiqiang. Six-degree-of-freedom simulation of UAV based on Simulink. Journal of System Simulation, 2007 (19): 4604-4606.

Authors Introduction

Ms.Yuping Mei

She is studying at Tianjin University of Science and Technology, majoring in Robotics Engineering. She is very interested in drones.

Ms.Ying Su

She is studying at Tianjin University of Science and Technology, majoring in Robotics Engineering. Passionate about drone machines.

Ms.Xue Yang

She is studying in College of Electronic Information and Automation, Tianjin University of Science and Technology, China. Her research interest is Robotics Arm.

The Application of Hexacopter UAV in The Field of Climbing Evasion

Ying Su*, Yuping Mei, Xue Yang

Tianjin University of Science and Technology, Tianjin, 300222 China

E-Mail: *sy789464@outlook.com

Abstract

With the development of tourism to climb peaks, the safety of high mountain walls plays an increasingly important role. In the past, the terrain of the mountain wall was too steep and too fast, and people could not patrol the mountain wall climbers, resulting in the safety of climbers could not be guaranteed, and the needs of climbers were difficult to solve in time. In order to solve the above problems, the method of real-time monitoring of UAVs and delivery of materials by UAVs was proposed. Hexacopter UAV is an unmanned small aerial vehicle equipment that can carry out vertical lifting, which is expected to improve the safety factor of climbing mountain walls and solve the needs of climbers in a timely manner.

Keywords: Hexacopter UAV，real-time monitoring，delivery of materials，mountion climbing

1. Introduction

In order to ensure the sustainable development of climbing tourism, it is urgent to improve the safety of high-altitude operations. This study mainly analyzes the feasibility and problems of hexacopter UAV in the field of climbing. The stability and load-bearing capacity of the hexacopter UAV determine the bright future of the hexacopter UAV in the field of climbing. Among them, the most significant application of hexacopter UAVs in the field of climbing is to provide magnificent high-altitude shooting conditions. As climbers conquer steep peaks and climb steep rock faces, hexacopter drones with high-definition cameras can record the scene. Aerial footage can not only be used to monitor the climbing process, but also can be used as promotional material to promote climbing projects and attractions, and attract more people to participate in climbing. In the case of very mature UAV monitoring technology, the real-time recording and monitoring of high-altitude conditions by the hexagoter can play a very good role in the climbing tourism industry.

2. Introduction of six-rotor UAV

2.1. The concept of a six-rotor drone

UAV is a small unmanned aerial vehicle with the advantages of flexible flight performance such as vertical lifting, hovering, and loading. The six-rotor UAV has three groups with a total of six motors placed coaxially up and down to provide lift, adjust the attitude by changing the rotor speed, and realize the position control through the change of attitude.

2.2. Features of the six-rotor UAV

The six propellers of the hexagoning UAV are distributed in hexagonal vertex positions, with six propeller speeds as six input forces and six degrees of freedom as outputs. The rotation direction of the six propellers is not the same, and when encountering strong external interference or part of the rotor is disturbed, the output of multiple propellers controlling the rest of the propellers can show good stability [1].

(1) The flight altitude of the hexcaopter

Hexacopter drones typically fly at altitudes between 1,500 and 3,000 meters, but there are some professional drones that can fly to higher altitudes. The technologies that determine the flight altitude of hexacopter UAVs mainly include the following aspects: flight control systems, aviation power systems, communication and navigation systems, and meteorological and environmental sensors. The combined effect of these technologies determines the flight altitude of the drone.

The flight control system is one of the key factors that

determine the flight altitude of the drone, which includes the flight controller, attitude stabilization system, and altitude control system. These systems work together to ensure that the drone can fly stably and maintain a predetermined altitude [2].

The aerodynamic system includes engines, propellers, and power system controls. These systems provide the power needed to fly and influence the drone's climb performance and altitude performance.

Communication and navigation systems also play a vital role in the flight altitude of the drone. Satellite navigation systems (such as GPS) and ground console devices provide precise navigation and positioning information for six-rotor drones in the air, allowing them to accurately control altitude and position.

Meteorological and environmental sensors can detect meteorological factors such as air pressure, temperature, humidity, etc., helping the drone adjust its flight altitude and attitude according to the current environmental conditions.

To a large extent, these technologies have been very mature, especially in the professional field, the flight altitude control technology of UAVs has been very sophisticated, and has provided reliable technical support for various application scenarios.

(2) The volume of the hexacopter

The size of drones can generally be divided into miniature, small, medium, and large. The size of small UAVs is generally between 30 cm and 1 meter, and it is suitable for aerial photography, agricultural plant protection, environmental monitoring and other fields. Among them, medium-sized UAVs are between 1 meter and 5 meters in size, which are suitable for use in search and rescue, logistics and distribution, geological survey and other fields. The carrying capacity of a drone is usually related to its size. Medium-sized UAVs have a large carrying capacity and can already carry relatively heavyweight equipment and materials, such as rescue equipment, survey and monitoring equipment, etc.

3. The application status of six-rotor UAVs

At present, the application field of hexacopter UAV is expanding day by day, covering military, civilian and commercial aspects. In the military field, hexacopter UAVs can be used for reconnaissance, surveillance, target positioning and strike, providing a new way of intelligence acquisition and combat. In the civilian field, UAVs are widely used in aerial photography, geological survey, agricultural plant protection, weather monitoring and forest fire prevention. Drones also play

an important role in commercial areas such as logistics and distribution, building inspection, and power inspection. With the continuous progress of technology, UAVs are gradually used in emergency medical rescue, disaster monitoring and rescue, providing a new means of ensuring the safety of people's lives and property. The development of hexacopter UAVs has not only enriched the application field, but also provided people with more efficient and safer solutions, becoming an indispensable part of modern society. This study introduces the application of hexacopter UAV in the field of climbing, and solves the gap that there is no hexacopter UAV involved in the climbing tourism industry.

4. Analysis of the advantages of hexacopter UAVs in the field of climbing

4.1. High-altitude monitoring function

The high-definition high-magnification zoom camera embedded in the six-rotor drone can clearly see objects one to two kilometers away from the drone. Picture of a high-altitude surveillance drone is shown in Fig. 1. Equipping the six-rotor drone with a stable gimbal gimbal stabilization function, coupled with the smoothness and regularity of the drone's flight, can obtain enough stable images and lenses for surveillance even at full zoom. Regardless of the aircraft's movement, the gimbal maintains its orientation and inclination. The drones responsible for surveillance can navigate autonomously. "Waypoints were previously introduced on the map via a desktop computer or tablet, where waypoints were added by clicking on the map and navigation was stored in the drone's memory. The waypoint can be grouped into one or more points of interest. Navigation can be easily initiated via radio control, so there is no need to use a computer or tablet in the field, and the drone can be in the air in a matter of minutes.

Fig. 1 A high-altitude surveillance drone

4.2. Navigation features

Hexacopter drones play an important navigational role in the field of rock climbing. Rock climbers often need to face complex terrain and changeable environments during climbing, and the navigation

function of drones can provide valuable auxiliary information for climbers.

Equipped with high-precision cameras and navigation equipment, drones can see the climbing route, rock wall structures, and the terrain faced by climbers from the air. This bird's-eye view provides climbers with a new perspective and helps them better understand the complexity and difficulty of the climbing route so that they can develop a targeted climbing strategy. Its real-time monitoring function can help climbers timely obtain changes in the climbing route, such as rock collapse, climate change, etc., and provide timely warnings and risk warnings for climbers. This real-time monitoring is critical for the safety of climbers, allowing them to better respond to unexpected situations and reduce the likelihood of accidents [3].

The navigation role of hexacopter drones in the field of rock climbing can provide climbers with a full range of information support, help them better cope with challenges, and improve the safety and success rate of climbing.

4.3. Aerial teleportation function

For the preparation of rock climbing competitions and climbing events, drones can be used as a fast transmission tool to transmit competition equipment and judging equipment to the required locations, improving the work efficiency of organizers and participants. Climbers often need to carry a variety of equipment, food and water during the climb, and drones can be used as a high-altitude teleportation tool to help climbers deliver these supplies to the designated location. For climbers, there are often difficult parts of the climbing route to pass, such as steep cliff walls, inaccessible caves, etc. Drones can carry ropes, gear, and food to reach places that climbers can't, delivering them safely to where they are, providing them with the support they need. Climbers can also be provided with emergency supplies during the climb, such as delivering much-needed clothing, food, and medicine to climbers in case of inclement weather to help them get through the storm. Photos of actual transportation of hexarotor drones is in Fig. 2.

Fig. 2 Actual transportation of hexarotor drones

In general, the high-altitude teleportation function of drones in the field of rock climbing provides convenient material support for climbers, provides more safety and convenience during climbing, and adds new possibilities and innovations to rock climbing activities [4].

5. Development prospects and challenges of hexacopter UAVs in the field of climbing

The application of hexacopter drones in the field of climbing shows great potential, but at the same time, it is also necessary to pay attention to issues such as privacy protection and environmental protection. When using drones for high-altitude shooting, it is necessary to comply with local aviation regulations to ensure that they do not interfere with the flight of other aircraft. In addition, when flying, it is necessary to strictly comply with local environmental protection laws and regulations to avoid affecting the local ecological environment.

6. Conclusion

The application of hexacopter drones in the field of climbing provides more safety and technical support for climbers, expands the possibilities of climbing activities, and brings new innovations to climbing sports. With the continuous progress of technology and the expansion of applications, it is believed that the role of UAVs in the field of climbing will become more prominent, bringing more convenience and safety to climbing sports.

Acknowledgement

and support in my studies and life. Finally, I would like to thank all the people who helped me in the literature research and research to complete this paper. Thank you from the bottom of my heart!

[1] Xiaohan Liao, Chenchen Xu, Huping Ye, et al. Unmanned Aerial Vehicle (UAV) Application Development Key Infrastructureand Low-altitude Public Navigation Network Planning (in Chinese). *Bulletin of the Chinese Academy of Sciences.* [ChinaXiv:202303.10019]

[2] Xinyao Hui. Design and implementation of composite UAV flight control system based on STM32. *North China University of Technology*, 2022. DOI:10.26926/d.cnki.gbfgu.2022.000674

[3] LIU Huan, HAO Qianqian, LIU Wenhui, et al. Development trend of UAV intelligent measurement and control technology. China Science and Technology Information, 2023, (19):67-71.

[4] Jinsheng Guan. Unmanned and emergency logistics support under smart logistics. China Storage and Transportation, 2021, (05): 137-138.DOI:10.16301/j.cnki.cn12-1204/f.2021.05.052

Authors Introduction

Ms. Yuping Mei

She is studying at Tianjin University of Science and Technology, majoring in Robotics Engineering. She is very interested in drones.

Ms. Xue Yang

She is studying in College of Electronic Information and Automation, Tianjin University of Science and Technology,China. Her research interest is Robotics Arm.

Deep Learning and Embedded Based Operational Safety System for Special Vehicles

Haoran Gong*, Yumei Huang, Jiahao Xie
Tianjin University of Science and Technology, Tianjin, China

*E-mail: *1465157817@qq.com*

Abstract

Aiming at the characteristics of special operation vehicles, such as complex operating environment, heavy vehicle weight, long braking distance, and many visual dead angles for drivers, this project designs and implements a special vehicle operation safety system based on deep learning and embedded system. Firstly, the deep learning model is applied to detect the humanoid target appearing around the vehicle, get the detection frame of the humanoid target, and estimate the distance from the target to the camera; using the advantage of the embedded system which is relatively small can be easily deployed on the heavy vehicle. According to different distances, the system will issue relevant voice prompts to remind the driver to take timely measures such as avoiding or braking.

Keywords：special operation vehicles，Deep Learning，Embedded system，Complex working.

1. Introduction

Blind spots in the field of vision of specialised vehicles during operations are a serious safety challenge that can lead to a wide range of accidents, as the vehicles tend to be large-bodied. These drivers may not be able to directly observe the blind spots in their field of vision, increasing the risk of hazards to workers, other vehicles and obstacles.

For example, in specialised vehicles such as excavators or loader trucks (Fig. 1 below). drivers may not be able to directly see pedestrians or workers around them due to the special construction of the equipment itself, increasing the risk of run-over or collision accidents. In addition, when turning, working or reversing, the blind spot in the field of vision may also result in the driver not being able to perceive other vehicles behind the side, which may lead to a collision accident. Drivers are also less likely to be aware of fixed obstacles such as rocks and pillars, which can lead to collisions [1].

Fig. 1 Common Heavy Vehicles

In order to counteract these risks, specialised vehicles can take appropriate measures. To this end, we have designed a system that helps drivers to better avoid and detect obstacles in a timely manner, for example by installing auxiliary cameras and sensors. The use of advanced assistance systems, the strengthening of the driver's field of vision and the incorporation of a hazard warning system reduce the probability of safety accidents caused by blind spots or obstacles in the field of vision of specialised vehicles during operations [2].

The rest of this paper is organised as follows. Part II describes the hardware usage scheme. The third part describes the innovative points of the design as well as the leading advantages. The fourth part explains the application scenarios using the design. Part V summarises the main points of the paper.

2. Hardware usage programme

Haoran Gong, Yumei Huang, Jiahao Xie

2.1 Jetson nano

Hardware selection, we take into account the cost-effectiveness and durability of the two major aspects of the main development board selected NVIDIA series of development boards - Jetson nano development board.

Here to introduce this development board type: Jetson nano is NVIDIA's development of high-performance GPU equipped with a microcontroller, is to achieve deep learning and parallel computing necessary microcontroller. It is broken down into four different types of development boards. They are: Jetson Nano, Jetson TX, Jetson Xavier NX, and Jetson AGX Xavier, where we choose Jetson Nano. Because of its low energy consumption and low cost it takes the lead among the other three microcontrollers. And its arithmetic power has been increased by tens to hundreds of times, even more than a small part of the current computer CPU, and NVIDIA also equipped with a full set of SDK for it. the following Fig. 2 is the picture of Jetson Nano development board, Fig. 3 is the development board core board [3].

Fig. 2 Development Board   Fig. 3 Core Board
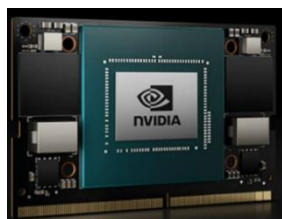
The jetson nano performance metrics (representative) are shown in Table 1.

Table 1 jetson nano performance metrics

arithmetic power	40TOPS
GPU	1024-core NVIDIA Ampere architecture GPU with 32 Tensor cores
CPU	6-core Arm® Cortex®-A78AEv8.2 64-bit CPU 1.5MB L2+4MB L3
RAM	8GB+68GB
Camera Interface	2 CSI interfaces
output	5W-10W

2.2 surveillance camera

Cameras are generally selected industrial cameras, such as Hikvision cameras or select Dahua cameras. The advantages of these two cameras are high-definition, clear images, large shooting angle, can be rotated and in the dim light conditions, has a strong night vision effect. They provide a more accurate video feed to the central processor. Fig. 4 below shows the Hikvision camera shooting debugging screen.
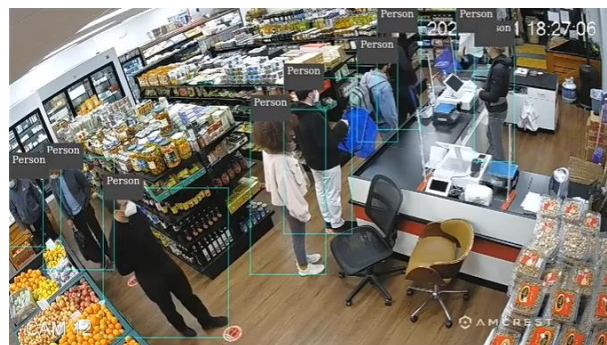
Fig. 4 Debug screen

3. Innovative design and advantages

3.1. Deep Learning

The initial idea of our design is to first recognise the human figure, and then to warn the driver by triggering an alarm when the human target walks into a relatively dangerous area of the vehicle's operation.

Due to the relative complexity of the special vehicle operating environment, traditional computer vision and image processing algorithms are relatively fuzzy and difficult to deal with to identify the humanoid target from the environment, and often triggers the alarm system incorrectly (e.g., a roadblock or reflective clothing), therefore, deep learning models are used in this project to complete the identification of the humanoid target, which is shown in Fig. 5.

Deep learning is a research direction in the field of machine learning. Deep learning uses deep neural networks for modelling and uses large amounts of data to optimise the parameters in the model to achieve high accuracy and robustness.

We have designed a deep learning model that can recognise targets within 50 to 100 metres in terms of distance, up to 9000 types of objects (yolo-9000) in terms of type, and hundreds or even thousands of targets in a single frame at the same time.

Fig. 5 Initial identification of deep learning models

3.2. *Technical difficulties*

There are many types of special vehicle operating scenarios such as harbours, construction sites, mining sites, as in Fig. 6. intensive construction of people, many obstacles and close vehicle traffic. Sand, gravel, goods, and materials are piled up. These serious obstacles to our acquisition of the surrounding image information.

Fig. 6 List of construction scenes at mines and construction sites

In addition, because it is assembled on the carrier, it is required to have low energy consumption; and the space inside the vehicle is narrow, which can only accommodate smaller equipment for normal operation, so we can only compress its size.

Due to the ever-changing environments in complex places, we cannot predict what will happen next. So we need as little false recognition or frame skipping as possible. Too many misjudgements can cause

unnecessary distractions for the driver. In addition, if the vehicle is travelling fast in a complex scene, there may be a requirement for the real-time performance of the model.

Generally in complex scenes there will be a large number of staff and supervisors. We want to recognise and estimate their distance to their own vehicle as much as possible.

3. 3. *Problem-solving responses.*

For the human-type and object-type densification problems, we have prepared datasets for training. A certain amount of data is collected to form the dataset for training and the dataset for testing. The amount of data is usually 2000 images, which contain at least 5000 humanoid targets. The dataset is labelled using manual annotation (Fig. 7). At the same time, the surrounding environment is constructed (Fig. 8).

Fig. 7 Dataset labelling
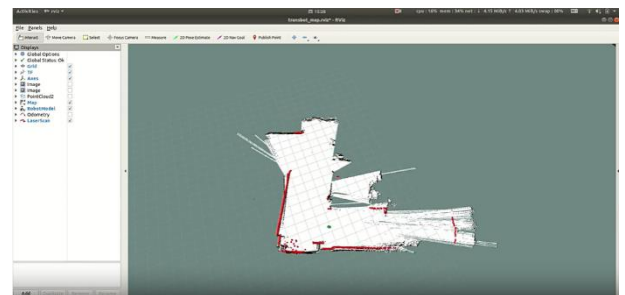
Fig. 8 Environmental scanning

The mainstream deep learning yolov5 and the deep neural network Pedestrian net, which is specialised for humanoid and vehicle detection, are deployed on jetson nano [4].

The model uses ResNet34 as the basic underlying layer and combines advanced techniques such as feature fusion, local zoom, and exact query. The original model is 84.62MB, after pruning 7.85MB, then use Tensor RT for assisted optimisation, generate the physics engine in

semi-precision form, and then deploy the physics engine into deep learning, then after a series of operations can be done on a development board to deal with four to eight video signal sources.

Finally, the target detection metadata is sent in the form of a request to the thread responsible for estimating the distance and voice announcement, and multiple calculations are performed to compare the estimated relative distance with a pre-set threshold to minimise the impact of erroneous judgements. Finally, an alarm is triggered to warn the driver to be careful of the surroundings.

4. Job Scenario Application.

4.1. *yard operation*

Heavy vehicles loading and unloading in port yards face serious visual challenges, with a large number of visual dead zones and blind spots. At the same time, the construction site is a noisy environment with a large number of people, which makes it difficult for the driver of the operating vehicle to detect people in the vicinity of the vehicle in a timely manner in complex situations, thus spawning many safety hazards.

We use a four-channel video signal source to assist the operation. The four video sources are processed simultaneously on a single jetson nano, which can achieve 10 frames per second for each source. The communication is done through a pos switch and finally a voice alarm system.

4. 2. *Shield Tunneling*

Shield Tunnel Battery Vehicle is an electric vehicle specially designed for shield tunneling project, and its main task is to transport materials and transfer personnel at the tunnel construction site.

However, due to the poor line of sight in the tunnel and the heavy load of the vehicle, there is a great safety hazard in the construction. The main reasons include: the existence of a blind field of vision, the existence of a no-signal section in the tunnel, the absence of reminders during operation, the possible failure of the brakes, and improper operation.

We made a solution to these situations in the follow-up processing, first of all, the tunnel front and rear of about fifty metres distance sweep, the construction workers to collect data, and then through our are reinforcing the night vision effect algorithm will be the situation in the

tunnel for light reinforcement, passed to the pos switch, and finally complete the exploration of the environment around the shield tunnel machine. Fig. 9 below shows a comparison of the colour of the enhanced picture in dark conditions for the night vision effect.
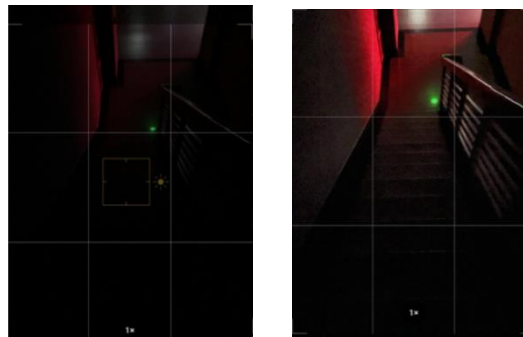
Fig. 9 Night Vision Enhanced Picture Colour Before and After Comparison

Fig. 10 below shows the final physical model drawing and the final physical picture.

Fig. 10 Physical models and final objects

5. Synthesis

This paper describes the various safety hazards of specialised vehicles in complex situations and how to deal with them. A machine has been designed to assist special vehicle drivers to avoid accidents. It can be adapted to various complex sites including yards, harbours and mines. By coping with various conditions

with ease. It is believed that in the future, it can help more drivers and escort the safe operation of special vehicles.

6. References

[1]. Zhao, J.F., Zhao, J.L., Li, Q., et al, Research on machine vision-based fast-passing technology for special vehicles at intersections, *Design and Application*, Vol. 31, No. 2023, pp. 240-246.

[2]. Liu X, Chacoco, An environmental target recognition method for special vehicle operations at airports, *Computers and Modernisation*, Vol. 2023, No. 8, pp. 18-24.

[3]. Fangxin Xu, Rong Fan, Xiaolu Ma, Improved YOLOv7 Algorithm for Congested Pedestrian Detection, *Computer Engineering*, November 2023, pp. 1-10.

[4]. Lu Beiyao, Zhang Tiecheng, Song Yang, et al, Control system of a hospital ordering robot based on JETSON NANO, *Intelligent Applications*, Vol. 14, 2021, pp. 39-41.

Authors Introduction

Mr. Haoran Gong

He is currently studying in the School of Electronic Information and Automation of Tianjin University of Science and Technology, and is proficient in embedded architecture with strong single-player working ability.

Ms. Yumei Huang

She studied at Tianjin University of Science and Technology in 2021. She is now the first major in her faculty and is expected to receive her Bachelor's degree in 2025.

Mr. Jiahao Xie

He studied Oceanographic Sciences at Tianjin University of Science and Technology in 2022 and is currently a bachelor student of science majoring in Marine Technology. He expects to receive a bachelor's degree in 2026.

"Green Fruit" - Intelligent Traceable Agricultural Product Production and Marketing Platform Based on Blockchain Technology

Yumei Huang*, Jiahao Xie, Haoran Gong, Ziyue Xiao
Tianjin University of Science and Technology, Tianjin, China

E-mail: *3033665674@qq.com

Abstract

The production process of agricultural products has been greatly improved and optimised through the intelligent agricultural greenhouse system. Using intelligent irrigation, measuring CO_2, IoT and other technologies, it creates a greenhouse intelligent control system that is simple to operate, highly automated and intelligent. In addition, using the double combination of front-end architecture and applets, it can record all the growth process, perfectly integrating the intelligent mode + big data application. Create a tailor-made ID for agricultural products. Under our design, we break the traditional automation mode, which is more conducive to understanding the whole growth process of agricultural products.

Keywords: Internet of Things; automation; big data application; intelligent agriculture.

1. Introduction

The aim of the paper is to discuss the many advantages brought by the traceable smart greenhouse model in terms of agricultural production, food safety and market competitiveness, as well as the initial system modelling and related technologies established [1].

From the perspective of food safety and security By integrating product traceability technologies, it is possible to trace the entire process of agricultural products from planting and growing to harvesting, processing and marketing. This helps to monitor potential sources of contamination, diseases or other risk factors, and to identify and deal with problems in a timely manner, thereby improving food safety. Automated control allows real-time monitoring of environmental parameters in the greenhouse, such as temperature, humidity and light, which can precisely control the growing conditions of crops. Through meticulous data recording, farmers can optimise planting strategies to improve the quality and yield of their produce. The increased resource efficiency enables smart greenhouse technology to enable precise irrigation and fertiliser application, as well as efficient energy use. This helps to reduce water wastage, fertiliser use and energy expenditure, and improve the sustainability of agricultural production. The Internet of Things (IoT) in this plays an extremely crucial role and offers many opportunities for innovation and efficiency gains in agricultural production. It can deliver real-time metrics of various parameters in the model to the network and feedback to the user's mobile phone. The user can observe the crop growth at any time. So, IoT can monitor various environmental parameters in the greenhouse in real time, such as temperature, humidity, light, soil moisture, and so on. This real-time monitoring enables farmers to understand the growth of their crops more accurately and quickly adjust the conditions of the greenhouse to optimise the production environment. When problems arise, such as disease, pest infestation or other production challenges, the traceability system can quickly pinpoint the source of the problem. This helps farmers take quick action to minimise losses and maintain continuity of production [2].

Each crop can be identified with its own identity card. Smart greenhouses generate large amounts of data through sensors and monitoring systems, which can be used for data storage. Farmers can make more informed decisions based on actual environmental and production data, and buyers can learn about their growth.

The rest of the paper is organised as follows. Section 2 presents the structure and principles of the system model. Part III presents the feasibility study of the system model. Part IV gives the physical pictures of the constructed model and the verification process to finalise the correctness of the designed. Part V summarises the main points of the paper [3].

2. Introduction to the system model

Intelligent greenhouse control system is a scientific system for intelligent management of greenhouses.

Yumei Huang, Jiahao Xie, Haoran Gong, Ziyue Xiao

Environmental data is collected through the corresponding sensor equipment. Then, through various commands set by the system, the management work is completed automatically.

At the same time, the data collected by the sensors will be analysed and the results will be fed back to the farmers. Under the guidance of agricultural experts, the yield is increased and the quality is improved. In the whole system, QCA9531 gateway module is used as the core component, and the core programme is burned into it. At the same time, all kinds of sensors, rolling shutters, fans, water valve controllers, cameras and other hardware devices will be accessed. When the power is to be turned on, it can be run and controlled automatically according to the programme. The use of sensors to obtain environmental data, when detected more than the threshold value of the fruits and vegetables, will start the corresponding actuator.

2.1 CO2 Sensor

Carbon dioxide (CO_2) sensor is composed of LED-PR measuring unit, LED driving circuit, LDO power supply, photoelectric signal processing circuit, temperature measuring circuit and microprocessor, which is shown in Fig. 1 below.
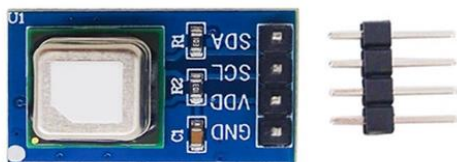
Fig. 1. Sample sensors

Firstly, a suitable location should be chosen to install the CO_2 sensor. The location should be chosen to represent the average CO_2 concentration of the plant growing area in the greenhouse. The sensor is usually installed in a hanging position to ensure that it can accurately measure atmospheric CO_2. before use, the CO_2 sensor needs to be calibrated to ensure the accuracy of its measurement. Calibration can be carried out by using a standard gas with a known CO_2 concentration, according to the model number of the sensor and the guidelines provided by the manufacturer. Then, connect to the monitoring system. CO_2 sensors are usually connected to the monitoring or control system of a smart greenhouse. This can be achieved through a wired or wireless connection. The sensor sends real-time measured CO_2 concentration data to the monitoring system, enabling the farmer to remotely monitor and manage CO_2 levels in the greenhouse. The farmer can set thresholds based on plant growth

requirements and optimal CO_2 concentration ranges. Once the measured CO_2 concentration exceeds or falls below the set threshold, the system triggers appropriate control measures, such as adjusting the ventilation system or CO_2 supply unit.

The CO2 sensor monitors the CO2 concentration in the greenhouse in real time and transmits the data to the monitoring system. These data are usually recorded and archived for future analysis and reference, as shown in Fig. 2 below.
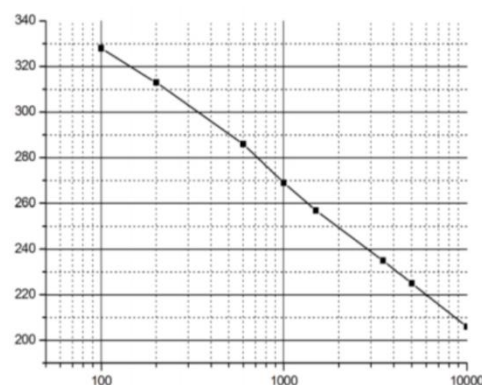
Fig. 2 Sensor characteristic curve

The Y-coordinate represents the output signal, with a maximum value close to 330mV; the X-coordinate represents the concentration of CO2, with a minimum value close to 210ppm.

2.2 temperature detector

Temperature detection is very important for greenhouse intelligent control system. As shown in Fig. 3, the temperature sensor is used to collect the surrounding environmental factors, read the temperature data, and through the signal processing function, the peripheral devices, as shown in Fig 4, respond accordingly and transmit the temperature data wirelessly to the cloud platform. When the temperature is higher than the preset temperature e, the MCU drives a relay to control the fan rotation to reduce the temperature. When the temperature is lower than the preset temperature, the MCU drives the relay to control the collector to heat up [14].

Fig. 3 Temperature sensor DHT11
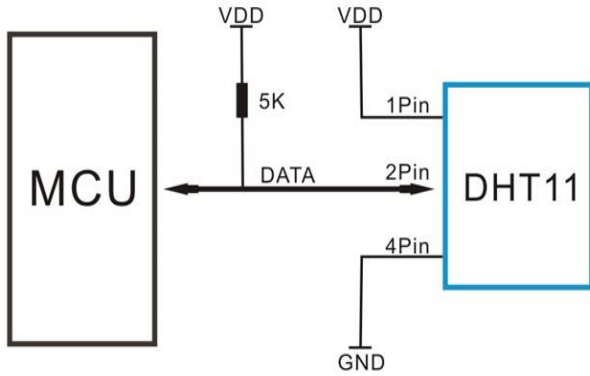
Fig. 4 Typical Circuit Applications

2.3 Soil moisture sensors and automatic irrigation systems

The soil moisture sensor is chosen to be installed per $1m^2$ of soil by selecting the LM393 soil moisture sensor, as shown in Fig. 5, based on the needs of the crop and the layout of the greenhouse. Typically, the sensor is buried near the root system of the plant to ensure that it can accurately measure the moisture in the soil. Soil moisture sensors usually require a power supply (Eq. (1)). The corresponding values for Equation I are shown in Table 1. Sensors usually use low-power electronic components that can be powered by batteries or other suitable power sources. Then connected to the monitoring system: The soil moisture sensor is connected to the monitoring system of the shed. Real-time soil moisture values are fed back to the platform to provide numerical feedback for subsequent soil rehydration. The premise is to set threshold values for soil moisture based on crop needs and optimal growing conditions. These thresholds will be used in the irrigation system [4].

Fig. 5 LM393 Soil moisture sensors

$$Q(t) = \int_0^t I(\tau)d_\tau + Q_0$$

(Battery power at time t)   (1)

Table 1 symbols denote meaning

Q(t)	Battery power at time t
$I(\tau)$	Current as a function of time
τ	integral variable
Q_0	Initial battery charge

The design of the automatic irrigation system is a multi-level, multi-component project as shown in Fig. 6 need to receive the sent data from the soil moisture sensor to determine whether irrigation is required. By setting a default upper and lower threshold of normal humidity ($\pm 3\%$) in the software. When the soil moisture is lower than the set lower threshold, the pump is activated to deliver water to the underground pipeline and spray it out to drench and moisten the soil to achieve the effect of irrigation. At the same time, the soil moisture sensor continuously checks whether the humidity has risen above the middle of the upper and lower thresholds, and if it meets the requirements, it stops the delivery; if it does not meet the requirements, it continues to cycle the process until it reaches the normal humidity range (Fig. 7).

```
if(sensor=="shadowing"){
  digitalWrite(IN1,HIGH);
  digitalWrite(IN2,LOW);
  analogWrite(ENA,100);
  delay(1000);
  digitalWrite(IN1,LOW);
  digitalWrite(IN2,LOW);
  analogWrite(ENA,255);
  delay(100);
  return true;
  }
if(sensor=="pump"){
  digitalWrite(PUMP,LOW);
  delay(500);
  return true;
}
  return false;
}
```

Fig. 6 Embedded part of the code

Yumei Huang, Jiahao Xie, Haoran Gong, Ziyue Xiao

```
/****************get DHT11 ****************/
int h=(int)dht.readHumidity();
delay(150);
/**********************/
digitalWrite(MH_power, HIGH);
delay(10);
float s = analogRead(MH_A0);
float s1 = map(s0, 0, 1023, 255, 0);
delay(150);
/************************/
uint16_t l=BH1750();
delay(150);
uint16_t DHT11=400;
if(SGP30(t,h)){
  co2=SGP30(t,h);
}
```

Fig. 7 Get DHT11 code

3. System feasibility study

The application of IoT technology in smart agriculture has enabled real-time monitoring and control of parameters in the greenhouses, providing a high degree of environmental control and visibility. This helps farmers to understand the situation in the greenhouse in real time and take necessary measures to meet the needs of the crop and improve yield and quality.

The traceability system provides a reliable guarantee of the quality and safety of agricultural products through QR codes and database technology [5]. Consumers can easily access detailed information about the products, which builds trust and improves their market competitiveness.

The high degree of automation of the greenhouse system makes it very easy and convenient for farmers to operate. Monitoring of environmental parameters, water and nutrient supply in the greenhouse can be automated through automatic control systems. This reduces the dependence on a lot of manpower and farmers can focus more on monitoring and managing the system rather than having to physically work in the greenhouses for long hours.

By accurately monitoring and controlling environmental conditions, the system is expected to significantly improve the yield and quality of produce. Because the growing environment in the greenhouse can be finely tuned, crops can be optimised for optimal light, temperature and humidity conditions at different stages of growth, leading to increased yields and improved product quality.

4. Physical pictures and verification process

The model is now completed in terms of technology development, with hardware models, WeChat small programmes, traceability systems [6], e-commerce website development. website, and in the future it will enter the data interaction stage and improve the audience

universality test [7], which are shown in Fig. 8, Fig. 9 and Fig. 10.
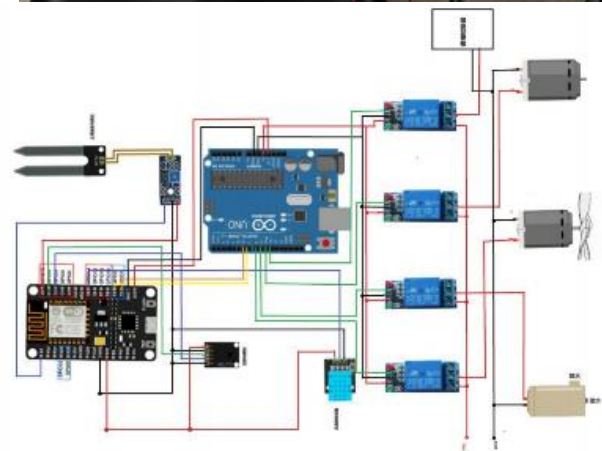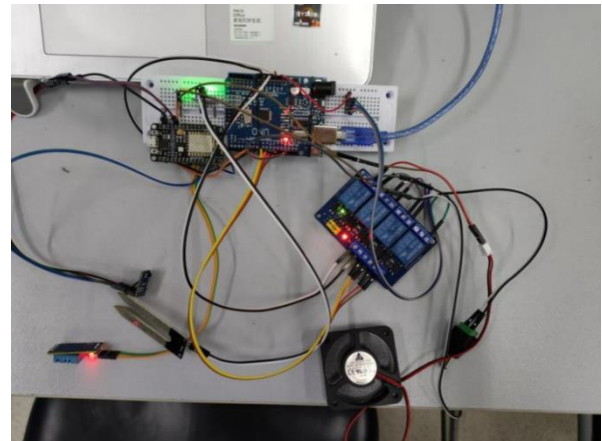
Fig. 8 on-site demonstration
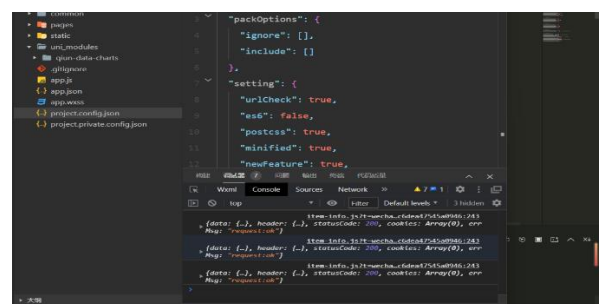
Fig. 9 Cloud Platform Home
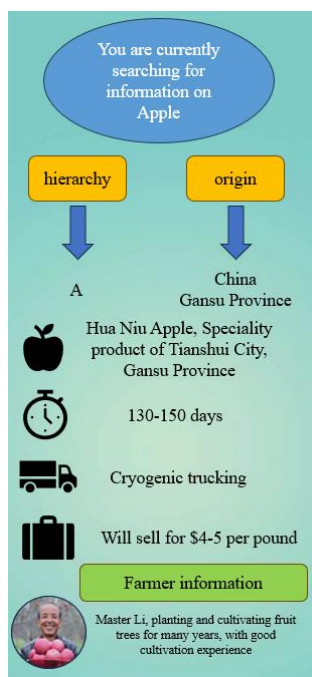
Fig. 10 Crop "ID"

5. Synthesis

The paper mainly describes the growing of crops in smart greenhouses under the fully automated mode. "Traceable Smart Greenhouse" is a modern agricultural production system that uses advanced technology and Internet of Things (IoT) to improve the efficiency of agricultural production and realise precise agricultural management. Including: intelligent sensor technology, Internet of Things, automation control, precise fertiliser application and water management. It improves the efficiency, quality and sustainability of agricultural production, while ensuring product quality through a traceability system to provide consumers with safer, traceable agricultural products [8].

6. References

[1]. He, Y. K., Song, Y. F., Tan, P., et al. "Design of Multifunctional Potted Plant Farming Device Based on STM32F4", *Computer Applications*, 2021, 4: pp. 88-90.

[2]. Miao, Junchen, Wang, Rui Hao, You Da, et al. "An intelligent agricultural greenhouse greenhouse detection system based on the Internet of Things", *Internet of Things Technology*, 2023, 1: pp. 16-18.

[3]. Gao Guanlei, Li Changpu, Fu Shuashuai, et al. "Design of intelligent greenhouse greenhouse control system based on Internet of Things", *Shandong Industrial Technology*, 2023, 2: pp. 71-77.

[4]. Yang, Hailong, Kou, Jian, Wen, Xiaodong, et al., "Design of Temperature and Humidity Detection System for Intelligent Buildings Based on ESP8266," *Journal of Hebei University of Architecture and Engineering*, 2023, 9: pp. 177-181.

[5]. Yang, Mina, "The design and development of "help farmers" sales system based on WeChat applet", *Software Development and Application*, 2021, 4: pp. 34-35.

[6]. Jinwen Chen, "Research on Key Technology of High Privacy Agricultural Traceability Based on Blockchain," *Master's thesis*, 2023, 6

[7]. Jian Nian, "Research on the Application of Traceability of Agricultural Products in Smart Agriculture," Smart Agriculture: monographs and reviews, 2021, no. 4: pp. 16-18.

[8]. Jingxuan Yang, "Current situation and future development trend of agricultural greenhouses in China," *Seed Science and Technology*, 2023, no. 14: pp. 142-144.

Authors Introduction

Ms. Yumei Huang

She studied at Tianjin University of Science and Technology in 2021. She is now the first major in her faculty and is expected to receive her Bachelor's degree in 2025.

Mr. Jiahao Xie

He studied Oceanographic Sciences at Tianjin University of Science and Technology in 2022 and is currently a bachelor student of science majoring in Marine Technology. He expects to receive a bachelor's degree in 2026.

Mr. Haoran Gong

He is currently studying in the School of Electronic Information and Automation of Tianjin University of Science and Technology, and is proficient in embedded architecture with strong single-player working ability.

Mr..Ziyue Xiao

Hold a positive and serious attitude towards work, have a strong sense of responsibility, be sincere, meticulous, optimistic, and stable, have a good team spirit, can quickly adapt to the work environment, and can continuously learn and improve oneself in practical work, and do their job well.

A Digital Twin Design Based on Robot Workstation

Huahao Li

College of Electronic Information and Automation, Tianjin University of Science and Technology, 300222，China

E-mail: 1983618432@qq.com

www.tust.edu.cn

Abstract

With the rapid development of intelligent manufacturing industry and the proposal of Made in China 2025 and Industry 4.0, Digital Twins DT has rapidly become a craze. Through the interaction between the physical object and the virtual model, the mapping is completed in the virtual simulation space, so as to reflect the actual operation process of the whole life cycle of the corresponding equipment workstation. The main structural content of this article is the research background and significance of digital twin technology, the current research status at home and abroad, the establishment of Solidworks physical models, the construction of PDPS (Process Designer&Process Simulate) workstations, and virtual simulation. Completed the construction of the robot seven color board assembly workstation, established a model of the assembly workstation, and combined this model with PDPS to simulate the workstation. Finally, the principle of connection communication and the achievement diagram were elaborated.

Keywords: Digital Twin, Robotic Assembly Workstation, Virtual simulation, PDPS

1. Introduction

"Made in China 2025" and "Industry 4.0" strategic plan proposed, so that intelligent manufacturing and artificial intelligence and a series of intelligent technologies and related development strategies continue to be proposed and implemented, therefore, digital twin technology has become a basic principle of intelligent manufacturing direction, and has received the general attention of all parties. Digital twin technology is the basis of intelligent manufacturing. It is gradually penetrating into manufacturing, medical care, urban management and other fields. The practical application of digital twin technology in control, perception, big data, artificial intelligence, modeling and other fields has also seen a centralized outbreak and growth [1].

As an advanced stage of digital development, digital twin will promote the integration of intelligent, information and digital development processes of integrated enterprises [2]. The development of digital twins has reached an unprecedented height, and now digital twins are not only an inevitable trend of national industrial development in the future, but also an important embodiment of the level of scientific and technological development.

The specific research content of this design:

First of all, the overall design and planning of the digital twin of the virtual simulation of the robot assembly workstation should be carried out. The following five aspects need to be studied and analyzed respectively:

(1) Understand and learn the overall structure and specific layout of the robot assembly workstation. Be fully prepared to build a layout in PDPS.

(2) Make and build models of ABB robot 1410 assembling workstation and other robot bases and other parts in SolidWorks 2019.

(3) Import all the models built in SolidWorks into the PDPS simulation software. Then build the workstation and write the robot motion instructions and paths. In the PDPS simulation software according to the actual layout of the assembly workstation.

(4) Carry out assembly simulation of all processes and route debugging in PDPS software to achieve the effect of workstation assembly and operation simulation.

(5) Communication steps and principles of virtual and real equipment of PDPS and robot assembly workstation.

2. General idea of system design

The system is mainly composed of off-line simulation module, twin module and auxiliary function module.

The overall process design flow chart of the robot assembly workstation system structure is shown in Fig. 1.
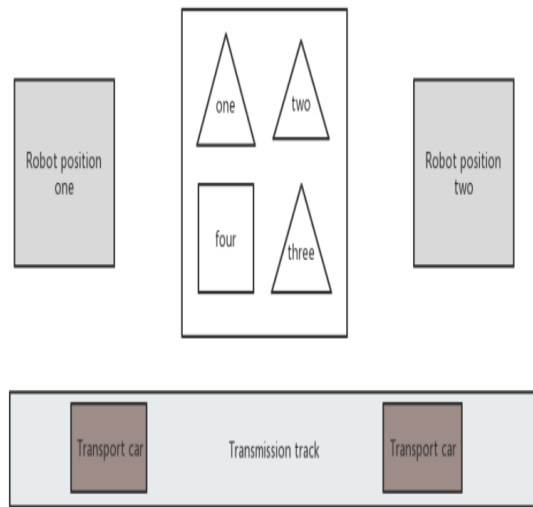
Fig.1.Process layout design flow chart of the system

The process mainly includes: the ABB1410 model industrial robot is placed in the No. 1 station and No. 2 station respectively, so that the robot carries out the handling and assembly work on the three seven-color panels and a cover plate on the storage platform in turn, and then transported to the No. 2 station robot operates the cover screw. After the screw is fixed, a set of processes is completed, and finally the object is transported to the finished product area by the transfer track.

3. Robot assembly workstation

The industrial robot used in the robot assembly workstation is an industrial robot specially designed for assembly. Compared with other industrial robots, the assembly robot has the advantages of strong flexibility, high precision, small working range, and can be used in combination with other operating systems. The main driving mode of the system is electric driving mode. The research scene of this paper should be carried out in the assembly of seven-color plates, and the advantages of industrial application are to improve production efficiency and reduce human capital.

For the model built above, the specific parameters are sorted out as follows. The model selection and size table of robot assembly workstation are shown in Table 1 below:

Table 1. Model selection and size table (Unit: mm)

No.	Name	Length	Width	Height
one	path	2400	575	800
two	Transport car	600	400	80
three	Bottom plate	200	200	30
four	Cover plate	200	200	20
five	Color palette one	150	150	20
six	Seven color palette number two	150	$75\sqrt{2}$	20
seven	Color palette number three	150	$75\sqrt{2}$	20
eight	Robot chassis	Radius 350	0	500
nine	Storage table	130	60	80

4. PDPS workstation construction and program design and simulation

The construction of the assembly workstation is mainly based on the layout of the robot training base to simulate the construction of the layout. Through the actual working principle and path of the robot, the coordinate position of the robot and the location of the material workpiece are first determined. Secondly, the trajectory and mode of movement of the robot are understood, and then the layout is carried out through the built model. As shown in the figure, the basic layout is composed of two ABB1410 robots, corresponding to station No. 1 and station No. 2 of the workstation. Station No. 1 is responsible for the assembly of 7 color board No. 1, 7 Color Board No. 2, 7 color Board No. 3 and the cover plate. It is transported to station No. 2 by the conveyor rail, and is responsible for fixing the screws on the four corners of the cover plate at station No. 2. Fig. 2 shows the preliminary model construction:
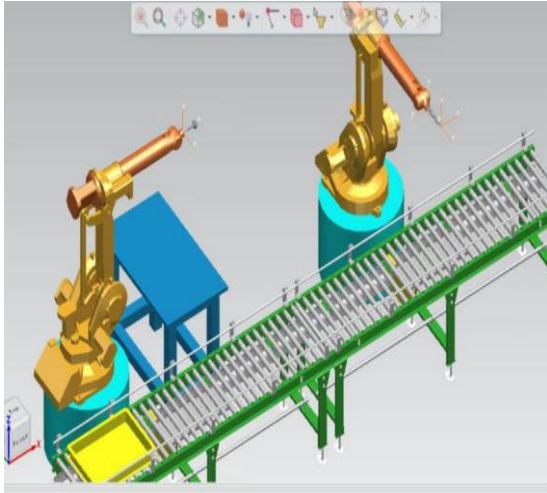
Fig. 2. Basic layout of robot assembly workstation

The robot assembly workstation is mainly composed of two stations, namely P10 station and P20 station. The robot at P10 station is mainly responsible for the handling and assembly of the 7-color board and the 7-color board cover plate, while the robot at P20 station is mainly responsible for the final assembly of the 7-color board by fixing the bottom plate and the 7-color board cover plate by screws at four corners.

The overall planning of the coordinate system of the robot assembly workstation, through which the coordinates of the robot and the robot base are connected, can be roughly divided into the establishment of the coordinate system of the seven-color plate and the cover plate, the coordinate system of the transport car and the chassis, and the motion coordinate system of the assembly robot OP10 station and the robot OP20 station. As shown in Fig. 3 below.
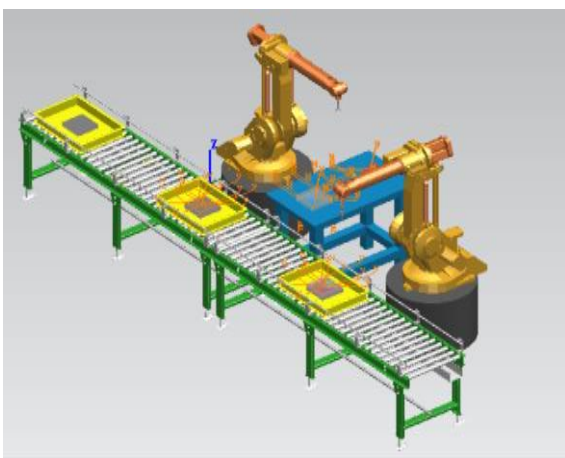
Fig.3. Coordinate diagram of robot and workstation

5. TIA+PLCSIM+NETTOPLCSIM+KEPServerEX virtual debugging

Open our robot assembly workstation in production line mode and click Add Zone, where the path and location can be observed [3]. Click on the motion simulation, and the robot station OP10 starts to carry and assemble the red triangle plate, yellow triangle plate, purple triangle plate and transparent cover plate successively. Meanwhile, the robot station OP20 also synchronously fixed the screws of the four stations of the cover plate. When the assembly is completed, sensor B receives the signal. The signal is received by NetToPLCsim. And the channel 4 signal shared by KEPServerEX, the PDPS workstation receives the virtual map of the OPC signal 1, and the conveyor begins to move. When the sensor A receives the position signal, which is the same as the signal B, the signal becomes 0, the conveyor belt stops moving, and the robot station OP10 and OP20 start handling and assembly. When the stop button is pressed, the signal is disconnected and the robot assembly workstation stops moving. The workstation of the robot is shown in Fig. 4.
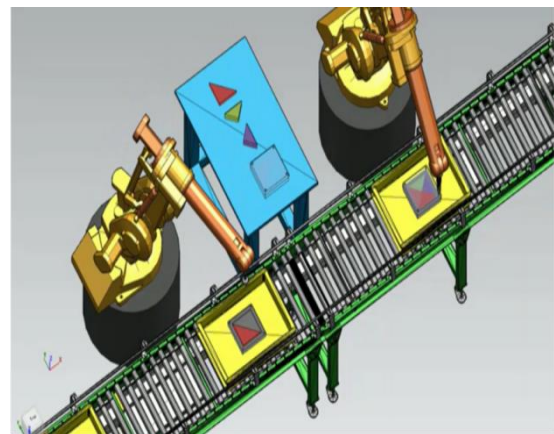
Fig.4. Robot motion diagram

6. Conclusion

In this paper, a simple control system can be realized through the construction and simulation of the basic model of PDPS software based on the working mode and basic method of robot assembly workstation. Do model construction, basic operation methods, path debugging. The virtual reality technology can realize human-computer interaction. The digital twin based on the robot assembly workstation mainly takes the assembly process of the seven-color panel as an example to simulate and simulate the actual robot assembly production line in the virtual space. The digital twin pursues the life cycle of the whole process, and this

design only adopts two stations of the robot assembly workstation for simulation and simulation. I hope it can provide some help for the research of relevant personnel.

References

1. Xuan Liang, ZHANG Chunfei, Liu Xiaohui, Yang Peijie. Research on digital twin modeling method of manufacturing production line under the background of Industry 4.0. Journal of Jianghan University (Natural Science Edition),2023, vol. 51 (2): 68-77.
2. S. Mihai et al. Digital Twins: A Survey on Enabling Technologies, Challenges, Trends and Future Prospects, Proceedings of IEEE Communications Surveys & Tutorials, vol. 24, no. 4, pp. 2255-2291.
3. J. Wang, H. Liu, Design and Algorithm of Multi-Robot Cooperative Welding and Assembling Workstation Based on RobotStudio, Proceedings of 2019 International Conference on Robots & Intelligent System (ICRIS), Haikou, China, 2019, pp. 14-17.

Authors Introduction

Mr. Huahao Li

He is a first-year master candidate in Tianjin University of Science and Technology, majoring in Intelligent multibody robot.

A Study of Chemical Reactor Simulation System Based on PCS7

Hongshuo Zhai

Tianjin University of Science and Technology, Tianjin, China

E-mail: 1936130557@qq.com
www.tust.edu.cn

Abstract

Since 1960, the reason for the rapid development of industrial control and its systems is the rapid development of the key process control theory, and their application in science and technology is more and more extensive. In general, the factory site has a series of hazards, and people tend to take a remote, accurate approach to remote control. In this research, PCS7 software is used to simulate the remote control of chemical reactor, and reasonable and correct simulation can be obtained after PID correction.

Keywords: PCS7, PID correction, process control, chemical reactor

1. Introduction

Normal human life is inextricably linked to a series of chemical reactions.

The chemical reactor is defined as a vessel for reaction. The device is widely used in various fields, and the ultimate goal is to obtain reactants that meet the requirements while ensuring production efficiency.

However, due to the presence of a large number of harmful gases in the chemical reaction site that are not conducive to the human body and the environment, affecting the physical and mental health of the practical operator, and because of the high requirements of the chemical reaction parameters, the previous on-site operation has been transformed into a more accurate and safe remote control.

This paper mainly adopts PLC-SIM as the controller, and the controlled object and control method are realized by PCS7. Firstly, the infrastructure of chemical reactor PCS7 is introduced, and then the structural model and automatic operation are built by CFC block and SFC block. Finally, the loop is designed and PID correction is carried out to obtain appropriate graphics.

2. PCS7 Project Basic Framework

In most cases, a project has ES stations, OS stations, AS stations and buses, and each station port has its special function [1], which is shown in Figure 2-1.
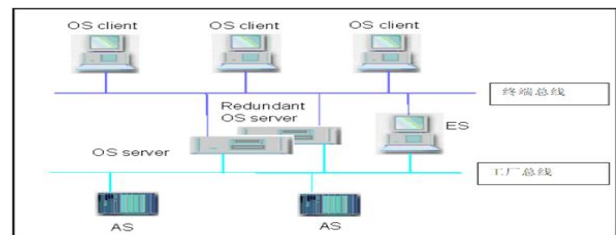
Figure 2-1 PCS7 framework

2.1. Engineer station

The engineer station (ES station) can achieve a series of operations such AS new construction, editing project files, and communication between AS station and OS station. In general, the design of ES station is completed in the software SIMATIC Mananger, and the configuration mainly consists of two links : AS station configuration and OS station configuration.

The configuration of automation station involves the compilation of continuous function diagram and sequential function diagram, factory level design, hardware configuration, communication network configuration and so on.

In the operator station, the process screen design and operation function design interface are mainly, and there are data archiving and protocol design.

2.2. Automation station

The automation station (AS station) is the control station of the Siemens DCS monitoring system, and the main modules of the station are: RACK, PS, CPU, I/O module, etc.

The CPU in PCS7 usually uses 400, in RACK, there are CP modules for the purpose of communication, but some CPU modules also have the function of communication directly, such as PROFIBUS DP interface or PROFIBUS DP serial communication module, through which to achieve device communication.

The automation station operates the system with fast industrial Ethernet access or the engineer station accesses the ET200 distributed I/O station using the Profibus-DP fieldbus. Note: The "block" in PCS7 is very critical, it is related to the work of the CPU in the software STEP7, the specific utility is shown in Table 2-1.

Table 2-1 Utility of blocks

Block	Utility
Tissue block (OB)	The function is to store the user master program and control successively.
System Function Block (SFB) and System Function Call Block (SFC)	They are stored in the S7 CPU so that they can implement some of the reserved functions.
Functional Block (FB)	It is a block that stores the user's own design.
Function call block (FC)	A subroutine that does not take up memory, but can be used repeatedly.
Background Data block (INSTANCE DB)	When a function block and a system function block are used, the block automatically connected with the block. No manual addition is required.
Data block (DB)	Its purpose is to store user data.
System Data Block (SDB)	It stores hardware configuration data.

2.3. Operator station

Most of the time operator stations are where we use the computer as a medium for control purposes. The OS configuration occurs at the engineer station, so it is feasible to consider the OS station project as part of the ES station project. Operator stations under distributed operating systems also have some basic two categories, the following is their introduction.

(1) Operator station server

The function of the operator station server is to pass the value in the program to another client that it has contact with, that is, the operator station client, and transfer the content of the OS station to the corresponding AS function block.

(2) Operator station client

The Data connection between the operator station client and the operator station Server is established through the Terminal Bus, and the data in the operator station server is accessed through Server Data.

The operator stands on the design, the configuration design of the above machine screen is the main work, should have the switch, operation panel CFC and SFC block writing and other functions, simulate the factory screen equipment diagram, in addition, there are data to archive and so on.

2.4. Factory bus and terminal bus

Plant Bus and Terminal Bus adopt Industrial Ethernet, which is consistent with IEC 802.3. This communication mode is mainly applied in network structure, most of which use optical fiber network as communication structure. Both the factory bus and the terminal bus will have good stability, and its role is to achieve reliable communication between various industrial communication buses.

For some medium or large specifications of industrial actual sites, it will have a high ball, for some industrial sites require a relatively smooth, fast contact device, PCS7 to provide it with optical fiber network and equipped with industrial Ethernet technology to achieve fast communication. The system uses advanced optical fiber connection technology and high-speed optical interconnection equipment, making it a new milestone in the field of communication. Among them, the latest and fastest Ethernet communication technology combines the scalable performance of switching technology with the high security of optical fiber ring network, and has a high data transfer rate and 1Gbit/S transfer rate, industrial twisted pair (ITP) and optical cable (FOC) can be used.

3. Mathematical Modeling and Automatic Operation of the Controlled Object

In PCS7, an engineering application, there are many ways to construct mathematical models.

3.1. CFC block

Also known as continuous function diagram, it is based on the function block to establish a numerical model of the controlled object, which has the function block of the

configured proportion, integration, differentiation, delay and operation, such function block is configured on the PCS7 has been programmed.

CFC block is used to build the object model, and operation block is used to construct the mathematical formula, so as to realize the mathematical modeling process of the controlled object.

For the establishment of the mathematical model of chemical reactor, the mathematical modeling of liquid level object, pressure object, feed object and temperature object was obtained by referring to the data and using CFC block.

1. Liquid level objects

Mainly: integral link and inertia link.

$$G(S)=\frac{3.25}{1+5.47S} * \frac{2.13}{0.97S} \qquad (3-1)$$

2. Objects of stress

Mainly: inertia link and integral link.

$$G(S)=\frac{2.78}{1+5.29S} * \frac{0.91}{5.58S} \qquad (3-2)$$

3. Feed flow

Mainly: the inertial link of its principal and subordinate reactants.

$$G(S)=\frac{1.87}{1+5.49S} \qquad (3-3)$$

$$G(S)=\frac{2.39}{1+4.68S} \qquad (3-4)$$

4. Temperature object

Mainly: the second stage of the lag time [2].

$$G(S)=0.6 * \frac{1}{4.29S} * \frac{1}{3.93S} \qquad (3-5)$$

3.2. SFC block

In the process of running, the simulation system automatically executes the sequence control program, uses SFC to initialize the initial value of the control parameters required by the whole control system, and executes each control loop in turn, which is shown in Figure 3-1.

Each control loop enters the automatic control state according to the initial value, and after the implementation of the END step, the system will automatically enter the stable operation state, and the operation switch guide operator mode is fully controlled, that is, the operator can complete the automatic control according to the initial control target value, so that each control loop enters the automatic control state after the parameter initialization.
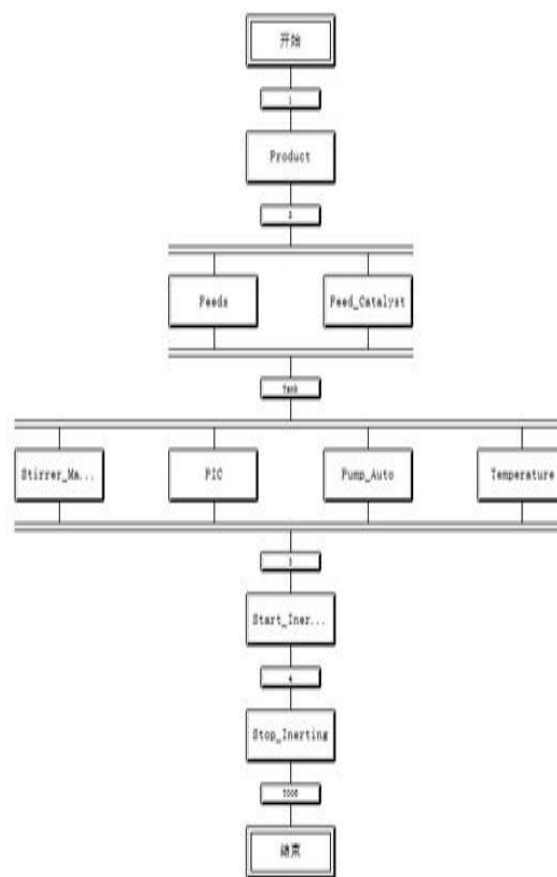
Figure 3-1 Simulating the SFC starting sequence

4. Design of Advanced Control Scheme for Simulation System of Chemical Reactor

The PID controller is used to control the reactor. The PID controller here uses the APL library of PCS7.

This topic has the feed part, the pressure part, the liquid level part and the temperature part.

The feed part is realized by a relatively simple ratio module, in which there are two modules, the master module and the slave module. The former controls the flow rate of the main reactant, the flow rate of the auxiliary reactant is controlled by the set value of the main module, the ratio controller provides two PID controls, and the latter controls the catalyst part of the reactant independently.

The pressure part is controlled by a separate PID module. To achieve this task, an initial setting value is set first. In actual simulation, if the running value exceeds our initial setting, the exhaust valve will be started; otherwise, if the running value is insufficient, the inert gas intake valve will be started until the process value each the set value. Then

the feedback control of the whole system is carried out by proportional integral and differential control method. In the liquid level part, a separate PID module is used to realize it.

In the temperature part, cascade control is used. The inner ring of cascade control adopts PID module to control the flow rate of hot and cold water valves to realize the control of jacket temperature [3]. The output MV of a PID module is adopted in the outer ring as the given value of the PID controller in the inner ring.

4.1. Liquid level control scheme

For the control of liquid level in the loop, a negative feedback closed-loop loop is used, as shown in Figure 4-1.
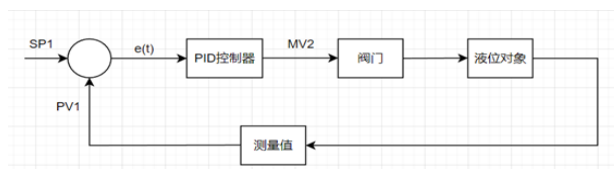
Figure 4-1 Pipes and devices

The value of the set value SP1 can be given by an external or internal PID controller whose parameters are tuned by the PID Tuner.

The liquid level control model, as shown in Figure 4-2 is built using the CFC block of PCS7.
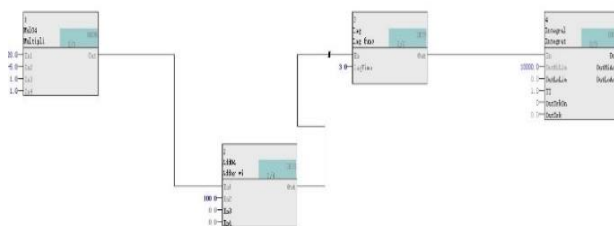
Figure 4-2 Liquid level simulation objects

For these four parts, it is usually used in a user-defined closed-loop control system set by users, integrating the input values according to the ladder rule, and output the results.

Figure 4-3 shows the pipe diagram and function block of the liquid level control loop. A control loop is provided in the figure. The CFC block in Figure 4-4 is the liquid level control loop pipe.
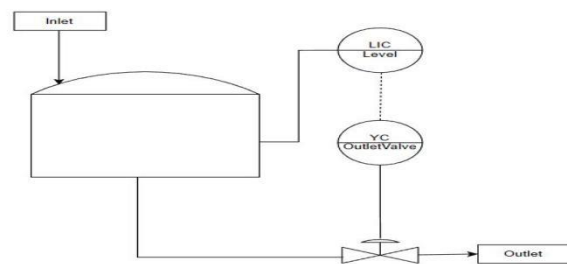
Figure 4-3 Pipes and functional blocks of the

liquid level control loop
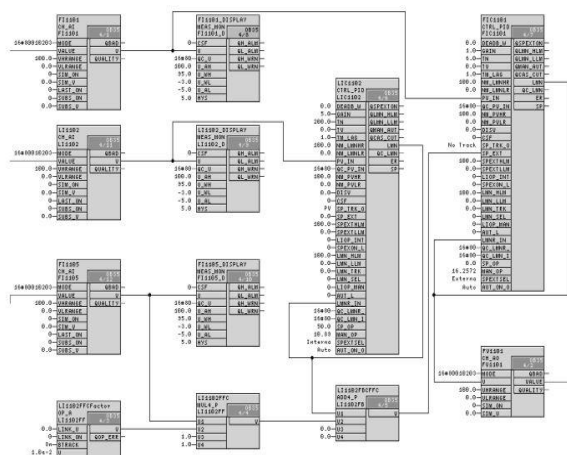
Figure 4-4 Connection structure of each CFC block in

the liquid level control loop

In addition, for the feed flow, pressure, temperature three parts, also to design the corresponding control scheme. Among them, the feed part is controlled by ratio, the pressure part is controlled by division, and the temperature is controlled by cascade.

4.2. Tuning of controller PID parameters

Using PID module to control the controlled object is very important for the parameters, and only appropriate parameters can meet our purpose. This topic only takes 4 control loops for example to study the liquid level control loop.

The PID Tuner tool is developed based on PCS7 software, and the data is collected and recorded through the background data block of PID function block, and the ideal data is obtained by simulation and applied in the controller. Different adjustment effects under various control modes can be achieved by setting. The tool

provides synchronous real-time data curve display of initial, median, and final values, and intuitive overall optimization functions are displayed.

Below, after proper PID tuning, the appropriate tuning pattern is obtained, which is shown in Figure 4-5.
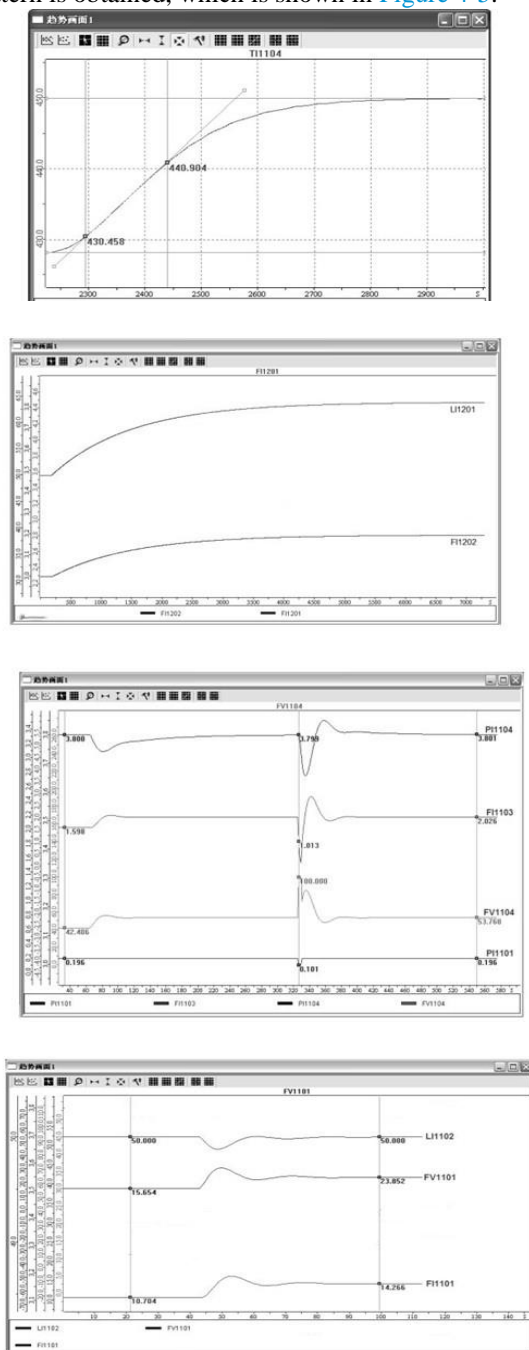
Figure 4-5 Setting results of liquid level, flow rate, pressure, and temperature

References

[1] HongBin Fang, et al. "Simulation Design and Implementation of Chemical Reaction Process Control System based on PCS7." Electric Drive 50.05 (2020): 102-107. doi:10.19457/j.1001-2095.dqcd20821.

[2] Jiangtao Zhang, et al. "Research Progress on Temperature Control of Chemical Reactor." Chemical Automation and Instrumentation 49.04 (2022): 419-423. doi:10.20030/j.cnki.1000-3932.202204002.

[3] Hengfei Wang, Yongliang Chen, et al. "the boiler superheated steam outlet pressure cascade expert PID control." automation instrument 36. 10 (2015) : 99-102. The doi: 10.16086 / j.carol carroll nki issn1000-0380.201510025.

Authors Introduction

Mr. Hongshuo Zhai

He is a first-year master candidate in Tianjin University of Science and Technology.

Analysis of Learning Quality Evaluation for University Student Courses with Process Assessment

Yuhao Zhang[1]*, Ying Gong[2], Xuran Wang[3]

[1]College of Artificial Intelligence, Tianjin University of Science and Technology, 300222, China

[2]College of Economic and Management, Tianjin University of Science and Technology, 300222, China

[3]College of Chemical Engineering and Materials, Tianjin University of Science and Technology, 300222，China

E-mail: *a15839176622@163.com

www.tust.edu.cn

Abstract

With the progressive exploration and application of formative assessment in university pedagogy, this evaluative method has become widely adopted for appraising students' everyday learning attitudes and conditions. Drawing upon pertinent data regarding classroom learning experiences of students at a specific university, this paper employs machine learning, K-means clustering, the TOPSIS evaluation model, and the entropy weighting method to investigate the relationship between formative assessment and the quality of university student learning, culminating in the creation of an evaluation model. This model allows us to pinpoint the key factors influencing student learning attitudes and offers support for formative assessment in the university context.

Keywords: University Process Assessment, attitude toward learning, machine learning, TOPSIS Evaluation model

1. Introduction

With the continuous deepening reform of education and teaching mechanism, the educational method of process assessment has been widely explored and applied by more and more universities. Compared with the outcome assessment, the process assessment can better evaluate the students' learning attitude and state [1].

In the process assessment, students' learning attitude is mainly affected by the objective environment and subjective initiative [2]. This paper will explore which factors have a greater impact on students' learning attitude, build a mathematical model that can reasonably evaluate the school's learning style and class style, and build an evaluation model that can reflect the level of students' learning enthusiasm.

The rest of this article is organized as follows. The second part analyzes the data and explores the factors that reflect students' learning attitude. The third part introduces the construction of evaluation model. The fourth part analyzes the evaluation model and the results obtained. The fifth part summarizes the main content of this paper.

2. Data Analysis

The data collected students' student number, college, class, course code, class code, teacher code, test name, test time, test full score, student score and normalized score. It should be noted that students choose different courses within the scope of core basic courses, some students choose multiple courses, some students choose only one course; Some students may have done work that was not posted by the class teacher; Some students don't take every test for some reason; Teachers in the same course don't publish pre-class tests the same often.

Yuhao Zhang, Ying Gong, Xuran Wang

In order to explore which factors have a greater impact on students' learning attitude, it is necessary to explore the control variables. The following will explore the influence of college factors, the number of courses selected by students, the attendance rate of students, and the factors of major, class, course, teacher and classroom on students' learning attitude.

2.1. College factor

By calculating the average scores of students in different colleges of the same course, and then summarizing all colleges, the average and variance of the scores of students in each college are obtained. The data are shown in Fig. 1.

Fig. 1 College factor visualization

The average score of students in different colleges is different, and the stability of grades is also different. Therefore, the learning atmosphere of different colleges is different, and students' learning attitude is also different. Therefore, students' learning attitude is affected by the factor of college.

2.2. The number of courses taken by students

We define the average of each student's normalized scores on several tests as the student's representative score. The number of courses selected by each student is 1, 2, and 1 of 3. The representative scores of students with 1, 2, and 3 courses selected respectively are calculated, and the corresponding scatter map is shown in Fig. 2.
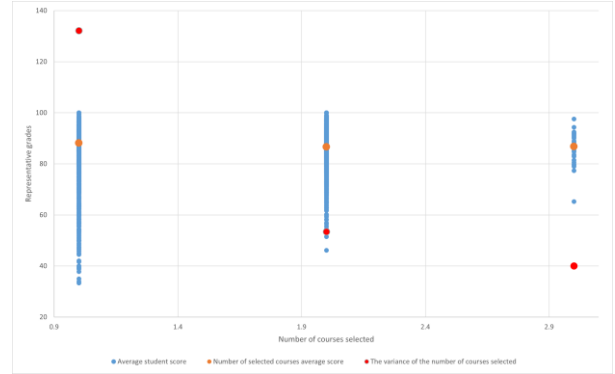
Fig. 2 figure of the number of selected courses

As can be seen from Fig. 2, with the increase in the number of courses selected by students, the average scores of students are basically the same, and the overall scores of students are higher and more and more stable. It can be inferred that the more courses a student chooses, the higher the student's enthusiasm for the course and the better the attitude.

2.3. Student attendance

When the total number of tests for each course is taken as the number of the last measurement for this course, that is, the total number of tests organized by the teacher for this course, the attendance rate of each student is obtained. The scatter plot corresponding to the attendance rate of students and the representative score is shown in Fig. 3.
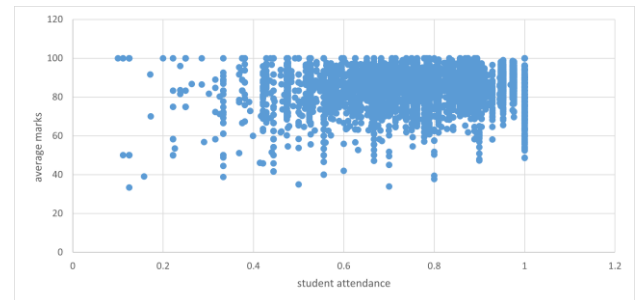
Fig. 3 Scatter plot of attendance

Count the number of people with scores greater than 80 in each attendance range and draw a bar chart, which is shown in Fig. 4.
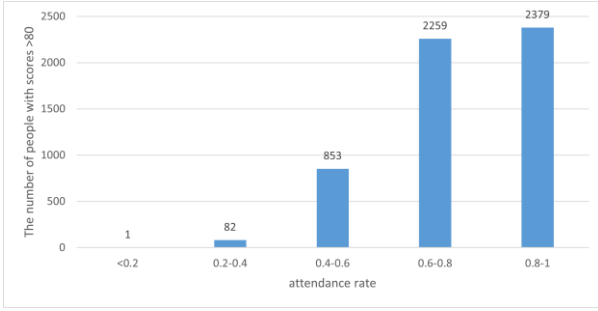
Fig. 4 attendance and high score statistics chart

As can be seen from Fig. 3 and Fig. 4, when students' attendance rate is small (0-0.4), students' high scores are compared. With the increase of students' attendance rate, students' scores are generally higher, and with the increase of students' attendance rate, the number of students who get high scores is increasing. Therefore, students' attendance rate has a great impact on students' scores and can better reflect students' learning attitude.

2.4. Major, class, curriculum, teacher factor

Calculate the average grade and variance of grades for each major, each class, each course, and each classroom, as well as the average grade and variance of students taught by each teacher.

After analysis, the following conclusions are reached:

(1) Students' learning attitude is influenced by their major, but the influence is small. (2) On the whole, students' grades and the stability of grades are also different depending on the class they belong to. Therefore, students' learning attitude is affected by class factors. (3) The overall performance of students taught by some teachers is poor and unstable; The rest of the teachers taught students different overall grades, but the fluctuation is not large, can reflect the true level. Therefore, students' learning attitudes are influenced by their teachers and courses.

3. Construction of Evaluation Model

Data for all students in each college and class is aggregated to get an average of average test scores, average test participation rates, and number of courses taken by students in each college and class. This will result in a dataset consisting of the college (or class), average test score, average test participation rate, and number of courses taken. Since there are few indicators, TOPSIS method is adopted here [3].

3.1. Data standardization

Since attendance, grade point average, and average number of courses taken are all positive indicators, it is only necessary to standardize the dataset consisting of college (or class), average test score, average test participation rate, and number of courses taken.

3.2. Calculate index weight

To ensure the objectivity and scientificity of the index weights, entropy weight method and analytic hierarchy process are used to determine the index weights, and the weights are shown in Fig. 5.

	attendance rate	grade	Average number of courses selected
College weight	0.24935	0.3984	0.35225
Class weight	0.28405	0.3831	0.33285

Fig. 5 index weight

3.3. Calculating the score

Define the maximum value Z^+ as the maximum value of each column in the vector matrix and define the minimum value Z^- as the minimum value of each column in the vector matrix. Define the distance of the i $(i = 1,2,…,n)$ evaluation object to the maximum value as:

$$D_i^+ = \sqrt{\sum_{j=1}^m \omega_j \left(Z_j^+ - z_{ij}\right)^2} \qquad (1)$$

Define the distance of the i $(i = 1,2,…,n)$ evaluation object to the minimum value as:

$$D_i^- = \sqrt{\sum_{j=1}^m \omega_j \left(Z_j^- - z_{ij}\right)^2} \qquad (2)$$

Then, we can calculate the non-normalized score of the i $(i = 1,2,…,n)$ evaluation object as:

$$S_i = \frac{D_i^-}{D_i^+ + D_i^-} \qquad (3)$$

It is evident that $0 \le S_i \le 1$, and the larger S_i the smaller D_i^+, indicating closer proximity to the maximum value. By incorporating the obtained indicator weights into the calculation of sample distances, we obtain the score for each sample [3].

4. Introduction of Result

According to the above model construction, MATLAB is used to calculate, and the final answer is obtained.

4.1. Evaluation hierarchy

The hierarchy of evaluation models is shown in Fig. 6.
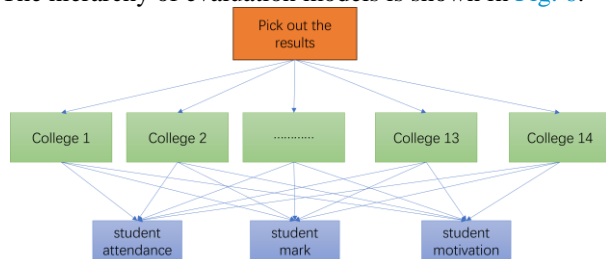
Fig. 6 evaluation hierarchy diagram

4.2. Personal evaluation model

Here, we want to find the best and worst people, and we first cluster them using K-Means into better and worse groups. Then, the two groups were evaluated by TOPSIS, and the evaluation index was the student attendance rate, the number of courses selected, and the individual performance [2].

5. Conclusion

Through the students' usual performance to construct the process evaluation model, get the factors that affect students' learning attitude, and can find the class, college, and students with better performance through the data.

This research will help colleges and universities to carry out process assessment better.

References

1. Li Q, Li Z, Han J, et al., Quality Assurance for Performing Arts Education: A Multi-Dimensional Analysis Approach, *Applied Sciences*, 2022, 12(10): 4813.
2. Zhang Rui, Xu Peng, Fang Mingfeng, a process-based and diversified curriculum assessment method for local comprehensive universities, *Journal of Chongqing University of Technology (Social Sciences)*, 2014, 28(04): 129-133.
3. Yu Xiaofen, Fu Dai, A review of multi-index comprehensive evaluation methods, *Statistics and Decision*, 2004(11): 119-121.

Authors Introduction

Mr. Yuhao Zhang

He is an undergraduate majoring in artificial intelligence at Tianjin University of Science and Technology. He learns about electronic information and artificial intelligence, and currently focuses on digital image processing, deep learning and pose estimation.

Ms. Ying Gong

She is currently a third-year student majoring in financial engineering at the School of Economics and Management of Tianjin University of Science and Technology, focusing on data mining.

Mr. Jiashuai Wang

He is currently studying Computer science and Technology at Tianjin University of Science and Technology, where his research interests are deep learning and computer vision.

Motion Analysis and Transfer Applications Based on Posture Recognition

Yuhao Zhang*, Mingyue Li, Jianhao Jiao

College of Artificial Intelligence, Tianjin University of Science and Technology, 300222, China

E-mail: * a15839176622@163.com

www.tust.edu.cn

Abstract

This paper investigates the posture trajectories in human motion using pose recognition technology based on the deep learning framework MediaPipe. By detecting key points on the human body and plotting and calculating these trajectories in the temporal dimension, we further conduct comparative analysis of these trajectories with professional sports coaches' motion guidance to assist athletes in correcting their posture. Additionally, this technology has been deployed on Jetson Nano, enabling its practical application in mobile scenarios, and providing robust tools and methods for fields such as rehabilitation therapy, sports training, and animal behavior analysis. This study offers insights into the transfer applications of posture recognition.

Keywords: deep learning, MediaPipe, posture recognition, Jetson Nano

1. Introduction

With the continuous development of artificial intelligence, artificial intelligence technology has played a huge role in the development of the sports and fitness industry. Moreover, with the continuous development of economy and society and the continuous improvement of people's living standards, people are more and more willing to strengthen their health through physical exercise. The application of artificial intelligence technology in sports will inject vitality into the sports industry.

People in the usual self-training, because there is no manual coach guidance, which will make the athletes can not get timely feedback when improving skills. Through the use of the MediaPipe framework developed by Goole company to realize the identification of human bone key points, through the changes of bone points in continuous time to draw the human movement trajectory, and then analyze the movement Angle of each action [1]. This will provide a wealth of data for athletes to continuously optimize their level.

The posture analysis of athletes has been deployed on the Jetson Nano, so that athletes can be monitored by moving the jetson nano, whether they are exercising outdoors or indoors [2]. In addition, we added the camera to the steering gear, so that the camera can follow the movement in real time and accurately analyze the attitude data.

In addition, the continuous actions in the same period of time can be transferred to the movement of animals, such as octopus, snake and other animals, through deep learning to train a model to identify key points, and then analyze the movement behavior of animals, which will help us better understand animals [3].

The rest of this article is organized as follows. The second part introduces the hardware. The third part introduces the software part. The fourth part analyzes the captured motion changes. The fifth part summarizes the main content of this paper.

2. The Hardware Structure

The main hardware used in this research is jetson nano, steering gear and high-definition camera. The camera is moved through the steering gear to better capture people's posture information, and the powerful image processing capability of jetson nano is used for data analysis.

Yuhao Zhang, Mingyue Li, Jianhao Jiao

2.1. Jetson nano

Jetson Nano is an embedded computing platform based on NVIDIA Maxwell architecture with a 128-core NVIDIA Maxwell GPU and quad-core ARM Cortex-A57 CPU that delivers high-performance deep learning and computer vision capabilities. The board supports a variety of peripheral connections, including GPIO pins, I2C buses, UART serial ports, and more, making it an ideal platform for integrating a wide range of sensors and actuators [2]. Jetson Nano runs Ubuntu Linux with strong software support for deep learning and computer vision applications. The jetson nano used in this study is shown in Fig. 1.

Fig. 1 Jetson Nano

2.2. IMX415 HD camera

Sony's IMX415 HD camera is our choice. The camera features 8.46M effective pixels, a frame rate of up to 120fps, and a field of view Angle of 130 degrees, providing the system with clear, high-resolution images. To provide the system with clear, high-resolution images. The IMX415 is an advanced image sensor for applications that require high-quality image capture, such as our attitude recognition system. The IMX415 HD camera used in this study is shown in Fig. 2.

Fig. 2 IMX415 HD camera

2.3. DS3230 digital steering gear

Two DS3230 digital steering engines are used in this study. This steering gear provides precise Angle control, enabling the system to dynamically adjust the direction of the camera according to the movement of the target object, achieving a 360-degree rotation. Through the steering engine, we can realize the full range of the target object trajectory capture. The DS3230 digital steering gear used in this study is shown in Fig. 3.

Fig. 3 DS3230 digital steering gear

3. Software Distribution

In software development, this design is based on MediaPipe open-source framework development, and combined with OpenCV to achieve the complete code.

3.1. Introduction to MediaPipe

MediaPipe is an open-source computer vision library developed by Google to provide developers with a flexible and efficient set of tools to build visually aware applications. It is designed to simplify the development process of computer vision tasks by providing pre-trained models and easy-to-use Apis that make it easier for developers to implement a variety of visual tasks. The main performance is as follows:

(1) Real-time performance: MediaPipe focuses on real-time computing, enabling low latency and efficient real-time image processing on multiple hardware platforms.

(2) Multimodal input: Support a variety of input sources, including images, videos, camera streams, etc., making it suitable for a variety of scenes and devices.

(3) Rich pre-trained models: MediaPipe offers a range of pre-trained models covering many areas such as human pose estimation, hand tracking, face detection, etc., which can be used to quickly build various computer vision applications.

(4) Cross-platform support: MediaPipe supports multiple operating systems (including Linux, Windows, Android, iOS) and a variety of hardware accelerators (such as Gpus, Tpus), making it widely portable.

In this design, we chose the MediaPipe framework as the basis to realize the key function of attitude recognition. By combining OpenCV, we were able to flexibly process camera input and use the pre-trained model provided by MediaPipe to accurately capture key pose information of the target object. This combination allows us to quickly build a powerful computer vision system for monitoring and analyzing the movements of target objects.

3.2. Establishment of trajectory analysis model

Let's take the motion analysis of one arm as an example. Using the MediaPipe framework, call the poseLandmarkerResult function to generate the three-dimensional coordinates of the key points. In Fig. 4 below, A, B, and C represent the three nodes of the arm. The poseLandmarkerResult function can be used to return the three-dimensional coordinates of the node, and can draw the trajectory motion [1].
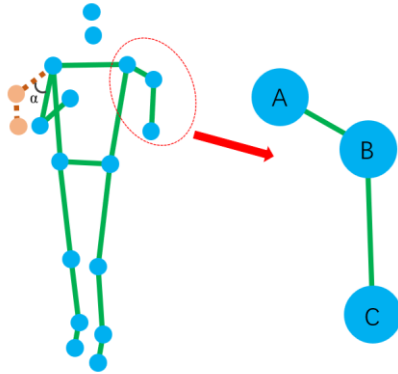
Fig. 4 Keypoint graph

Then according to the following formula, the angular distance change of the node relative to the previous state can be calculated.

$$\theta = arc\,cos\left(\frac{\overrightarrow{AB}*\overrightarrow{BC}}{|\overrightarrow{AB}|\cdot|\overrightarrow{BC}|}\right) \qquad (1)$$

4. Introduction of Result

Running the code on the jetson nano shows the trajectory on the screen and can calculate the Angle between the nodes.

4.1. Effect of display

As shown in Fig. 5 below, the trajectory of the key points of the arm in a period of time can be drawn, so that the movement posture of the professional athlete can be compared, and the change of the Angle of the joint can be accurately viewed to help the athlete find the correct movement posture.

Fig.5 effect display diagram

Through the object detection algorithm, the camera can capture objects such as tennis balls and ping-pong balls, which can better identify the position of people [2]. The display effect is shown in Fig. 6.
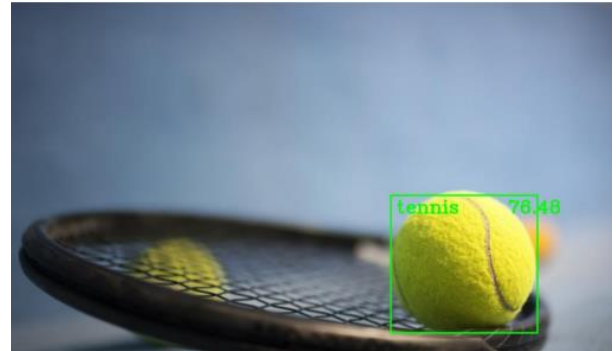
Fig. 6 object identification

4.2. Idea transfer application

It has far-reaching and multi-level significance to transfer human posture recognition technology to animals to realize key point detection and trajectory tracking. This not only provides new research avenues in the field of biology, deepening the understanding of animal ecology, behavior, and physiology, but also brings substantial contributions to wildlife conservation, livestock management and scientific education. Through the application of this technology, we can monitor the movement patterns, posture changes and behavior patterns of animals in real time, providing a scientific

basis for protecting endangered species, monitoring wildlife migration, and improving the efficiency of livestock production [3].

5. Conclusion

Through the movement posture analysis of human arm as an example, the trajectory of the athlete is successfully analyzed, which also reflects the key points of the whole body, and the movement posture analysis will help the athlete to correct the posture in a more three-dimensional way. And this model is deployed on the jetson nano, which will facilitate device movement. The transfer of posture analysis to animal movement behavior will greatly help the study of animals.

Acknowledgements

References

1. Han K, Li X, Research Method of Discontinuous-Gait Image Recognition Based on Human Skeleton Keypoint Extraction, *Sensors*, 2023, 23(16):7274.
2. Zhang Shang, WANG Hengtao, RAN Xiukang, Lightweight traffic sign detection method based on YOLOv5, *Electronic Measurement Technology*, 2002, 45(08): 129-135.
3. Li Yandong, HAO Zongbo, Lei Hang, Convolutional neural networks research review, *Journal of Computer Applications*, 2016, 36(09): 2508-2515+2565.

Authors Introduction

Mr. Yuhao Zhang

He is an undergraduate majoring in artificial intelligence at Tianjin University of Science and Technology. He learns about electronic information and artificial intelligence, and currently focuses on digital image processing, deep learning and pose estimation.

Ms. Mingyue Li
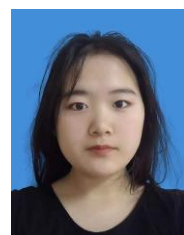
She is an undergraduate majoring in artificial intelligence at Tianjin University of Science and Technology. She currently focuses on digital image processing and deep learning.

Mr. Jianhao Jiao

He is an undergraduate majoring in artificial intelligence at Tianjin University of Science and Technology. He learns about electronic information, and currently focuses on deep learning and pose estimation.

Functional Safety Assessment of the Safety Protection System Based on Petri Net

Peng Wang*, Mengyuan Hu

*College of Electronic Information and Automation, Tianjin University of Science and Technology,
300222, China*

*E-mail: *1090465299@qq.com*

www.tust.edu.cn

Abstract

In this paper, the functional safety evaluation of the safety protection system of gasoline hydrogenation unit was carried out using Petri net. Firstly, the principle and framework of the gasoline hydrogen refueling unit was described. Secondly, the safety integrity level was introduced, and the influencing factors of the safety integrity level were summarized. Thirdly, the Petri net model and the Markov model are compared and the Petri net model is used to verify its security integrity level. Finally, the calculation result demonstrated that the SIL did not reach the target level, and then reached the target level after improvement. This analysis method can provide reference for the safety integrity level evaluation of similar devices.

Keywords: Gasoline hydrogenation, Petri net, Functional safety assessment

1. Introduction

With the passage of time, modern science and technology continue to develop, the number of vehicles climbed. The sulfides produced by automobile exhaust have a great impact on air quality. Therefore, the production of clean gasoline is very important for environmental protection. In response to this problem, we have proposed the hydrogenation of gasoline. The purpose of gasoline hydrogenation is to desulfurize at a small octane loss. Gasoline hydrogenation equipment is a device used for filling the car fuel tank, which includes liquid level sensors, flowmeters, valves and other equipment. Because it is usually operated in high temperature, high pressure and hydrogen environment, there are safety risks, which requires us to do a good job of safety protection, equipped with safety instrument system. Safety instrument system is composed of sensor logic controller and actuator, it can perform one or more absolute safety control instrument functions. Petri networks, which can intuitively describe the relationship between system states and events, are widely applied in many fields. Therefore, the functional safety assessment of the safety protection system of the gasoline hydrogenation unit based on Petri network is a topic worth discussing.

The rest of this article is organized as follows. The second section introduces the research status at home and abroad about the gasoline hydrogenation technology and the safety analysis. In the third part, the gasoline hydrogenation unit is presented. In the fourth section, a example is given to analysis the risk of facility. In the fifth part verifies and improves the safety integrity level of gasoline hydrogenation unit. The sixth part summarizes the main content of this paper.

2. Research Status at Home and Abroad

2.1. Research status of gasoline hydrogenation technology

At present, most of the catalytic raw materials selected by the manufacturers in China have the problem of excessive metal content, as well as excessive sulfur content and non-hydrocarbon components. These problems will not only cause the decline of product quality in the gasoline production process, but also cause the problem of environmental pollution in view of the existence of many

sulfur and other elements in the production process. In the production process, using diolefin saturation treatment can effectively reduce the occurrence of olefin coking phenomenon; in addition, in the relatively mild reaction environment, the influence of impurities can be further reduced, so as to reduce the octane value factors and losses, and better meet the relevant national technical standards and production requirements.

2.2. Research status of safety analysis

The gasoline hydrogenation equipment of a chemical plant in China uses a set of independent safety instrument system (SIS) for safety control. Its process is complex, and the operating environment is high temperature and high pressure, involving flammable and explosive gases, which requires close monitoring. Since its launch, the application of SIS system has significantly improved the stability and reliability of the equipment, realized safe, stable and efficient operation, avoided the occurrence of equipment shutdown and failure problems, at the same time brought economic benefits to the enterprise, but also guaranteed the personal safety of workers.

2.3. Application of petri network in system security analysis

Petri was first proposed by Kal Adam Peri. Petri network represents the system model in the form of a mesh structure model. Petri network is a visual mathematical modeling tool that contains elements such as place, transition, arc and token, which can represent the static function and structure of the system. In addition, Petri networks combine data flow, control flow, and various logical relationships to support a more rich and complex system modeling, allowing for a better analysis of the dynamic behavior of a specific system.

3. Analysis of the Gasoline Hydrogenation Unit

The raw oil is introduced into the filter through an external device, and the purpose of the overall filtration is to separate the particles in the raw material. Among them, after the large particles are removed, they enter the selective hydrogenation feed equipment. The feed can be used as a way to select the hydrogenation reactor feed pump to control the flow rate to achieve a reasonable effect, and then use the supplementary hydrogen mixing mode to form a mixed feed. First, the mixture is sent to the selective hydrogenation reaction device, and the heat exchange is

carried out by means of the hydrodesulfurization reaction. The heating is stopped after the temperature is repeatedly raised to a suitable temperature. Next, the diene to olefin reaction is carried out to meet the requirements. In order to ensure the smooth progress of the whole reaction process, we introduce the catalyst under suitable catalytic conditions for the reaction. After the reaction is completed, we will use a tandem selective hydrogenation reactor to reprocess the product. Through the hydrogenation static desulfurization treatment, and the selective hydrogenation reaction redox reaction, the heat exchange is finally realized through the heat exchanger, and then after the treatment of the catalyst, the saturated state of the olefin product is successfully obtained.

During this process, the reactant is reinjected into the heat exchanger for heat exchange. When it reaches a certain extent, the subsequent reaction of the reactants can be continued. By adjusting the temperature and the interaction of the catalyst, they enter the second reactor for desulfurization treatment, and the catalyst can efficiently achieve the requirement of complete desulfurization. The product of hydrodesulfurization reaction enters the heat separation tank after heat exchange, and the gas phase is added to the product cold separation tank and circulating cleaning equipment after air cooling and cooling of the desulfurization reaction. After the treatment of desulfurization and hydrogenation reaction, the obtained product will be transported to the separation tank after gas phase cooling. In the separation tank, part of the oil phase material will be separated. After it is stable, the non-condensable gas will be transported to the liquid separation tank for further treatment. The separated liquid is discharged after condensation treatment through the condenser, and some of the gas is sent to the gas-liquid separator to mix with water vapor and return to the heat exchanger to continue heating. Then, the product after pressure stabilization treatment is transported to the absorption tower, and then processed again by the liquid separation equipment of the circulating hydrogen compressor. After the mixed hydrogen treatment, it meets the needs of desulfurization production.

4. Device Risk Analysis

4.1. Confirmation of the SIL level

Safety Integrity Level is a measure of the safety of safety instrument system and enterprise safety instrument system management level, its value represents the order of risk

reduction level. The three foundations for determining the level of security integrity are: structural constraints, system capabilities, and hardware security integrity. The safety integrity of hardware mainly depends on the reliability of hardware in the case of dangerous failure. The influence of these factors mainly includes the failure model, system structure, functional test cycle and component reliability level. In the IEC61508 standard, the safety integrity level is divided into four levels, and the operation mode of the system is divided into low requirement operation mode and high requirement operation mode. In the low requirement operation mode, the SIL level is based on the average failure probability, while in the high requirement or continuous operation mode, the SIL level is based on the average failure probability per hour. The requirements for SIL levels for different operating modes, which are shown in Table 1 and Table 2.

Table 1 The SIL level for the low-requirement operation mode

Low-required operation mode		
Safety integrity level	Average probability of failure on demand	Risk reduction factor
（SIL）	（PFD_{avg}）	（RRF）
4	$10^{-4}\sim10^{-5}$	10000~100000
3	$10^{-3}\sim10^{-4}$	1000~10000
2	$10^{-2}\sim10^{-3}$	100~1000
1	$10^{-1}\sim10^{-2}$	10~100

Table 2 The SIL level for the high requirement operation mode

The SIL level for the high requirement operation mode	
Safety Integrity Level	Meverage hourly failure probability
（SIL）	PFH_{avg}
4	$10^{-8}\sim10^{-9}$
3	$10^{-7}\sim10^{-8}$
2	$10^{-6}\sim10^{-7}$
1	$10^{-5}\sim10^{-6}$

The SIS system is more focused on monitoring whether there are risk conditions in the production process and reducing the possibility of risk occurrence. The scheme is a passive system, which is generally in a non-dynamic state. Only when necessary, it will play a role and will not actively participate in the normal operation of the basic process control system. The safety life cycle covers the entire period from the conceptual design of the project to

the discontinuation of SIF. Under certain time and conditions, the possibility of safety-related systems performing their prescribed safety functions can be called safety integrity level (SIL), which represents the level of reducing the risk of safety instrumented systems. It is very important to select the appropriate SIL level. If the selection is too high, it will cause cost waste, and if it is too low, it will bring unacceptable risks. According to the IEC61508 standard, SIL4 level is the highest and SIL1 level is the lowest.

4.2. Influencing factors of safety integrity level
- Failure mode
- Redundant structure
- Common cause failure
- Diagnostic coverage rate
- Periodic function test
- Maintenance mode

4.3. Safety integrity level verification

Hardware security integrity is a part of the overall security integrity of SIS, and its influencing factors include failure mode, component failure rate and detection cycle. When evaluating an interlocking circuit, it is necessary to determine its failure probability to determine its SIL level. Finally, the SIL verification results are compared with the grading analysis results to verify whether the circuit meets the requirements. In order to meet the requirements of SIL grading analysis, it is necessary to reset the layout structure, component selection, redundant design and maintenance plan of the interlocking loop. IEC standard provides a variety of methods to verify SIL level, including simple formula method, reliability block diagram method, fault tree analysis method and Markov model method.

For the transition of the system state, the Markov model is represented by a circle and an arrow, which represent the state of the system and the transition between states, respectively. The Markov model of the 1oo1 structure is shown in Fig. 1.
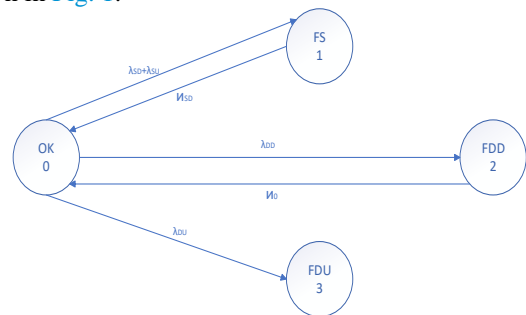
Fig.1. 1oo1 redundant structure Markov model

The state transition matrix of 1oo1 redundant structure is P [1].

$$P = \begin{bmatrix} 1-(\lambda_s + \lambda_D) & \lambda_{SD}+\lambda_{SU} & \lambda_{DD} & \lambda_{DU} \\ \mu_{SD} & 1-\mu_{SD} & 0 & 0 \\ \mu_0 & 0 & 1-\mu_0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

To calculate the mean time between failures, you can use the system state transition diagram, the system state transition diagram is shown in Fig. 2.

Fig.2 1oo1 structure Markov model (MTTFS)

The corresponding matrix is:

$$Q' = 1 - (\lambda_{SD} + \lambda_{SU}) \quad (2)$$

And because:

$$N = [I - Q']^{-1} \quad (3)$$

Therefore:

$$N = \frac{1}{\lambda_{SD} + \lambda_{SU}} \quad (4)$$

The sum of the elements of the N rows of the matrix is the result of MTTFS:

$$MTTF = \frac{1}{\lambda_{SD} + \lambda_{SU}} \quad (5)$$

The formula of system dangerous failure probability is:

$$PFS = S_0 P^{8760} \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}^T \quad (6)$$

The formula of system safety failure probability is:

$$PFD = S_0 P^{8760} \begin{bmatrix} 0 & 0 & 1 & 1 \end{bmatrix}^T \quad (7)$$

Markov model is a model that can represent the dynamic behavior of the system. It can consider multiple reliability factors, and multiple reliability indexes can be obtained by one modeling. However, it will face problems such as space explosion. Especially when the complexity of the system increases, the difficulty of this method will increase greatly.

The stochastic Petri net is used to analyze the hardware failure probability, which solves the problem of Markov state explosion, reveals the dynamic characteristics of the system, and can evaluate the hardware safety integrity of the safety instrumented system.

5. Verification of safety integrity level of gasoline hydrogenation unit

In this section, SIL is evaluated for one of the dangerous scenarios SIF5 of the gasoline hydrogenation unit. The main fuel gas of the fractionator reboiler F9101 of SIF5 is selected to set low pressure and low PSLL6049 A / B / C (2oo3) interlock shutdown XV6011 as a representative. In this process, the components to be considered are as follows.

- SIF serial number:SIF5
- Accident scenario：The pressure of F9201 fuel gas in the hydrodesulfurization heating furnace is too low, which leads to the extinction of the main nozzle. If the fuel gas continues to enter the furnace, the furnace explosion accident may occur.
- Consequence description:The pressure of F9201 fuel gas in the hydrodesulfurization heating furnace is too low, which leads to the extinction of the main nozzle. If the fuel gas continues to enter the furnace, the furnace explosion accident may occur.
- SIF function description:F9101 main fuel gas setting pressure is low PSLL6049A / B / C (2oo3) interlock closed XV6011.
- SIL target: SIL2

In this process, the components to be considered include as shown in Table 3.

Table 3 Components to be considered for SIF5

Sensor Part Pressure Sensor	PSLL6049A / B / C (2oo3)
	Safety barrier (1oo2)
Logic controller part	PLC （2oo3）
	Relay 1 (1oo1)
Actuator	Emergency Shut-off Valve XV6011 (1oo1)
	Relay 2 (1oo1)

In order to evaluate random hardware failures, complete failure data of systems and components are required, considering functional test cycles and common cause failures. These factors will be included in the evaluation category in order to more accurately determine the random hardware failure probability of the system, and then evaluate its hardware security integrity level. The failure

probability of the system needs to meet the system requirements.

$$PFD_{avg} = PFD_{avg\ sensor} + PFD_{avg\ PLC} + PFD_{avg\ actuator} \quad (8)$$

The relay 1, relay 2, emergency cut-off valve and solenoid valve are all 1oo1 redundant structure, while the safety gate is 1oo2 redundant structure to ensure the stability and reliability of the system. Based on the Petri net model, the average failure probability of the emergency shut-off valve is calculated.

The state transition matrix Q is :

$$Q = \begin{bmatrix} -7.48\times10^{-6} & 2.34\times10^{-6} & 4.73\times10^{-6} & 4.13\times10^{-6} \\ 5.71\times10^{-5} & -5.71\times10^{-5} & & \\ 0.04167 & & -0.04167 & \\ 0.125 & & & -0.125 \end{bmatrix} \quad (9)$$

Equation solving:

$$\begin{cases} XQ = 0 \\ \sum_i Xi = 1, 1 \le i \le n \end{cases} \quad (10)$$

Solution:

$$PFD_{avg\ Cut\text{-}off\ valve} = 1.98\times10^{-2}$$
$$PFD_{avg\ Relay\ 1} = 2.455\times10^{-6}$$
$$PFD_{avg\ Solenoid\ valve} = 7.454\times10^{-4}$$
$$PFD_{avg\ Safety\ barrier} = 7.902\times10^{-4}$$
$$PFD_{avg\ Relay\ 2} = 2.5164\times10^{-6}$$

PLC,the system sensor is a 2oo3 redundant structure, and the reachability diagram of the 2oo3 structure Petri net is shown in Fig. 3.
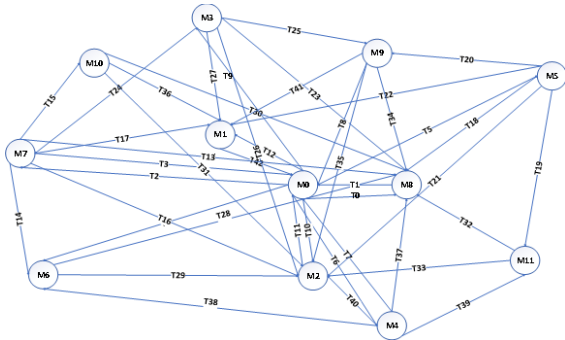
Fig. 3 2oo3 structure reachable graph

The state matrix is:

$$Q = \begin{bmatrix} -0.125 & 4.38\times10^{-7} & 6.54\times10^{-7} & 1.75\times10^{-6} & 2.62\times10^{-6} & 1.81\times10^{-6} \\ 5.71\times10^{-5} & -5.71\times10^{-5} & 0 & 0 & 0 & 0 \\ 0.125 & 0 & -0.125 & 0 & 0 & 0 \\ 0 & 1.31\times10^{-6} & 1.97\times10^{-6} & -6.66\times10^{-6} & 0 & 0 \\ 0.125 & 0 & 3.28\times10^{-6} & 0 & -0.125 & 0 \\ 0 & 1.46\times10^{-7} & 2.18\times10^{-7} & 0 & 0 & -6.66\times10^{-6} \\ 0 & 0 & 1.82\times10^{-6} & 0 & 0 & 0 \\ 2.71\times10^{-6} & 1.46\times10^{-7} & 2.18\times10^{-7} & 0 & 0 & 0 \\ 0.0417 & 0 & 0 & 0 & 0 & 0 \\ 0.125 & 7.28\times10^{-7} & 1.08\times10^{-6} & 0 & 0 & 0 \\ 0 & 7.28\times10^{-7} & 1.08\times10^{-6} & 0 & 0 & 0 \\ 0 & 0 & 1.82\times10^{-6} & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 0.125 & 2.71\times10^{-6} & 1.13\times10^{-6} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.80\times10^{-6} & 3.76\times10^{-6} & 1.20\times10^{-6} & 0 & 0 \\ 1.80\times10^{-6} & 0 & 3.76\times10^{-6} & 0 & 0 & 1.20\times10^{-6} \\ 0 & 0 & 3.38\times10^{-6} & 1.16\times10^{-6} & 0 & 1.75\times10^{-6} \\ -3.7\times10^{-6} & 0 & 1.88\times10^{-6} & 0 & 0 & 0 \\ 1.75\times10^{-6} & -9.37\times10^{-6} & 3.88\times10^{-6} & 0 & 1.16\times10^{-6} & 0 \\ 0 & 0 & -0.0417 & 0 & 0 & 0 \\ 0 & 0 & 1.88\times10^{-6} & -0.125 & 0 & 0 \\ 0 & 0 & 1.88\times10^{-6} & 0 & -3.69\times10^{-6} & 0 \\ 0 & 0 & 1.88\times10^{-6} & 0 & 0 & 3.7\times10^{-6} \end{bmatrix} \quad (11)$$

Equation solving:

$$\begin{cases} XQ = 0 \\ \sum_i Xi = 1, 1 \le i \le n \end{cases} \quad (12)$$

Solution:

$$PFD_{avgPLC} = 1.21\times10^{-6}$$
$$PFD_{avg\ sensor} = 2.3\times10^{-4}$$

Then $PFD_{SIS} = PFD_{avg\ cut\text{-}off\ valve} + PFD_{avg\ relay1} + PFD_{avg\ solenoid\ valve} + PFD_{avg\ relay2} + PFD_{avg\ safety\ barrier} + PFD_{avg\ PLC} + PFD_{avg\ sensor} = 2.15\times10^{-2}$, which does not meet the SIL2 requirements and needs to be adjusted. It can be seen from Table 3-10 that the shut-off valve has the greatest influence on the total PFD of the equipment is the shut-off valve,which is shown in Table 4.

Table 4 The contribution of subsystem to system PFD

Systematic name	Proportion of impact on PFD
Cut-off valve	90%
Solenoid valve	3%
Pressure sensor	4%
Safety barrier	3%

Therefore, we can take the transformation of the emergency shut-off valve and increase the emergency shut-off valve group to realize the transformation from the 1oo1 structure to the 1oo2 redundant structure. After the transformation, the $PFD_{avg\ shut\text{-}off\ valve'} = 4.36\times10^{-3}$. After the transformation, $PFD_{SIS} = PFD_{avg\ shut\text{-}off\ valve'} + PFD_{avg\ relay1} + PFD_{avg\ solenoid\ valve} + PFD_{avg\ relay2} + PFD_{avg\ safety\ barrier} + PFD_{avg}$

$_{PLC}$ + PFD$_{avg\ sensor}$ = 6.13×10^{-3}, which meets the SIL2 safety requirements.

6. Conclusion

In recent years, China has paid more and more attention to production safety, and the government has also issued relevant policies and regulations for industries with high risks. In this context, the automation level of China 's petrochemical enterprises has been continuously improved, which provides good conditions for ensuring production safety. Driven by the policy, with the superposition of new facilities and demand and the increase in the demand for facilities upgrading, SIS system will be more widely used and play a vital role in today 's society and the country. The main conclusion of this paper is that the security level of the system depends on many factors, among which structural constraints, system capability and hardware security integrity level are crucial factors. In order to ensure the safe operation of the system, it is necessary to make reasonable analysis and judgment on these factors. Explore a variety of factors that affect the functional safety of safety instrumented systems, and use Petri net models to model and optimize them, so as to improve the level of safety instrumented systems.

References

1. Guo H, Yang X. Automatic creation of Markov models for reliability assessment of safety instrumented systems. Reliability Engineering & System Safety, 2008, 93(6):829-837.

Authors Introduction

Ms. Peng Wang

She is a postgraduate tutor of Tianjin University of Science and Technology. In 2014, she received a doctorate from North China Electric Power University. The research direction is the functional safety assessment of safety instrumented systems.

Ms. Mengyuan Hu

In 2023, she received her Bachelor of Engineering degree from the School of Electronic Information and Automation, Tianjin University of Science and Technology, China. She is pursuing a master's degree in engineering from Tianjin University of Science and Technology.

Pedestrian Attribute Recognition Based on Deep Learning

Peng Wang*, Qikun Wang, Shengfeng Wang

*College of Electronic Information and Automation, Tianjin University of Science and Technology,
300222, China*

E-mail: *2324365941@qq.com

www.tust.edu.cn

Abstract

This paper studied pedestrian attribute recognition based on deep learning, for its importance in the fields of smart city construction. Firstly, the research status of pedestrian attribute recognition and common deep learning models was introduced. Secondly, considering the accuracy decline problem and gradient problem of the neural network, the residual network was used as the main body of the neural network model. Thirdly, the model was trained to classify multiple person attributes through two data sets, Market-1501 and DukeMTMC-reID. Finally, the pedestrian attribute recognition model was tested, and good results were obtained.

Keywords: Neural networks, Pedestrian attributes, ResNet50

1. Introduction

The contribution of attribute recognition technology in medical, security, intelligent furniture and other fields has attracted more and more attention. After the introduction of deep learning algorithms in the computer field, the computer has realized the processing and application of massive information through continuous learning. The person attribute recognition based on deep learning extracts the feature information of a known pedestrian photo through the convolution and pooling network model, and classifies it to obtain several attributes about the person. The acquisition of these attributes brings important practical applications to the fields of smart city construction and military security.

In this paper, we first introduce the research background and current research status of pedestrian attribute recognition. In the second chapter, the theoretical basis of deep learning will be introduced, and the attribute recognition of pedestrians will be mainly studied and verified by relevant experiments. In the third chapter, the full text is summarized, discussed and analyzed.

2. Methods and results of pedestrian attribute recognition

In the field of computer vision, deep learning has become a basic tool. In this study, a large number of deep learning concepts will be used, and deep learning neural networks will be used to identify pedestrian attributes. Finally, results will be obtained in experiments..

2.1. The emergence of artificial neural networks

T Landahl et al. first proposed artificial neural network, which is a network model built by imitating the connection of nerve cells in the nervous system of animals. As shown in Fig. 1, each circle represents neurons and arrows represent the direction of signal transmission, which is similar to the synapses in nerve cells to transmit signals.
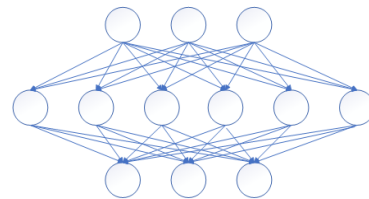
Fig. 1 Artificial neural network structure diagram

Peng Wang，Qikun Wang, Shengfeng Wang

As the basic part of the neural network model, the calculation formula of neurons is as follows:

$$y = f_{active}(\sum_{i=1}^{n} w_i * x_i + b) \quad (1)$$

Where, xi represents the ith input of the neuron, wi is the weight parameter of the input, which can be changed, n represents a total of n neuron inputs, and b is the bias value. Here is called the activation function, a neuron multiplies the input with a weight and then adds a bias value, which is output to the next layer through the activation function.

In the deep learning network, the weight value and bias value are the parameters that need to be learned. Through continuous learning, the gap between the output value and the real value becomes smaller and smaller.

2.2. Principle analysis of convolution layer

As shown in Fig. 2, the function of the convolution layer is to use multiple convolution operations to extract multi-channel eigenvalues from the output of the upper layer, and then send these eigenvalues as outputs to the next layer. The convolution layer reduces the parameters to be trained in the process of layer by layer extraction. In a given image, the local pixel information is converted to the corresponding information of the output image after weighted addition. The part of the convolution layer used for convolution is called the convolution kernel, its size can be defined, and the parameters in it are the parameters we want to train.

Other parameters we define in the convolution kernel determine the effect of image data extraction, including fill, number of output channels, number of input channels, and step size. The size of the convolution image is:

$$H_{out} = \left[\frac{H_{in}+2\times padding[0]-dilation[0]\times(ker\,nelsize[0]-1)-1}{stride[0]} + 1\right] \quad (2)$$

$$W_{out} = \left[\frac{W_{in}+2\times padding[1]-dilation[1]\times(ker\,nelsize[1]-1)-1}{stride[1]} + 1\right] \quad (3)$$
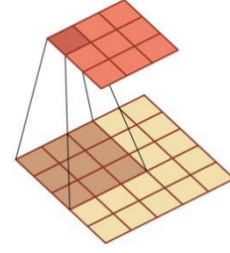
Fig. 2. Schematic diagram of convolution operation of 3*3 convolution kernel

2.3. Principles and advantages of ResNet network

With the more and more extensive application of deep learning in computer vision, the depth of neural network is also deepening. However, it is found that with the deepening of network depth, the results obtained are not getting better and better, and the problem of gradient disappearance is becoming more and more serious. Therefore, He et al proposed residual convolutional neural network. The residual block structure is added to the network structure, and the short circuit design is added next to the convolutional layer to solve the problem of the gradient disappearing and the accuracy of the training set.

ResNet is divided into 18, 34, 50, 101 and 152 layers, which differ in the number of layers with 4 groups of convolutional layers in the middle.

2.4. Meaning of loss function

The loss function is an operation function used to measure the difference between the predicted value f(x) of the model and the real value Y. It is a non-negative real value function, usually expressed by L(Y, f(x)). The smaller the loss function, the better the robustness of the model. We use is called the two-dimensional cross-entropy loss function. This function is a binary classification function, BCELoss is the binary loss function between the target value and the predicted value, the formula is:

$$l_n = -w_n * [y_n * log\,x_n + (1 - y_n) * log(1 - x_n)] \quad (4)$$

Where Wn represents the matrix of weights, Xn represents the prediction matrix of function output, and Yn represents the target matrix.

2.5. Introduction and application of data sets

In recent years, thanks to the exploration of many researchers on pedestrian attribute recognition, there are many open source data sets that can be used, and we will mainly use the following two data sets.

First, the first data set is Mark-1501 data set.

3-1(a)Picture under camera a

3-1(b)Pictures in other cameras
Fig. 3 Markrt-1501 data set

As shown in Fig. 3, the Mark-1501 dataset was originally used for gender reidentification research, and was filmed on the Tsinghua campus using six cameras with different viewing angles. Lin et al. [1]. annotated attributes for each person in this dataset. In this dataset, a total of 1501 pedestrians and 32,668 character rectangles are included, as shown in Table 1. In so many images, there are two parts: the training set and the test set.

Table 1 Labeling probabilities of some data attributes for Mark-1501

label	probability	label	probability
Young	0.0186	Adult	0.2130
Old	0.0107	Bag	0.2463
Handbag	0.1145	Downblue	0.1638
Downbrown	0.0919	Downgreen	0.0186
Downpink	0.0386	Downwhite	0.0772
Downyellow	0.0133	Upblue	0.0613
upgreen	0.0746	Uppurple	0.0399
upred	0.3901	Upyellow	0.3901
clothes	0.3901	Up	0.9481
hair	0.3262	gender	0.4261

Then, the second data set is DukeMTMC-reID.

In the training set, there were 12,936 images, including 751 pedestrians, with an average of 17.2 images per pedestrian. In the test set, there were 750 pedestrians with 19,732 images, an average of 26.3 images per person.

As shown in Fig. 4, the DukemtMC-Reid dataset is a subset of the DukeMTMC dataset. On the campus of Duke University, eight cameras were used to capture the video stored in the form of a pedestrian border box, each frame of which was manually marked by someone. In the video, images are captured every 120 frames, and the resulting images make up the data set. There were 1,404 pedestrians in the dataset, most of whom were captured by two or more cameras, and 36,411 images made up the DukeMTMC-reID dataset. In this data set, the training set consists of seven hundred and two images of people in rows randomly selected, and the remaining images are used as the test set. It is worth noting that there are 408 people caught by only one camera, and they are included as interference items in the data set. Lin et al. labeled a total of 23 attributes in this dataset, including gender, whether to wear boots, whether to wear a hat, whether to have a backpack, shoe color, seven lower body clothing colors and eight upper body clothing colors.

4-1(a)Picture under camera a

4-2(b)Pictures in other cameras
Fig.4.DukeMTMC-reID

2.6. Evaluation index design

The number of correct identifications for each semantic attribute and the number of all samples is calculated, and the quotient of these two numbers is evaluated as the accuracy of each attribute identification. The average accuracy of all attributes can indicate the extent of the model effect. The formula used is as follows:

$$acc_i = \frac{T_i}{N} \tag{5}$$

$$acc_{mean} = \frac{\sum_{i=1}^{c} acc_i}{c} \tag{6}$$

In addition, due to the great imbalance of pedestrian attribute sample data, the negative sample of a certain

attribute may occupy more than 90%. Therefore, a balanced indicator is also needed, and the average accuracy rate (AP) can balance the identification accuracy of positive and negative samples. The mAP of the average accuracy of each attribute can show the superiority of the algorithm on unbalanced data sets. For the ith attribute, TPi represents the number of positive samples correctly predicted, Pi is the total number of positive samples, TNi represents the number of negative samples correctly predicted, and Ni is the total number of negative samples, then:

$$AP_i = \frac{\frac{TP_i}{P_i} + \frac{TN_i}{N_i}}{2} \qquad (7)$$

$$mAP = \frac{\sum_{i=1}^{c} AP_i}{c} \qquad (8)$$

For the research of pedestrian attribute recognition, we should not only pay attention to the accuracy of individual attributes, but also pay attention to how many attributes in each image can be successfully recognized. This index mainly statistics the accuracy (acc.), accuracy (prec.), recall rate (Recc.) and F1 values of attribute recognition at the sample level.

Accuracy: In both positive and negative cases, the proportion of the predicted correct number to the total number is expressed by the formula:

$$ACC = \frac{TP+TN}{TP+FP+FN+TN} \qquad (9)$$

Precision is relative to the prediction results of positive examples. The accuracy of the predicted positive examples is evaluated by the proportion of the real positive examples in the predicted positive examples. The formula is as follows :

$$precision = \frac{TP}{TP+FP} \qquad (10)$$

The recall rate is judged according to the actual sample. Its main purpose is to judge the proportion of the predicted positive example in the actual positive example, and can be expressed by the formula :

$$recall = \frac{TP}{TP+FN} \qquad (11)$$

The F1 score is the harmonic mean of the correct rate and the recall rate, defined as :

$$F1 = \frac{2*precision*recall}{precision+recall} \qquad (12)$$

2.7. experimental result analysis

The conclusion on Mark-1501 is the average accuracy is 0.9631; the average F1 score is 0.6492. The following figures (Fig. 5 and Fig. 6) show the recognition results of some attributes.
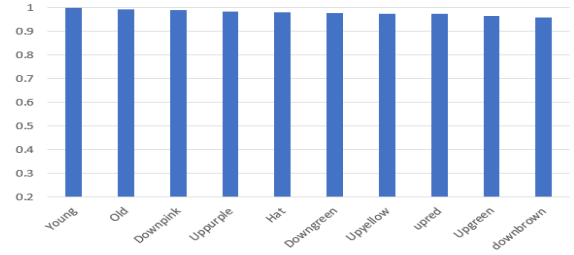
Fig.5. Accuracy of some attributes
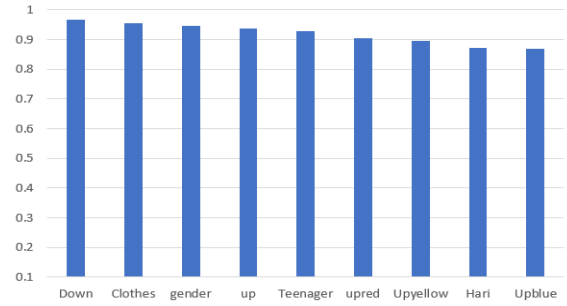
Fig.6. Precision of some attributes

On the DukeMTMC-reid dataset, we concluded that the average accuracy was 0.9152 and the average F1 score was 0.5739.The following figure (Fig.7 and Fig.8)shows the recognition results of some attributes.As shown in the figure, figure 7 shows partial accuracy, and figure 8 shows partial precision.

Fig. 7 Accuracy of some attribute

Fig. 8 Precision of some attributes

3. Conclusion
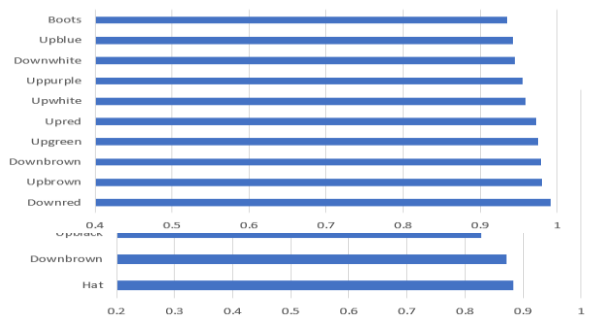
In this paper, nonlinear functions are introduced into a four-dimensional conservative chaotic system to generate multiple scrolls. After the introduction of nonlinear function, the equilibrium point of the system changes from a fixed point to a set of equilibrium points. The system carries out basic characteristic analysis, and discusses its divergence, equilibrium point and whether the energy is

conservative. The equilibrium points obtained by introducing one-dimensional nonlinear function are divided into two categories. For its Lyapunov exponent analysis, after introducing the sine function without multiple angles, The Lyapunov exponents of the system equations show similar periodic characteristics to the sine function, and the Lyapunov exponents obtained by changing the initial values are very different. With the change of initial value, the phase diagrams obtained are also different, and the number of vortex attractors formed is also different, which verifies the multi stability.

Then, the nonlinear function is extended, two nonlinear functions are introduced, and the system with two nonlinear functions is further analyzed. The obtained phase diagram changes from one-dimensional to two-dimensional scroll attractor, and the relationship between the number and arrangement of scroll and the threshold width is obtained. By changing the initial value, the phase diagram with different internal distribution but the same number of scroll is obtained.

References

1. LIN Y, ZHENG L, ZHENG Z, et al. Improving Person Re-identification by Attribute and Identity Learning. arXiv:1703.07220 [cs.CV], 2017.

Authors Introduction

Ms. Peng Wang

She is a postgraduate tutor of Tianjin University of Science and Technology. In 2014, she received a doctorate from North China Electric Power University. The research direction is the functional safety assessment of safety instrumented systems.

.

Mr. Shengfeng Wang

In 2023, he received his Bachelor of Engineering degree from the School of Electronic Information and Automation, Tianjin University of Science and Technology, China.He is pursuing a master's degree in engineering from Tianjin University of Science and Technology.

Mr. Qikun Wang

In 1996, he received his Bachelor of Engineering degree from the School of Electronic Information and Automation, Tianjin University of Science and Technology, China. He is a senior engineer, research direction is HVAC, building technology.

Simulation of office air conditioning air supply based on COMSOL

Peng Wang*, Mengda Liu, Qikun Wang

College of Electronic Information and Automation, Tianjin University of Science and Technology, 300222, China

E-mail: *1789025463@qq.com

www.tust.edu.cn

Abstract

This paper analyzed the summer thermal environment of an office in Tianjin University of Science and Technology based on COMSOL software. Firstly, the principle of indoor thermal environment distribution was introduced. Secondly, according to the actual position of the furniture in the office, the mathematical model is constructed based on the basic theory of computational fluid dynamics. Thirdly, the COMSOL software was used for simulation and calculation, combining boundary conditions. Finally, the simulation results are analyzed through the simulated indoor three-dimensional velocity field and temperature field. The simulation results show that the air conditioning supply can well achieve indoor occupants' comfort.

Keywords: Temperature field, Wind speed field, Indoor thermal environment, CFD simulation, COMSOL

1. Introduction

Indoor temperature and velocity are one of the most important and frequently measured parameters in indoor environments, affecting personnel thermal comfort, energy balance and air flow [1]. Because the indoor temperature field and velocity field is a complex physical field, and there are many influencing factors. The change process of indoor temperature and velocity has the characteristics of large inertia, non-linearity, and easy to be affected by external environmental factors. Therefore, the indoor temperature distribution of most buildings is uneven, and the temperature data of different areas such as the corner, center, floor and ground of the room are very different. Therefore, it is necessary to use COMSOL fluid dynamics simulation software for indoor temperature and velocity simulation.

The first chapter introduces the necessity of simulating indoor temperature and velocity fields. The second chapter mainly introduces the construction of the indoor model, including the indoor objects, the heat transfer coefficient of the wall, the heat transfer coefficient of the window, etc. Meshing of indoor models. In Chapter 3, the simulated temperature and velocity fields are analyzed. In Chapter 4, simulation conclusions are drawn.

2. Indoor Physical Model Simulation

2.1. Model selection

The research object is an office on the fourth floor facing south. The north window and the corridor are adjacent. Due to the different functions of the room, the interior structure will also be different. In addition, the parameters of the office building walls, indoor personnel activities and room ventilation times and other factors, all greatly affect the office internal thermal environment. Since there are many objects in the office, not all objects can be considered, so the physical model only considers two types of objects, one is large objects such as desks, chairs and cabinets, and the other is hot objects such as computers, notebooks, human bodies and refrigerators. All objects are simplified to cuboids, which is conducive to grid division and calculation convergence.

This paper takes an office of Tianjin University of Science and Technology as the research object. The geometric size of the office is 6m×7.m×3.5m, and there is an air conditioner in the southwest corner of the room. The air supply port of the air conditioner is 1m×0.5m. The return air port is simulated as the return air port of the air conditioner, which is located below the air outlet. One of the Windows is located on the south wall with dimensions of 4m×2.3m and 1m above the ground. Suppose there is no

Peng Wang, Mengda Liu, Qikun Wang

turnover in the office，the number of people in the office is six. The physical model of the office is shown in Fig. 1 and Fig. 2.
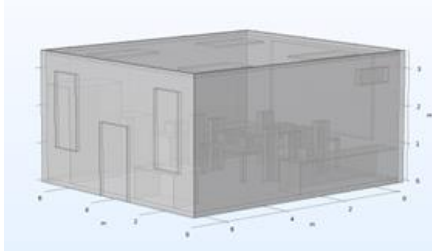
Fig. 1 The exterior of the office physical mode
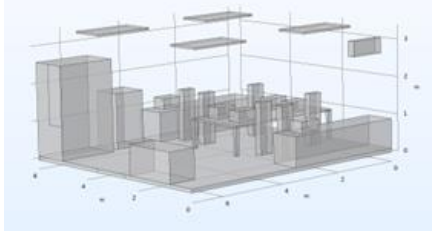
Fig. 2 The interior of the office physical model

2.2. Governing equation

Mass conservation equation is the embodiment of mass conservation law in fluid mechanics [2], as shown in Eq. (1), u,v,w--velocity component, unit m/s,ρ--Fluid density in kg/m³.

$$\frac{\partial(\rho)}{\partial t} + \frac{\partial(\rho u)}{\partial x} + \frac{\partial(\rho v)}{\partial y} + \frac{\partial(\rho w)}{\partial z} = 0 \qquad (1)$$

The hysical meaning of the energy conservation equation is that the increment of the total energy of the fluid controlled in the body per unit time is equal to the sum of the energy increment generated by the fluid capacity of the inflow control and the external action, which includes the energy increment generated by thermal radiation, heat conduction, work done by volume force, work done by pressure and work done by viscous force.as shown in Eq. (2).

$$\frac{\partial(\rho T)}{\partial t} + \frac{\partial(\rho uT)}{\partial x} + \frac{\partial(\rho vT)}{\partial y} + \frac{\partial(\rho wT)}{\partial z} =$$
$$\frac{\partial}{\partial x}\left(\frac{\mu}{pr}\frac{\partial T}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{\mu}{pr}\frac{\partial \emptyset}{\partial y}\right) + \frac{\partial}{\partial x}\left(\frac{\mu}{pr}\frac{\partial \varphi}{\partial z}\right) + S_T \quad (2)$$

2.3. Mesh partition

COMSOL 6.0 software was used to model the office, the transient solver was used to solve the established model, and the dynamic simulation of the air conditioning supply situation of the office was carried out. The computed boundary conditions include the properties of the wall structural materials, the outdoor temperature of 30 degrees Celsius, the corridor temperature of 30 degrees Celsius, the air outlet speed of the air conditioner of 1m/s, and the air outlet temperature of the air conditioner of 20 degrees Celsius. Assuming that the indoor and ambient temperatures are the same at the beginning, the whole office is divided using a conventional grid to calculate laminar flow, solid and fluid heat transfer, and non-isothermal flows. The state of the office temperature field and velocity field within 6000s of air supply is simulated and calculated. The boundary conditions and heat sources are shown in Table 1 and Table 2.

Table 1 Boundary condition

Boundary	Heat transfer coefficient
Body of wall	5(W/(m2*K))
Windows	10(W/(m2*K))
Door	4(W/(m2*K))

Table 2 Source of heat

Source of heat	Flux of heat(W)
computer	100
people	50
80	80

3. Image Analysis

The temperature cloud image at room x=1, x=3, x=5m (as shown in Fig. 3, Fig. 4, Fig. 5) and velocity cloud image (as shown in Fig. 6, Fig. 7, Fig. 8) were captured respectively. The scale on the right of the temperature cloud image is the size of the temperature, corresponding to 20 to 30 degrees Celsius from blue to red, and the scale on the right of the velocity cloud image corresponds to the size of 0 to 2m/s. Fig. 3 shows the temperature cloud map at x=1m. The upper left corner is close to the air outlet of the air conditioner, where cold air blows into the room and the temperature is low. The temperature is about 20 degrees Celsius according to the scale. Fig. 4 and Fig. 5 are x=3m and x=5m respectively. The reason why the temperature in the lower left corner is relatively low is that the cold air from the air conditioner blows out and falls down due to gravity, and the cold air gathers to the ground on the left side, where the temperature is about 25 degrees Celsius. The combination of the three images shows that the temperature on the upper floor of the room is evenly distributed, about 27 degrees Celsius. However, the

position temperature of the air outlet is low, which is not conducive to the comfort of indoor personnel.
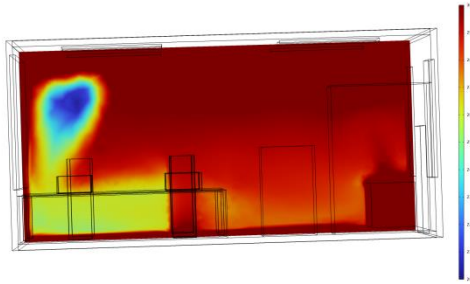
Fig. 3 The temperature field at x=1m
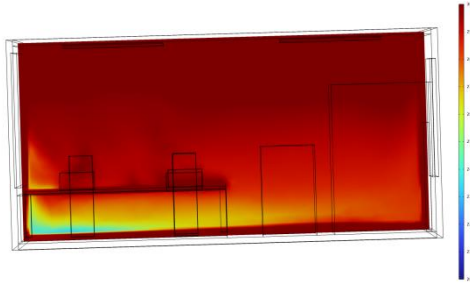
Fig. 4 The temperature field at x=3m
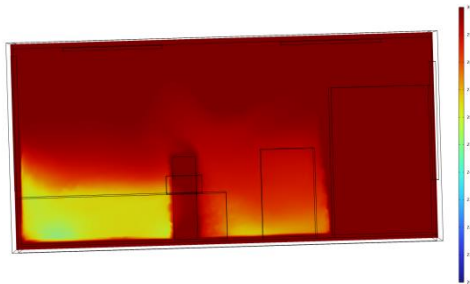
Fig. 5 The temperature field at x=5m

Fig. 6 shows the indoor velocity cloud map at x=1m. The reason for the large wind speed in the upper right corner is that it is located at the air conditioning outlet, and the wind speed is 2m/s. The reason for the large wind speed in Fig. 7 and Fig. 8 is that the wind speed in other places is evenly distributed due to the cold air blown by the air conditioner. The comprehensive wind speed cloud map shows that the wind speed in other places is evenly distributed except for the air conditioner outlet, and the indoor personnel are more comfortable.
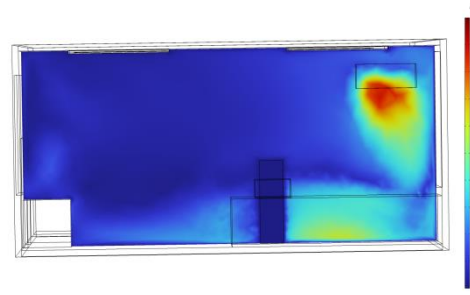
Fig. 6 The velocity field at x=1m
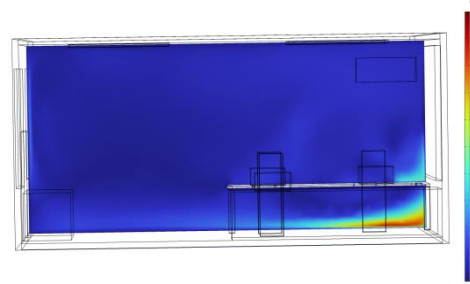
Fig. 7 The velocity field at x=3m
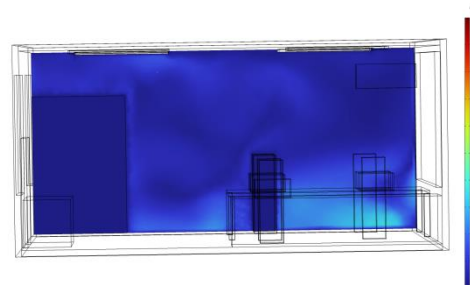
Fig. 8 The velocity field at x=3m

The cross sections at y=1m and y=3.5m were intercepted, respectively, where Fig. 9 and Fig. 10 show the temperature cloud, and Fig. 11 and Fig. 12 show the velocity cloud. Fig. 9 shows the position of the air outlet of the air conditioner. It can be seen from the figure that after the air from the air conditioner blows out, it falls to the position of the table due to gravity, and the temperature is low. Fig. 10 shows that the temperature distribution in the upper and lower parts of the room is relatively uniform for the location located in the center of the room at y=3.5m.

Fig.9 The temperature field at y=1m

Peng Wang, Mengda Liu, Qikun Wang
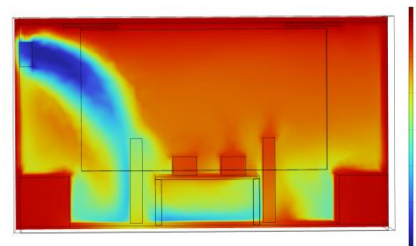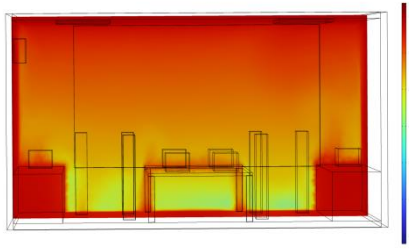
Fig. 10 The temperature field at y=3.5m

Combined with Fig. 11 and Fig. 12, the comfort of indoor personnel is not high under the air conditioning, and the comfort of personnel is good in the middle position of the room.
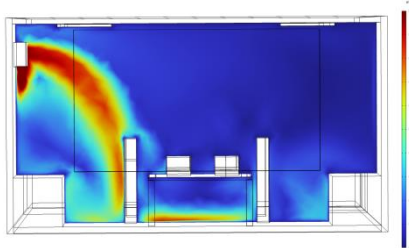
Fig. 11 The velocity field at y=1m
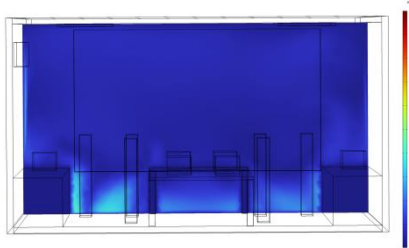
Fig. 12 The velocity field at y=3.5m

Screenshots at z=0.2m and z=1.75m were taken, respectively, where Fig. 13 and Fig. 14 show the temperature cloud, and Fig. 15 and Fig. 16 show the velocity cloud. Combined with the above analysis, the upper temperature of the room is higher than the lower temperature, and with the height increase, the temperature also increases. The temperature distribution is uniform in the z plane.
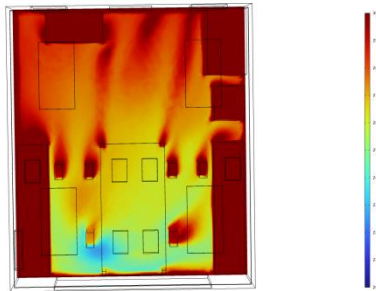
Fig. 13 The temperature field at z=0.2m
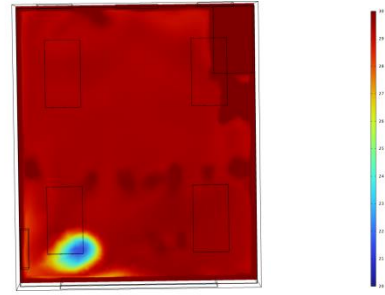
Fig. 14 The temperature field at z=1.75m
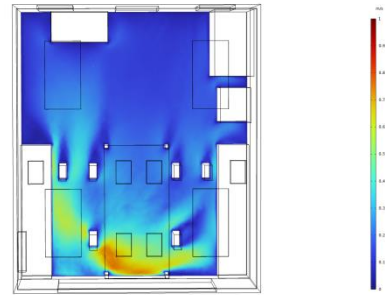
Fig. 15 The velocity field at z=0.2m
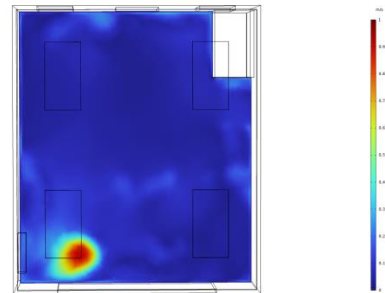
Fig. 16 The velocity field at z=1.75m

Fig. 17 shows the 3D diagram of the room, where the green line represents the path of the wind sent by the air conditioner, and the red arrow represents the direction and magnitude of the wind speed. The figure shows that the air provided by the air conditioner can cover the whole room. In addition to the position of the air outlet of the air conditioner, the wind speed is also moderate, which can meet the comfort of indoor personnel.
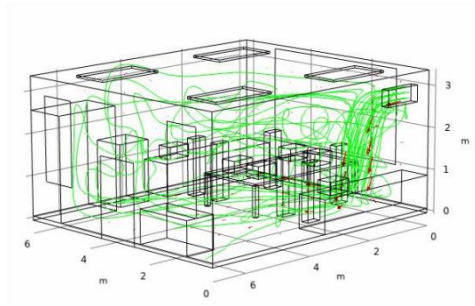
Fig. 17 3D air supply path diagram

Mr. Qikun Wang
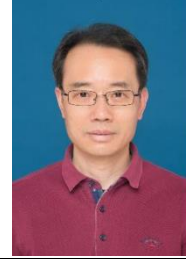
He graduated from Tianjin University of Science and Technology, is a senior engineer, research direction is HVAC, building technology.

4. Conclusion

Based on the above simulation analysis, the air provided by the air conditioner can cover the whole room, and the air supply by the air conditioner can meet the comfort of indoor personnel.

References

1. Jin, X., He, Y., Liu, S., Liu, Z., & Liu, P. (2022). Analysis of Distribution Characteristics and Influencing Factors of Ambient Temperature Field in Buildings. E3S Web of Conferences.
2. Khatri R, Singh A P, Khare V R. Identification of Ideal Air Temperature Distribution using different location for Air Conditioner in a room integrated with EATHE – A CFD based approach. Energy Procedia, 2017, 109: 11-17.

Authors Introduction

Ms. Peng Wang

She is a postgraduate tutor of Tianjin University of Science and Technology. In 2014, she received a doctorate from North China Electric Power University. The research direction is the functional safety assessment of safety instrumented systems.
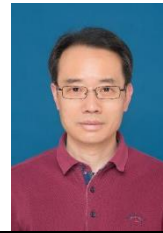
Mr. Mengda Liu

In 2023, he received his Bachelor of Engineering degree from the School of Electronic Information and Automation, Tianjin University of Science and Technology, China.He is pursuing a master's degree in engineering from Tianjin University of Science and Technology.

Solo Wheel Technology-Self-balancing Wheelbarrow

Ziyue Xiao*, Yumei Huang, Zhencheng Chang, Mingxuan Li

Tianjin University of Science and Technology, Tianjin, China

E-mail: *1418249500@qq.com

Abstract

As a new type of personal transportation, self-balancing unicycle has attracted wide attention with its unique design and advanced control system. We aim to study and optimize the self-balancing performance of the unicycle to improve its stability and maneuverability. Firstly, we gain an in-depth understanding of its operation principle. On this basis, an advanced control algorithm is proposed to realize real-time attitude adjustment. Second, for applications in complex environments, we propose an intelligent sensing system to enhance its environment sensing capability.

Keywords: Cascade PID, Angular feedback, MPU6050, Motor driver IC 33886.

1. Introduction

Since they all end up controlling the same control object (the motor of the car model), there is a coupling between them. For the sake of analysis, one of them is assumed to be stable with the other control objects. For example, when controlling the speed, the car model needs to be able to maintain upright control; When controlling the direction, the car model needs to be able to maintain balance and constant speed; Similarly, when it comes to balance control, speed and direction control are also required to be smooth. Of these three tasks, keeping the model balanced is key. Since the car model is affected by three controls at the same time, from the perspective of the car model balance control, the other two controls become its interference. Therefore, the control of the speed and direction of the car model should be kept as smooth as possible to reduce the interference of balance control. Taking speed adjustment as an example, it is necessary to change the inclination angle setting value of the car model in the balance control of the car model, so as to change the actual inclination angle of the car model. In order not to affect the balance control of the model, the change of the inclination of the model needs to be carried out very slowly. This will be discussed in more detail later in Speed Control.

The intuitive experience of controlling the balance of the car model comes from people's daily life experience. The average person can keep a straight wooden stick upright on the tip of a finger with simple practice. This requires two conditions: one is that the palm of the hand holding the wooden stick can be moved; The other is that the eye can observe the tilt angle and inclination tendency (angular velocity) of the stick. Counteract the angle and tendency of the stick by moving the palm of your hand, thus keeping the stick upright. These two conditions are indispensable, and in fact they are the negative feedback mechanism in the control [1], see Fig. 1.
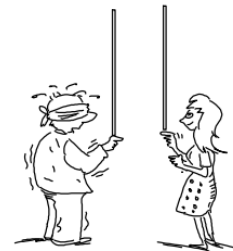
Fig. 1 Feedback control to keep the stick upright

Balance control is also achieved through negative feedback, which is relatively simple compared to keeping the stick upright above. Because the model has two wheels on the ground, the body will only tilt in the direction of the wheels. Controlling the rotation of the wheels to counteract the tendency to tilt in one dimension keeps the car balanced. The gravitational field, which uses thin lines to suspend heavy objects, is simplified to form an idealized pendulum model. An upright model can be seen as an upside-down pendulum placed on a

platform that can be moved from side to side [2]. This is shown in Fig. 2.
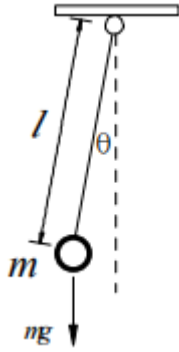
Fig. 2 Simple pendulum

In the case of a small offset angle, the restoring force is proportional to the magnitude of the offset angle and is in opposite directions. Under the action of this restoring force, a single pendulum undergoes a periodic motion. A pendulum that moves in the air will eventually stop in a vertically balanced position due to the damping force of the air. The damping force of the air is directly proportional to the speed of the single pendulum motion and is in the opposite direction. The greater the damping force, the faster the pendulum will stabilize in the vertical position.

2. Hardware Usage Programme

2.1 Motor driver IC 33886

The DSC F8013 in Fig. 3 is a 3.3V device, and its IO output voltage is up to 3.3V, which does not meet the requirement that the 33886 must be greater than 3.5V, so a 5V power supply is specially designed in the circuit to pull the drive signal of the 33886 up to 5V. Since the IO port of the F8013 can tolerate 5V, the above circuit can make the 33886 drive signal voltage reach 5V.
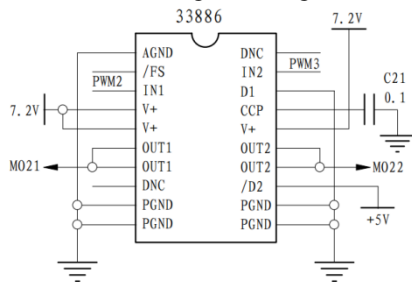
Fig. 3 Dual-motor drive circuit

In order to improve the application efficiency of the power supply, the PWM waveform of the drive motor adopts a unipolar drive method. That is, the voltage applied to the motor is a voltage in a PWM cycle.

Therefore, in order for each motor to be able to achieve forward and reverse rotation, two PWM signals are required. The two motors require a total of 4 PWM signals. The drive circuit is shown in Fig. 4.
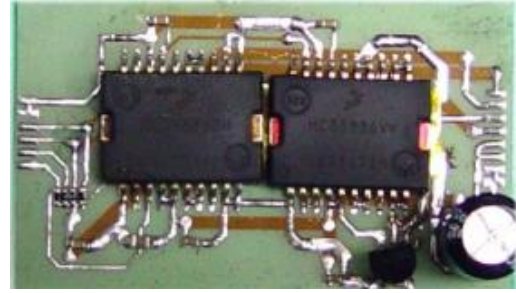
Fig. 4 Motor drive circuits

2.2 Surveillance camera

Each photocell outputs two pulsed signals, which have the same waveform, but are $90°$ out of phase. If the motor is forward, the second pulse is $90°$ behind, and if the motor is reversed, the second pulse is $90°$ ahead. This relationship can be used to determine whether the motor is rotating forward or reverse. In the actual circuit, only one pulse signal is detected, and the speed of the motor is measured by its frequency. The steering of the motor is judged by the positive or negative voltage applied to the motor. Although due to the inertia of the car model, it is possible that in the case of low motor speed, the direction of rotation of the motor and the voltage applied to it may be different, resulting in inaccurate measurement of motor speed. Due to the racing during the race, the motor is generally running at high speed, and the direction of the motor is consistent with the direction of the voltage. Experiments show that this method can effectively determine the rotation direction of the motor and control the speed.

2.3 Surveillance camera

In the article on the control principle of the car model, there are two ways to measure the angle and angular velocity of the car model, the first is through the gyroscope and the acceleration sensor, and the second is through two acceleration sensors. The angular velocity is integrated with feedback to obtain a smooth angle signal that is consistent with that of the accelerometer. This part of the signal processing can be done by the microcontroller software, or it can be realized by external op amp circuits. In order to verify the performance of this circuit, the ENC-03 measurement signal mounted on the model car was also acquired in Fig. 5.
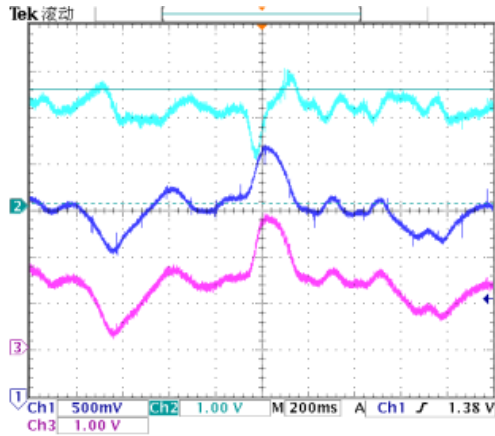
Fig. 5 Measurement angle waveform plot

From the waveform, it can be seen that the angular velocity of the car model integrated by the accelerometer is the same as the waveform of the angular velocity measured by the gyroscope, which verifies the correctness of the method.

3. Innovative Design and Advantages

3.1. *Vehicle tuning*

The debugging of the car model is divided into various links such as debugging preparation, static parameter tuning, dynamic parameter tuning, mechanical adjustment of the car model and competition strategy formulation. There are many parameters and components involved in the whole commissioning, and these parameters are closely influenced by each other. If you don't have a clear understanding of the physical process and whether there are correct debugging steps, then all kinds of errors in the debugging process will cover up the correct cause, affect the entire debugging process, and even shake the confidence of the production.

3.2. *MCU program download and communication*

The DSC 56F8013 has on-the-system programming (ISP) capability. After the MCU came out, there was a bootloader program inside, and the user program could be downloaded through the UART serial port. If the program itself is bootstrapped for the first time, all future program downloads can be done through UART. For more information, please refer to the materials available on the Freescale website. UART can not only improve the download of the program, but also display the serial characters sent during the operation of the program with the assistance of the upper serial port monitor, which is convenient for debugging the program. It is

recommended that the Bootloader download function of F8013 and serial port monitoring be jointly developed into a host computer software, so that the development and debugging of the entire car model program can be completed through the serial port.

3.3. *Analog acquisition*

The F8013 has six channels of AD conversion. Five of these channels are applied in the reference design. Write an AD conversion program and send the collected values to the upper computer for display. By observing the collected values of each channel, the operating status of each sensor, the dynamic range of the signal, and the influence of noise on each channel are confirmed. It is best to display the values collected by each channel on a curve, so that you can intuitively observe the changes in the signal and the amplitude of the noise. The following is the acquisition waveform of the gyroscope and accelerometer displayed on the upper microcomputer.

Each analog quantity will have a different proportion of noise, and low-pass filtering is required for each acquisition. The simple way to do this is to average it over multiple acquisitions. The waveform data shown in Fig. 6 is averaged over 20 acquisitions [3].
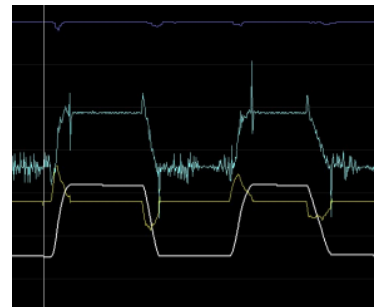
Fig. 6 AD acquires the signal

4. Job Scenario Application.

4.1. *Personal travel*

The wheelwheel can be used as a lightweight, portable means of transportation for short distances in the city. The wheels are fast and flexible, they can shuttle between the flow of people, and they can control the direction and speed of the wheels by tilting their bodies, which is very convenient and fast.

Ziyue Xiao, Yumei Huang, Zhencheng Chang, Mingxuan Li

4. 2. *Industrial applications and military applications*

The wheels can also be applied to industrial scenarios, such as warehouses, factories, etc. The uniwheel can control the direction and speed of travel by tilting the body, making it suitable for transportation and handling work in narrow, crowded environments. The uniwheel is also used in military fields, such as reconnaissance, surveying, etc. The single wheel can carry out fast and maneuverable action and detection in severe environments through its lightness and flexibility. For example, on the battlefield, a single wheel can be used for quick patrols, monitoring the situation, gathering intelligence, etc.

5. Synthesis

The unicycle is a balanced vehicle based on the principle of an inverted pendulum, consisting of a body with only one wheel and a balance control system. The movement and control of the uniwheel is realized by the user's body tilt, which is lightweight, flexible and fast, and is suitable for a variety of application scenarios such as personal travel, toys and entertainment, industrial applications, and military fields.

References

[1]. De Xiao Kong, A Yi Wang, Jun Hagi Li, Research on two-wheeled self-balancing vehicle control system based on quaternion, Electronic Measurement Technology, 2023.6, pp. 75-76.
[2]. T.S. Chen, S.Y. Chen, An engineering solution design based on STM32 two-wheeled balance bike, China Science and Technology Information, Vol. 20, No. 20, 2023, pp. 121-123.
[3]. Yuanwei Chen, Theoretical study and modelling of embedded-based self-balancing vehicle, Instrumentation Technology, Vol. 3, No. 2023, pp. 75-76.

Authors Introduction

Mr. Ziyue Xiao

He is hold a positive and serious attitude towards work, have a strong sense of responsibility, be sincere, meticulous, optimistic, and stable, and can continuously learn and improve oneself in practical work, and do their job well

Ms. Yumei Huang

She studied at Tianjin University of Science and Technology in 2021. She is now the first major in her faculty and is expected to receive her Bachelor's degree in 2025.

Mr. Zhencheng Chang

He is currently pursuing his undergraduate degree at the School of Electronic Information and Automation, Tianjin University of Science and Technology.

Ms. Mingxuan Li

Her major is environmental science. She is currently a undergraduate student in Tianjin University of Science and Technology

Second-order Self-balancing Inverted Pendulum

Ziyue Xiao*, Zhencheng Chang, Mingxuan Li, Yumei Huang

Tianjin University of Science and Technology, Tianjin, China

*E-mail: *1418249500@qq.com*

Abstract

The team utilizes the second-order inverted pendulum cart based on LQR controller for steady pendulum control with light rods. First, the second-order inverted pendulum is used as a research object to obtain its set of dynamical equations, then, the set of dynamical equations is written in the form of state-space expressions, and finally, the second-order inverted pendulum system of a balanced trolley with a light rod is controlled by a stabilized pendulum using the LQR controller. The inverted pendulum is a typical nonlinear, underdriven and unstable system, which can realize the all-round wind resistance of the rod and occupy less space with high stability.

Keywords：LQR controller，Second-order handstand pendulum，Lagrangian mechanics.

1. Introduction

In real life, many structural systems need to be stable in a variety of environments, such as high-rise buildings, bridges, tower structures, etc. One of the important factors is the influence of climatic conditions, especially wind on the building structure. Due to the uncertainty and variability of wind forces, buildings and structural systems need to be designed to be stable enough to resist the effects of wind. In practical application, in order to ensure the stability of the structural system, it is usually necessary to take various resistance measures, such as adding dampers, support devices and traction devices to the structural system, This is shown in Fig. 1.

Fig. 1. Telephone poles in kind

For slender structures such as rods, due to their special morphological and material properties, the study of their stability under wind force is more complex and critical. Therefore, in order to ensure the wind resistance and stability of the rod, it is necessary to carry out relevant research and application.

To this end, it is possible to study the dynamic characteristics of the rod under the action of wind, etc., the morphological and material characteristics of the rod, and the structural design of the rod to explore how to achieve the stability of the rod. At the same time, some practical application devices can also be developed to help the rod stabilize under the action of wind and other forces, such as adding support, traction and adjustment devices [1].

However, these methods are too traditional and require the support of manpower and material resources, and when the device is damaged, it is not easy to repair, time-consuming and labor-intensive. Therefore, it is urgent to develop a new type of stable and non-falling rod device.

In view of the above shortcomings, our group proposed a second-order inverted pendulum system based on LQR controller to stabilize the pendulum control of the light bar. Our team first took the second-order inverted pendulum system of the balance trolley with light bar as the research object, analyzed its dynamics based on the Lagrangian mechanics

method, obtained its dynamic equations, and then wrote the dynamic equations into the form of state space expressions, and finally, after a brief overview of the LQR controller, the LQR controller was used to stabilize the second-order inverted pendulum system of the balance trolley with light bar. The inverted pendulum is a typical nonlinear, underdriven, and unstable system. Through the research and optimization of the inverted pendulum system, our team has obtained new control algorithms, which have application prospects in robot control, missile interception control, aviation docking control and general industry.

2. Hardware Usage Programme

2.1 Mpu6050

MPU6050 is a commonly used six-axis inertial measurement unit (IMU) consisting of a gyroscope and an accelerometer. It is capable of measuring changes in the rotation and acceleration of an object in three axes.

MPU6050 has a built-in 3-axis gyroscope and a 3-axis accelerometer that can communicate with the microcontroller via an I2C or SPI interface. It is widely used in robots, drones, game controllers, and other applications that require gesture perception or motion tracking [2].

With MPU6050, you can obtain the angular velocity and acceleration data of an object, which can be used to calculate the object's pose or perform motion tracking. This is shown in Fig. 2,
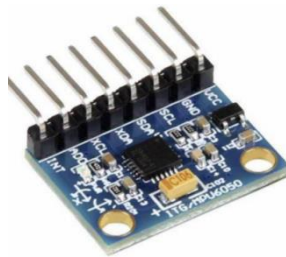
Fig, 2 MPU6050

2.2 STM32C8T6

STM32C8T6 is a 32-bit ARM Cortex-M3 core microcontroller from STMicroelectronics. It has high performance, low power consumption, a wide range of peripheral interfaces and rich development resources, and is one of the commonly used MCUs in embedded system development.

STM32C8T6 has expandable memory and peripheral interfaces, including high-speed USB, CAN bus, multiple serial communication interfaces, analog-to-digital converters and digital-to-analog converters, etc., and has a wide range of applications in industrial control, automotive electronics, home appliance control, medical equipment, smart home and other fields.

In addition, STM32C8T6 has a wealth of development resources, including development boards, online IDE tools, official documentation, and community support, which facilitate engineers to quickly develop and test embedded system applications.

3. System programming

3.1 LQR control

After the above qualitative description of the LQR controller, the simulation calculation was carried out by the relevant functions of Matlab (the mass of the body is 0.9 kg, the height is 0.126 m, the mass of the rod is 0.1 kg, and the length is 0.390 m [3], which is shown in Fig. 3.
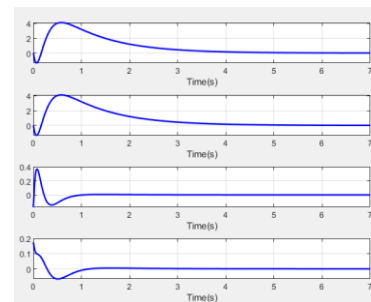
Fig. 3 MATLAB simulation results

The device was then simulated with Simulink, and the results are shown in Fig. 4.
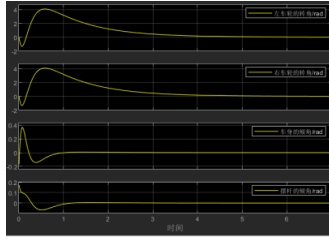
Fig. 4 Simulink simulation results

It can be seen from Fig. 2 that under the action of wind, the rotation angle of the wheel of the second-order inverted pendulum system with light rod reaches the maximum at 0.7s, and the rotation angle of the wheel reaches the maximum at 0.05s, and the second-order inverted pendulum system with light rod gradually tends to stabilize at 2s and basically achieves the stable effect at 3s [4].

3.2 Wind simulation

Wind Simulation We use Fluid Dynamics (CFD) simulation, which is a method of simulating fluid flow and force on objects through numerical calculations. When simulating the action of wind, the flow of the wind field can be simulated by establishing a mathematical horizontal shape, and the object to be tested can be added to the horizontal simulation, and the wind force on the object can be calculated through numerical calculation.

We perform a CFD simulation of a second-order inverted pendulum model with a light bar, and the results are shown in Fig. 5.
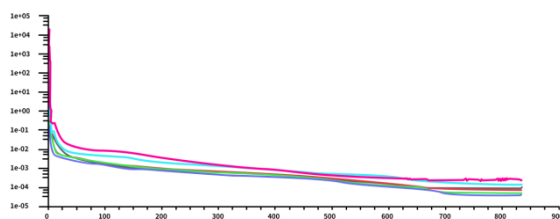
Fig, 5 CFD simulation

Wind power is the key to test the stability of the second-order inverted pendulum trolley with light bar, and the control variable method is used to test the wind power of the second-order inverted pendulum trolley with light bar. In physics, the method of controlling factors is often used to control multi-factor problems, turning multi-factor problems into multi-factor problems. Only one of the factors is changed at a time, and the remaining factors are controlled

unchanged, so as to study the influence of the changed factor on things, study them separately, and finally solve them comprehensively, this method is called the control variable method. It is an important method of thought in scientific inquiry, and is widely used in various scientific explorations and scientific experiments.

4. Introduction of Functional Module

4.1 Simulink emulation

For the test of the second-order inverted pendulum, a combination of simulation and reality is used. First, the system was simulated using Simulink to obtain the state data of the device. It is then combined with the real thing to verify the correctness of the simulation data. The controllable FS-75 industrial fan was used to perform wind tests on the second-order inverted pendulum system with light rods, and the device data was measured by changing the length, weight, and material of the rods. Through continuous simulation and testing, the efficiency of simulation testing is improved and the trial and error time is reduced. Extensive testing to ensure the correctness of the simulation and the ability to improve the simulated model in real time complement each other for continuous improvement.

4.2 FS-75 Industrial Fan Test

Experiments with a controllable FS-75 industrial fan and anemometer allow for better control of single variables for experimental measurements. After the experimental model is established, the system can be simulated and measured by Simulink rods of different lengths, weights, and materials, which greatly reduces the experimental time and improves the accuracy of the experimental data, which is shown in Fig. 6.

Fig. 6 FS-75

Ziyue Xiao, Zhencheng Chang, Mingxuan Li, Yumei Huang

5. Synthesis

In this experiment, the second-order inverted pendulum system was used to carry out the non-inverted experiment on the light rod, and the second-order inverted pendulum system with light rod was simulated by Simulink, and the system gradually stabilized in 2s. Then, the second-order inverted pendulum system with different rod lengths, different system states and different materials was simulated respectively, and the data state of the optimal second-order inverted pendulum system and device with light rod was obtained, and the influence of using different controllers on the stability of the device was discussed, and the time error was calculated to obtain the best controller. Then, combined with the real thing, the wind test of the second-order inverted pendulum system with light rod was carried out with the controllable FS-75 industrial fan and the stability test of the device was carried out by changing the length of the light rod, and it can be concluded that the experimental device has high stability, and the time to reach the stable state at the wind speed of $1 \sim 3$ m/s is 0.72s, 0.91s and 1.33s respectively, and the length of the rod is increased within a certain range, and the system can achieve stability, which meets the requirements of physical experiments [5].

In this experimental setup, the second-order inverted pendulum handstand model was studied, the optimal handstand state was discussed under different rod lengths and materials, the stability of the system was simulated under different external disturbances (wind speeds), and the influence of different controllers on the stability of the device and the time error were discussed. In the future, it is planned to change the length of the trolley (lower rod part) in the experimental device and simulate the measurement. Multiple fans were used to test the device to discuss whether the device could be more stable in more complex situations.

6. References

[1]. Hui-Hui Zhang, B.J. Hou, J. Gao, et, al, Research and implementation of stable pendulum control for linear inverted pendulum based on LQR, *Mechanical Design and Manufacturing*, 2023.7, pp. 1-6

[2]. Jiang, Lingyun,Xu, Bingji,Zhang, Fenghua, et, al, A study on a stabilisation control algorithm for linear two-stage inverted pendulums, *Computer Simulation*, 2023.Vol. 40, pp. 314-320

[3]. Hui-Wen Shi, Jie Zheng, MATLAB and ADAMS simulation study of linear inverted pendulum based on model predictive control, *Industrial Technology and Vocational Education*, 2023 XXI, pp. 1-5

[4]. Liu, Jiawei, Song, Nan, et, al, Optimisation and real-time simulation of inverted pendulum control system, *Chemical Automation and Instrumentation,* Vol. 50, 2023, pp. 453-458

[5]. Ji Hong, Wang Yingli, Zhao Zhongyi, et, al, Mathematical modelling and system performance analysis based on a two-stage inverted pendulum, Mechanical Design*, Vol.* 2023, No. 40, pp. 137-142

Authors Introduction

Mr. Ziyue Xiao

He is hold a positive and serious attitude towards work, have a strong sense of responsibility, be sincere, meticulous, optimistic, and stable, and can continuously learn and improve oneself in practical wor

Mr. Zhencheng Chang

He is currently pursuing his undergraduate degree at the School of Electronic Information and Automation, Tianjin University of Science and Technology.

Ms. Mingxuan Li

Her major is environmental science. She is currently a undergra student in Tianjin University of Science and Technology.

Ms. Yumei Huang

She studied at Tianjin University of Science and Technology in 2021. She is now the first major in her faculty and is expected to receive her Bachelor's degree in 2025.

Design of Nanny's Abnormal Behavior Recognition Bracelet Based on Human Activity Recognition(HAR) Deep Learning Model

Dengpeng Wang*, Yingfan Zhu, Yande Xiang, Ziyue Xaio
Tianjin University of Science and Technology, Tianjin, China

Email: 3264418107@qq.com

Abstract

This article designs a nanny abnormal behavior recognition bracelet. The bracelet is equipped with multiple sensors and a powerful control board, forming a complete nanny abnormal behavior recognition system, which realizes nanny abnormal behavior recognition and alarm in various environments. It uses a gyroscope to collect the three-axis acceleration information of the nanny, and uses the HAR model to infer the nanny's real-time behavior. When the abnormal behavior of the nanny is inferred, STM32 sends the GPS collected positioning information to the employer through ESP32 for timely alarm.

Keywords: Behavior recognition, Nanny, HAR model, Alarm

1. Introduction

With the increasing aging situation in China and the gradual implementation of the two-child and three-child policy, the number of elderly and young people in China has soared, and more and more elderly and children need to be taken care of, and the market scale of the nanny industry has expanded year by year. However, it is difficult to regulate the behavior of nannies, and the incidents of nannies abusing the caretakers are frequent. At present, employers can only monitor nannies by installing surveillance cameras, but there are large dead corners where nannies can still commit violent acts against the caretakers.In particular to dynamics analysis with moving bodies, the selection of the numerical integration method is crucial for the realization of the actual dynamics occurring in the real world. It cannot be solved by a simple way to chip the time step of the integration in the explicit numerical method. In the implicit numerical integration, it will help to refine the time step adaptively.

A nanny's abnormal behavior recognition bracelet based on HAR deep learning model was studied to solve such pain points. The nanny needs to wear the bracelet anytime and anywhere, and the employer can check the nanny's behavior in real time through the mobile phone. When the nanny's behavior to the caregiver is identified, the location of the nanny is directly sent to the employer so that the employer can timely understand the situation and report to the police.

The rest of this paper is organized as follows: The second part introduces the construction and deployment of HAR model. The third part introduces the hardware selection. The fourth part introduces the software design of the system. The fifth part summarizes the main content of this paper.

2. Construction and deployment of HAR model

2.1. *HAR Model Introduction*

Activity recognition has recently gained attention as a research topic because of the increasing availability of accelerometers in consumer products, like cell phones, and because of the many potential applications [1]. HAR model is an efficient human behavior recognition model, which aims to accurately monitor complex human behavior with limited hardware resources. The model uses a convolutional neural network (CNN [2]) architecture, a deep learning architecture that has achieved remarkable success in the field of image recognition and is particularly suitable for processing multidimensional time series data.

The reasoning principle of the model is that when the human body performs different actions, the three-axis acceleration curves of Roll, Pitch and Yaw are significantly different. Fig. 1 and Fig. 2 show the three-axis acceleration curves of the human body while sitting and jogging respectively, and it can be seen that the two curves are significantly different. The model can judge the specific behavior of the human body by the triaxial acceleration when performing different actions.

Dengpeng Wang, Yingfan Zhu, Yande Xiang, Ziyue Xaio
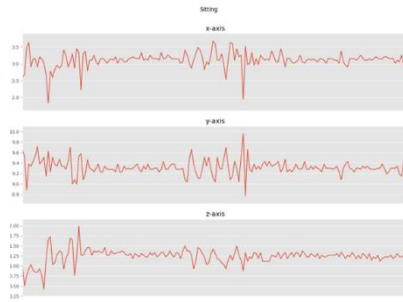
Fig.1 Three axis acceleration curve of human body during sitting still.
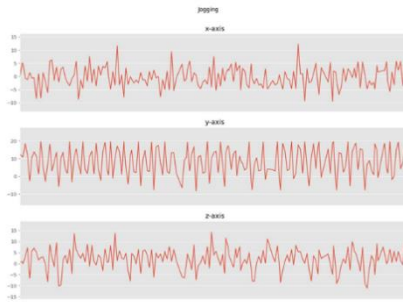
Fig.2 Triaxial Acceleration Curve of the Human Body During Jogging.

2.2. *Construction of HAR model*

In the process of model construction, we fully consider the advantages of convolutional neural network in time series data processing. The convolutional layer can effectively capture local patterns in the data, while the maximum pooling layer helps to reduce the sensitivity of the model to the input data and improve the robustness of the model. The fully connected layer is responsible for integrating these local information to form a holistic understanding of different behaviors. This hierarchical structure allows the model to better learn and understand complex behavioral features. The overall structure is shown in Fig. 3.
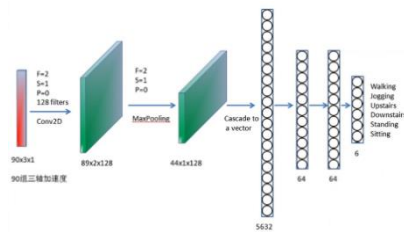
Fig. 3 Overall Network Structure of HAR Model

In the design of the network structure, the parameter configuration of each layer is paid attention to to ensure that the model can effectively capture the key features in the timing data. First, by setting the number of

convolutional nuclei to 90, the size to 3*3, and the step size to 1, the local pattern of the input data is effectively extracted, which helps to capture subtle changes in the behavioral data. The activation function of the convolutional layer is ReLU [3] to introduce nonlinear characteristics and enhance the expressiveness of the model.

In the maximum pooling layer of the second layer, a 2*2 pooling kernel is adopted and its step size is set to 2 to reduce the spatial dimension of the data and the sensitivity of the model to the input data. This step helps to improve the generalization performance of the model so that it can better adapt to changes in different samples.

The design of the third convolutional layer continues to follow the above principles, the number of convolutional cores is 44, the size is 3*3, the step size is 1, the zero padding is 0, and the activation function is ReLU. The setting of this layer further deepens the feature extraction of time series data, and provides more informative input for the subsequent processing of the full connection layer.

The next three fully connected layers play a key role in the overall structure. The 5632 nodes of the fourth fully connected layer are responsible for integrating the local information extracted by the convolutional layer to form a holistic understanding of different behaviors. Two fully connected layers with a number of 64 nodes further improve the abstraction capability of the model, enabling it to better capture the abstract characteristics of the behavior.

The final output layer is a fully connected layer with six nodes, each corresponding to a behavior category (walking, jogging, going up, going down, standing, sitting, etc.). This design ensures that the model can accurately classify different behaviors, providing a reliable basis for real-time monitoring.

2.3. *Deployment of the HAR model*

In order to deploy the model on a small-capacity embedded device, we first upgraded the HAR model iteratively for several times, and successfully compressed the size of the HAR model from the initial 387.43KB to 19.75KB through operations such as pruning, quantization and optimization of the algorithm structure. The model size optimization effect is shown in Fig. 4.

```
Total Flash:        396730 B  (387.43 KiB)
      Weights:      382488 B  (373.52 KiB)
      Library:      14242 B  (13.91 KiB)
Total Ram:          95028 B  (92.80 KiB)
   Activations:     92216 B  (90.05 KiB)
      Library:      2812 B  (2.75 KiB)
      Input:        1080 B  (1.05 KiB  included  in Activations)
      Output:       24 B  (included  in Activations)
Done
```

a. Before optimization

```
Total Flash:        20222 B  (19.75 KiB)
      Weights:      5824 B  (5.69 KiB)
      Library:      14398 B  (14.06 KiB)
Total Ram:          94744 B  (92.52 KiB)
   Activations:     92216 B  (90.05 KiB)
      Library:      2528 B  (2.47 KiB)
      Input:        1080 B  (1.05 KiB  included  in Activations)
      Output:       24 B  (included  in Activations)
Done
```

b. After optimization

Fig. 4 Size comparison of HAR model before and after optimization

After lightweight and efficient processing, HAR model not only has a smaller storage footprint, but also can run on low-resource devices. In the training and design phase of the model, we employ a series of efficient algorithms and neural network structures, including advanced techniques such as convolutional neural networks (CNN) and recurrent neural networks (RNN [4]). These choices greatly reduce the training time and prediction time of the model to ensure both accuracy and timely reasoning in real-time applications. The improved model has achieved remarkable results in more than 100,000 test sets, and the accuracy rate of the model has increased from 85.19% to 89.22%, as shown in Fig. 5.

```
x86 c-model #1      85.19%  0.1922276  0.0628854  0.5124106  -0.0000000  0.1922308  0.7339588
original model #1   85.21%  0.1922357  0.0628677  0.5123944   0.0000000  0.1922390  0.7339362
X-cross #1           99.86%  0.0011135  0.0003314  0.0029682  -0.0000000  0.0011135  0.9999890
```

a. Before optimization

```
x86 c-model #1      89.22%  0.1651415  0.0618931  0.4544591   0.0000000  0.1651443  0.8036504
original model #1   89.22%  0.1651415  0.0618931  0.4544591   0.0000000  0.1651443  0.8036504
X-cross #1          100.00%  0.0000001  0.0000000  0.0000002  -0.0000000  0.0000001  1.0000000
```

b. After optimization

Fig. 5 Comparison of accuracy before and after HAR model optimization

The deployment of the model is realized by STM32CubeMX tool, and the Cube AI component in the software can compress the model according to different magnifications and convert it into C language, so as to facilitate the deployment of the model to the embedded development equipment.

3. Hardware selection

3.1. *Main control chip*

The main control chip adopts STM32F407VET6 developed by ST Company. The chip adopts Cortex-M4 core, and the peripherals are mounted with multi-channel I2C, SPI, CAN and other communication interfaces. It has rich functions and its main frequency reaches

168MHz. With 512KB Flash and 192KB RAM, the powerful computing power and sufficient memory are enough to meet the real-time requirements of nanny abnormal behavior recognition bracelet behavior reasoning. The outline design of the chip is shown in Fig. 6.

Fig. 6 STM32F407VET6 chip shape design

3.2. *WIFI module*

This design uses ESP32-WROOM-32WIFI module, its appearance is shown in Fig. 7. The ESP32-WROOM-32 is a powerful universal Wi-Fi+BT+BLE MCU module with powerful features and a wide range of application areas. It can be used in low-power sensor networks and high-performance tasks such as speech coding, audio streaming, and MP3 decoding.

Fig. 7 WIFI module appearance diagram

The module combines high performance computing and low power consumption, which is very suitable for iot project development application scenarios. It uses asynchronous serial port to communicate with STM32 to receive data, acts as a Wi-Fi access point or client, accesses Aliyun iot platform to achieve wireless network connection, and supports wireless communication between iot devices.

3.3. *Gyroscope*

The gyroscope in this design uses MPU6050 attitude sensor, and its chip appearance is shown in Fig. 8. The MPU6050 is a 6-axis attitude sensor that integrates a 3-axis MEMS gyroscope, a 3-axis MEMS accelerometer, and a scalable Digital Motion Processor (DMP). The range of the accelerometer can reach ±16g, and there is a configurable digital low-pass filter inside, and the

configurable register can select the output data for low-pass filtering. The MPU6050 communicates with the STM32 via the IIC bus to transmit the nanny's three-axis acceleration.

Fig. 8 MPU6050 chip

3.4. *GPS Module*

The GPS module uses TAU1201 positioning module, and its module appearance is shown in Fig. 9. TAU1201 is a high-performance dual-band GNSS positioning module, equipped with BDA Beidou's CYNOSURE III GNSS SoC chip, the module supports the new generation of Beidou III signal system, and supports all civil navigation satellite systems in the world. These include BDS, GPS, GLONASS, Galileo, IRNSS, QZSS and SBAS (WAAS, EGNOS, GAGAN and MSAS). The TAU1201 integrates an efficient power management architecture to provide a high-precision, high-sensitivity, low-power solution for GNSS navigation applications.

Fig. 9 Appearance of the TAU1201 module

4. System software design

4.1. *Acquisition of triaxial acceleration*

The STM32 reads the contents of the relevant registers of the MPU6050 through the IIC communication protocol, a serial communication bus using a multi-master-slave architecture developed by Philips in the 1980s to allow motherboards, embedded systems, or mobile phones to connect low-speed peripherals. The IIC communication rate of STM32F407VET6 can reach up to 400KHz, which is enough to meet the requirements of HAR model inference.

Fig. 10 shows each meaning of the acceleration range register of the MPU6050 attitude sensor and the configuration guide. During configuration, the accelerometer range can be controlled by writing data to the register. Fig. 11 is an introduction of the acceleration read register of the MPU6050 attitude sensor. When the MPU6050 is used, the raw data of the three-axis acceleration is obtained through the 16-bit ADC value of the acceleration read register after the accelerometer RANGE is configured. The relationship between the three-axis acceleration (ACC) and the acceleration range (R) and the raw data (ADC) of the three-axis acceleration can be expressed as:

$$ACC = \frac{R * ADC}{32768} \tag{1}$$

Fig. 10 MPU6050 aceleration configuration register

Fig. 11 MPU6050 acceleration read register

4.2. *Nanny location acquisition*

The nanny location is obtained through the TAU1201 module. TAU1201 module can capture the signal of the satellite to be tested according to a certain height cutoff Angle of the satellite, track the operation of these satellites, transform, amplify and process the received GPS signal, so as to measure the propagation time of the GPS signal from the satellite to the receiver antenna, and interpret the navigation message sent by the GPS satellite. When the TAU1201 is finished interpreting, it communicates with STM32 through its own serial port to send the nanny location information to STM32.

4.3. *Behavioral inference data and location information upload*

Data upload is completed through the ESP32-Wroom-32wifi module. When the nanny's behavior is deduced on STM32, the HAR model inference result will be sent to ESP32 through STM32's own serial port. After the ESP32 makes the judgment of dangerous behavior, the judgment result will be sent to the employer's mobile phone server through MQTT protocol. When it is deduced that the nanny is in danger, the nanny's real-time location information is uploaded to the employer's mobile phone so that the employer can take the next step.

5. Conclusion

The nanny abnormal behavior recognition bracelet, to a large extent, solves the problems of traditional surveillance cameras, such as large monitoring dead Angle and not easy to carry, so that the nanny's behavior can be better supervised and standardized, has a good role in promoting the development of the nanny industry, and indirectly protects the life and property safety of the caretakers.

References

1. Jennifer R. Kwapisz, Gary M. Weiss, Samuel A. Moore, Activity Recognition using Cell Phone Accelerometers, ACM SIGKDD Explorations Newsletter.Volume 12,Issue 2.2011: pp.74-82.
2. Wang Qinghua;Li Ziwei;Zhang Shuqi;Chi Nan;Dai Qionghai,A versatile Wavelet-Enhanced CNN-Transformer for improved fluorescence microscopy image restoration,Neural Networks.Volume 170,Issue. 2024:pp.227-241.
3. He Zihao;Shu Qianyu;Wang Yinghua;Wen Jinming,A ReLU-based hard-thresholding algorithm for non-negative sparse signal recovery,Signal Processing.Volume 215,Issue.2024.
4. Xu Ziming;Leung Juliana Y,A novel formulation of RNN-based neural network with real-time updating – An application for dynamic hydraulic fractured shale gas production forecasting,Geoenergy Science and Engineering.Volume 233,Issue.2024.

Authors Introduction

Mr. DePeng Wang

He is studying at the School of Electronic Information and Automation at Tianjin University of Science and Technology, China.

Ms. YingFan Zhu

She is currently pursuing her undergraduate degree at the School of Electronic Information and Automation, Tianjin University of Science and Technology.

Mr. YanDe Xiang

He is currently pursuing his undergraduate degree at the School of Electronic Information and Automation, Tianjin University of Science and Technology.

Mr. ZiYue Xiao

He is currently pursuing his undergraduate degree at the School of Electronic Information and Automation, Tianjin University of Science and Technology.

Research on dynamic obstacle avoidance and complex path planning strategies based on ROS robots

Yi-Wei Chen

Bachelor's Degree Program in Intelligent Robotics, National Yunlin University of Science & Technology.123 University Road, Section 3, Douliou, Yunlin 64002, Taiwan, R.O.C

Jr-Hung Guo*

Department of Electrical Engineering, National Yunlin University of Science & Technology,123 University Road, Section 3, Douliou, Yunlin 64002, Taiwan, R.O.C

Kuo-Hsien Hsia

College of Future National Yunlin University of Science and Technology, 123 University Road, Section 3, Douliou, Yunlin 64002, Taiwan, R.O.C

Email: b10811150@yuntech.edu.tw, jrhung@yuntech.edu.tw, khhsia@yuntech.edu.tw

Abstract

Robot Operating System (ROS) is a software system framework used by many robot systems. Although ROS provides a good development environment and related frameworks, ROS is not suitable for public places such as restaurants because of the coming and going of people. Dynamic obstacle avoidance is often handled by stopping the robot, or when there are frequent environmental map changes or when sensors such as optical radar fail, the stop action is also used. However, this often causes path obstructions or delays in completing tasks. Therefore, this study attempts to use images, auxiliary sensors, and various path avoidance strategies to solve the problem of the robot stopping and waiting for the obstacles to disappear. The problem of rapid changes in map paths.

Keywords:Robot Operating System (ROS), Dynamic obstacle, Dynamic obstacle avoidance.

1. Introduction

Robot Operating System (ROS) [1] is a very commonly used development platform for developing and researching robot systems in recent years. Because this software uses open-source code and uses resources provided by many robot developers. Let this ROS system grow rapidly. Whether mobile robots can actually operate in actual environments is also a very important research topic for mobile robots. Because the mobile robot is in the actual environment, cannot avoidance dynamic or static obstacles. Or without the ability to re-plan the correct path, the mobile robot well cannot actually operate in a real environment. Of course, this also limits the application of mobile robots. Therefore, this study uses ROS as a research platform to focus on how mobile robots dodge dynamic obstacles in the environment. And re-chart the path to try to find a solution. This can reduce the chance of the mobile robot stopping when encountering obstacles and improve the efficiency of the mobile robot. It also enables mobile robots to be applied in real environments. In previous research, dynamic obstacle avoidance is an important research topic for many mobile robots. For example, CHEN, Chin S., et al. [2] use the ROS architecture to evaluate the path cost of autonomous mobile robots (AMR) for obstacles in the environment. Then choose the better path and let the mobile robot run. CHOI, Jaewan., et al. [3] use reinforcement learning to improve the reliability of dynamic window approach (DWA) [4] and timed elastic band (TEB) [5]. Make mobile robot obstacle avoidance more efficient.

It can be found from previous related research that many mobile robots mostly use small controllers such as Raspberry Pi [6] due to cost and volume. And it uses LiDAR Light Detection and Ranging (LiDAR) as the main sensor. Although such an architecture is lower in cost, it results in insufficient computing power. And because when detecting obstacles, lidar can only detect obstacles and measure distances on the plane that light or laser can scan. It often happens that when an obstacle is not in the LiDAR light scanning area, the obstacle is misjudged or cannot be detected. Therefore, the development and application of mobile robots are hindered. Therefore, this research will use the architecture of x86 computers. In addition to using LiDAR, it will also use imaging technology. Through this architecture, the mobile robot's computing power and obstacle detection can be faster and more accurate. Allowing mobile robots to truly overcome dynamic obstacles in the environment and enable practical applications.

2. System Architecture

The architecture diagram of the mobile robot is shown in Fig. 1. The robot is equipped with cameras, LiDAR and STM32 control boards, etc. It also has many various interfaces for expansion of different hardware. It can be widely used in actual or simulated various task development and applications. This robot architecture uses ROS as the

Yi-Wei Chen, Jr-Hung Guo, Kuo-Hsien Hsia

main core and adds modules such as targets following systems. The system can be adjusted according to different applications to provide a diverse robot development environment. At the same time, the development results can also be quickly transferred to different mobile robot platforms.
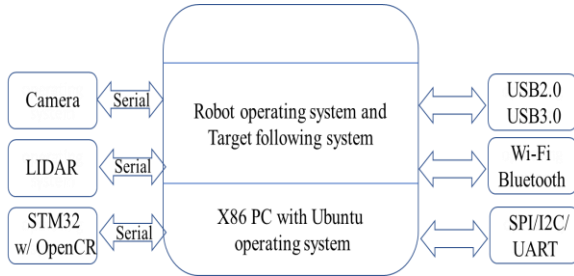
Fig.1 ROS mobile robot architecture block diagram.

In the ROS system, the sensors commonly used for object distance measurement are mainly LiDAR and imaging lenses. In order to consider the calculation amount and get the object distance, many robot platforms are based on LiDAR. However, general LiDAR is only 2D, that is, it can only measure one plane, and this also causes LiDAR to be unable to detect objects outside the 2D plane it scans. Therefore, in order to avoid obstacles that cannot be detected using LiDAR, this study uses images to detect obstacles and calculate distances.

There are two algorithms commonly used in ROS for object detection and distance calculation. One is KCF [7] (Kernelized Correlation Filter), which is a method to find the image features being tracked and then use this feature to track the object. Therefore, the operation speed of this method can be very fast. However, when obtaining object features and image tracking, this algorithm will lose the image data, making it impossible to perform accurate tracking. The other is DSST [8] (Discriminative Scale Space Tracker). DSST mainly improves the shortcomings of KCF that only uses gray scale to obtain object features and does not calculate the object scale. It makes the judgment of objects more accurate and enables distance estimation. The most important thing is that it will not have a big impact on the overall calculation amount. Although the above two algorithms are sufficient for tracking and following general mobile robots. However, because this research hopes to use images to process more complex, even more obstacles and more targets, the TLD [9] (Tracking-Learning-Detection) algorithm is also added. This algorithm can accurately identify the target, even when the object reappears after being obscured, it can still correctly identify the target. And if the target object rotates, TLD can also correctly identify it. Of course, the number of calculations required for this algorithm will be relatively large.

2.1 Research methods

In this study, in order to make ROS have a higher object recognition rate and calculate a more accurate object distance, we integrated KCF and DSST. As the main object

recognition and distance calculation core, the integrated operation structure diagram is shown in Fig. 2. In this processing block, KCF is mainly used to quickly identify targets and establish preliminary samples for subsequent identification. The main purpose is to identify possible obstacles or targets as quickly as possible, and then use DSST to further process the identification model and identification samples. Because KCF has a faster processing speed, it can first determine which objects may be closer to the mobile robot. Then, when necessary, it will be handed over to LiDAR and DSST for confirmation. Once a possible obstacle or path impact on the mobile robot is confirmed, it will continue to be tracked with KCF.
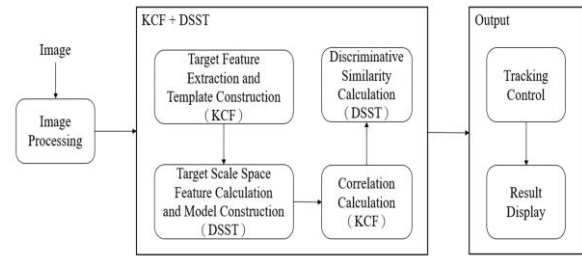
Fig.2 Block diagram of KCF+DSST target recognition system.

In the TLD algorithm part, because the preset requirements are relatively high for the computing power of the system, the calculation will only be performed when KCF+DSST determines that the object is closer to the mobile robot. Its block diagram is shown in Fig. 3. Moreover, the frequency of calculation can be set in the system, such as once per second, 10 times per second, etc. In addition, the maximum number of tracking can also be set. This prevents the CPU from consuming too much computing power in object tracking. Next, we will conduct experimental tests on the above methods.
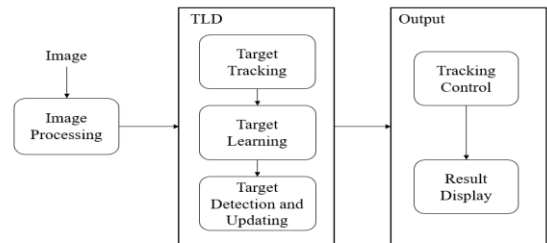
Fig.3 TLD target identification system block diagram.

3. Experimental Results

We use an x86 computer as the main control core, and the specifications used are: CPU, Intel® Processor N-series

N100, 8GB DDR4 DRAM, 512GB SSD. And directly install Ubuntu, and then install Gazebo and ROS in the system. First, we first confirm the object detection capabilities of the KCF+DSST and TLD algorithms. Fig. 4 is our test results on ROS. We use the same obstacles to identify objects and calculate scale distances using the KCF+DSST and TLD algorithms. According to the experimental results, both algorithms can operate normally. The operation speed of KCF+DSST is about 20% faster than the TLD algorithm. The calculation results of the scale distance between the two are almost the same.
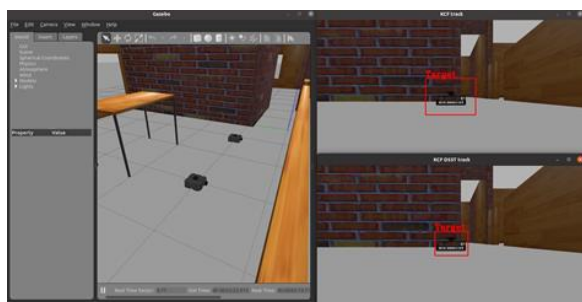
a. Objects move away from the robot.

b. Object approaches the robot
Fig.4 Object tracking and recognition results

But when we dynamically generate multiple objects with different sizes, distances, and positions. Although KCF+DSST can still find objects quickly. However, when objects are interlaced or have large displacement or rotation, KCF+DSST will lose the tracked object. As a result, the risk of collision of mobile robots becomes higher. When using TLD, this algorithm can quickly and correctly handle the interlacing of objects, as well as rapid displacement, rotation, etc., and can also correctly calculate the distance of objects. But when more than 150 objects appear, there will be a short pause based on the hardware we are using. Although it does not affect the operation of the system, there will be obvious pauses. KCF+DSST will only pause for a short time when there are about 300 objects.

4. Conclusions

This paper uses imaging technology to deal with the detection problem of multi-obstacle objects, and develops a mobile robot platform using an x86 CPU computer. The actual photo of this platform is shown in Fig. 5. We verified on ROS that KCF+DSST and the TLD method were used to verify the path re-planning that can be used to solve multi-dynamic obstacles. That is to say, obstacle avoidance can be performed when there are many dynamic obstacles in the environment. Of course, the current structure still cannot completely allow mobile robots to pause. However, based on the current experimental results, more than 90% of suspension situations can be avoided. The other 10% are almost all dynamic obstacles that completely block the path of the mobile robot. We will also try to solve this part in the future so that mobile robots can overcome the obstacles of dynamic emergence and movement in the environment and serve people in real environments.
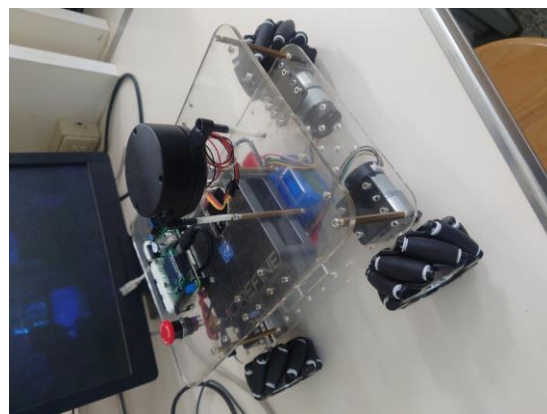
a. Mobile robot left view.
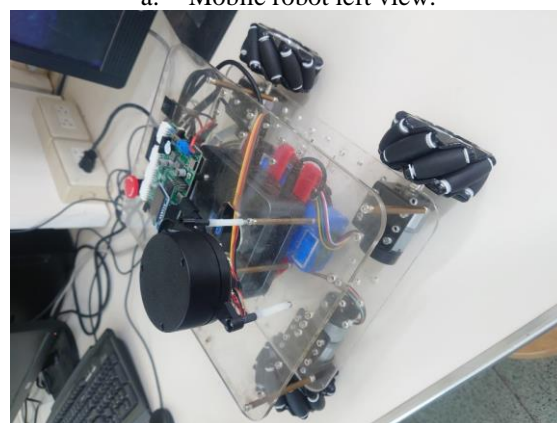
b. Mobile robot right view
Fig.5 Mobile robot platform

References

1. ROS Website, "https://www.ros.org/ ",2023/12/02"
2. CHEN, Chin S., et al. "Velocity estimation and cost map generation for dynamic obstacle avoidance of ROS based AMR.", Machines, 2022, 10.7: 501.
3. CHOI, Jaewan; LEE, Geonhee; LEE, Chibum. "Reinforcement learning-based dynamic obstacle avoidance and integration of path planning. Intelligent Service Robotics.", 2021, 14: 663-677.
4. OX, Dieter; BURGARD, Wolfram; THRUN, Sebastian. "The dynamic window approach to collision avoidance.", IEEE Robotics & Automation Magazine, 1997, 4.1: 23-33.
5. RÖSMANN, Christoph; HOFFMANN, Frank; BERTRAM, Torsten. "Timed-elastic-bands for time-optimal point-to-point nonlinear model predictive control.", In: 2015 european control conference (ECC). IEEE, 2015. p. 3352-3357.

6. Raspberry_Pi Wikipedia Website, "https://en.wikipedia.org/wiki/Raspberry_Pi ",2023/12/02
7. HENRIQUES, João F., et al. "High-speed tracking with kernelized correlation filters.", IEEE transactions on pattern analysis and machine intelligence, 2014, 37.3: 583-596.
8. DANELLJAN, Martin, et al. "Discriminative scale space tracking.", IEEE transactions on pattern analysis and machine intelligence, 2016, 39.8: 1561-1575.
9. KALAL, Zdenek; MIKOLAJCZYK, Krystian; MATAS, Jiri.,"Tracking-learning-detection.",IEEE transactions on pattern analysis and machine intelligence, 2011, 34.7: 1409-1422.

Authors Introduction

Mr. Yi-Wei Chen

Yi-Wei Chen, a graduate of National Yunlin University of Science and Technology, Taiwan, holds an engineering degree with a focus on robotics-based technologies. His academic pursuits are deeply rooted in the exploration and advancement of robotic applications.

Dr. Jr. Hung Guo

He is an Assistant Professor at the Department of Electrical Engineering, National Yunlin University of Science & Technology, Douliou, Taiwan. He received his Ph.D. Degree from National Yunlin University of Science & Technology, Taiwan in 2012. His research interests include sensor network, intelligent systems, intelligent robot. and technical and vocational education

Dr. Kuo-Hsien Hsia

He received the Ph.D. degree in electrical engineering from the National Sun Yat-Sen University, Taiwan, in 1994. He is currently an Associate Pofessor of National Yunlin University of Science and Technology in Taiwan. His research interests are in the area of mobile robotics, fuzzy control and image processing.

Research on Multi-Robot Formation on Two-Dimensional Plane

Kuo-Hsien Hsia

National Yunlin University of Science and Technology, 123 University Road, Section 3,Douliou, Yunlin 64002, Taiwan

Chun-Chi Lai

National Yunlin University of Science and Technology, 123 University Road, Section 3,Douliou, Yunlin 64002, Taiwan

Yi-Ting Liu

National Yunlin University of Science and Technology, 123 University Road, Section 3,Douliou, Yunlin 64002, Taiwan

Yu-Le Chen

National Yunlin University of Science and Technology, 123 University Road, Section 3,Douliou, Yunlin 64002, Taiwan
Email: {khhsia, cclai, M11112017}@yuntech.edu.tw, 1048576yc@gmail.com

Abstract

Mobile robots are playing an increasingly important role in both service and manufacturing industry. The management of multiple mobile robots is a very important issue on the research of mobile robotics. From a mathematical perspective, this paper discusses the problem of multiple robots on a two-dimensional plane reaching the designated positions in the shortest time to complete formation transformation. We improved the algorithm proposed by Hsia, Li and Su and proposed a new algorithm using a determinant and the Munkres assignment algorithm. Finally, the new algorithm is compared with the path distribution obtained by the Monte Carlo method under different numbers of robots and the excellence of the new algorithm has been verified.

Keywords: Multiple mobile robot, Pattern formation, Two-dimensional plane, Munkres assignment algorithm.

1. Introduction

In this society where time is money, saving time is an important factor that must be considered on task completion. How to complete tasks in the shortest time is what task assignment issues aim to achieve. Assignment problem [1] refers to the process of assigning specific work, tasks or responsibilities to specific people or teams. In an organization or team, assigning issues is a way of allocating work to ensure that matters are carried out efficiently and achieve desired goals. How to make an assignment that the tasks could be completed in the shortest time has always been the most important issue.

The shortest-path planning problem [2] is to find an assignment way for shortest-time running. Path planning refers to the movement path generated in order to enable one or more objects to complete a certain task in a designated area. In order to complete the task assignment problem in the shortest time, it is important to find a way to have the shortest path. The importance of the algorithm for this problem is that it can have the solution with higher efficiency, and the obtained result is repeatable.

This paper discusses the assignment problem of formation change of multiple robots. For a centralized robot system, there is a control center that makes decisions and assignments, while the other robots execute the received commands. In this paper, the multi-robot system is a centralized system and an algorithm is used to determine the destination that each robot is planning to move to. There are many types of formations in discussion, such as wild goose formation, front arrow formation, T-shaped formation, etc. The robots are asked to transfer their formation from one to another. The robots could only move laterally or vertically on a plane, so the robot is limited to a two-dimensional grid plane. In order to allow multiple robots to complete the formation in the shortest time, an algorithm can be used to find the assignment way that makes the formation transformation successfully in the shortest time.

2. Preliminary
2.1. *Problem description*

In the multi-robot formation transformation problem, there are n robots on a two-dimensional grid plane. The robots move horizontally or vertically from the starting positions to the target positions in the same speed. In order to find the fastest way to complete the formation change, we have to decide the target position for every robot and the robot with the longest distance will be assigned with a fastest way. This is a min-max problem.

Kuo-Hsien Hsia, Chun-Chi Lai, Yi-Ting Liu, Yu-Le Chen

2.2. *L1-norm between two points*

The L1-norm between two points is defined by the total distance of the projection of the line segment formed by each point on the axes on the Cartesian coordinate system, that is, the horizontal distance between the two points plus the longitudinal distance between them. Suppose the two points are at (x_1, y_1) and (x_2, y_2), their distance is $|x_1 - x_2| + |y_1 - y_2|$.

2.3. *Determinant*

To complete the assignment problem of *n* robots, each robot at his starting position will have a corresponding target point position. When all robots are assigned, their corresponding relationships will be presented in the distance matrix, and unassigned positions can be filled with 0. Under such circumstances, there will be *n* non-zero values in the distance matrix on different rows and columns, forming a square matrix of full rank. The fastest way to check for full rank is to check whether its determinant is 0 or not. Therefore, we use the determinant as a tool for checking the feasibility of formation transformation planning.

2.4. *Munkres assignment algorithm*

Munkres assignment algorithm [3] is an optimization algorithm for solving assignment problems. Given a cost or benefit matrix, it finds a set of assignments that minimize the total cost or maximize the total benefit, in which each row and column only can be assigned once.

2.5. *Monte Carlo method*

The Monte Carlo method is a numerical calculation method based on random sampling and statistical simulation [4]. When assigning tasks using the Monte Carlo method, one can randomly find an allocation way, and then find the maximum of the solution. One can get the number of steps for completing the tasks by using the Monte Carlo method one time. The more times of sampling, the more assignment ways there will be. By minimizing the maximum value of these, one can get a assignment way almost the most suitable.

3. Proposed new algorithm

3.1. *Deviation of the new algorithm*

We refer to the algorithm proposed by Hsia, Li and Su [5], but instead of finding the assignment way from the entries of the distance matrix, we find the minimum path from the positional relationship of the values in the matrix, and use the determinant to make the decision. The program searches sequentially from 0, and fills the searched position into another decision matrix. When a value is searched in the distance matrix, the

corresponding position in the decision matrix is filled with prime numbers from small to large, and the rest are not searched. The reached position is represented by 0 in the judgment matrix. When the search is completed, find the determinant of the judgement matrix. If the result is 0, add 1 to the search value and search again. Repeat the above steps until the determinant value is not 0, and then use the Munkres assignment algorithm to assign tasks.

3.2. *The algorithm*

Step 1: Calculate the L1-norm of all starting points to target points, and form the results into a matrix named distance matrix. The columns of the matrix correspond to the starting points, and the rows correspond to the target points.

Step 2: Initialize the judgement matrix of same dimension as distance matrix with all entries 0.

Step 3: Set the search value as 0.

Step 4: Fill in prime numbers in sequence to the corresponding position in the judgment matrix where the entry of the distance matrix is equal to the search value. The selection of prime numbers starts from 2, and each prime number filled in is the smallest prime number larger than the previous one.

Step 5: If the determinant of the decision matrix is not 0, jump to step 7.

Step 6: Increase the search value and back to step 4.

Step 7: Replace all positions of the distance matrix by a sufficiently large number, e.g. 999, if the corresponding position of the judgment matrix is 0.

Step 8: For a distance matrix, subtract the minimum value of that row from each row and subtract the minimum value of that column from each column. Thus at least one element in each row and each column will be zero.

Step 9: Use the least straight line to pass through all 0s, so that each 0 has at least one line passing through it.

Step 10: When the number of lines equals the number of robots, the position of 0 in the matrix can be used for task assignment. Starting from the row with the fewest 0s, only one 0 can be circled in each row and column. If the number of circles is equal to the number of robots, the assignment of the task is completed. If not, go to the next step.

Step 11: Subtract the minimum value in the undrawn row /column, and add the minimum value to the intersection position of the lines. After completion, back to step 9.

4. Comparison to Monte Carlo method

This comparison was performed on Matlab 2021, simulating the formation transformation situations of 10,

30, 50, 100, 500 and 1000 robots respectively. The starting and target points of the robots are randomly generated each time. Then the proposed algorithm and the Monte Carlo method are used for task assignment. The Monte Carlo method takes the best allocation method out of 10,000 times as the representative value. After comparison, the smaller one is the better assignment in the simulation. For each number of robots, we have 1000 simulations for both methods. This comparison index is to count the number of simulations that the proposed algorithm's result is not greater than the Monte Carlo method's result in 1000 simulations with different numbers of robots. The simulation results are summarized in Table 1. It is clear from Table 1 that the proposed algorithm can always have the same or better assignment way. Hence the proposed algorithm is better than Monte Carlo method for the formation transformation problem.

Table 1. Comparison results.

Number of robots	10	30	50	100	500	1000
Number of the results of proposed algorithm not worse than Monte Carlo method	1000	1000	1000	1000	1000	1000

5. Illustrated Examples

This paper takes three formation transformations as examples, namely the flying goose formation, the front vector formation and the T-shaped formation, and MATLAB is used for simulation. The robot is located on a 7x7 grid plane. The configuration of each formation is shown in Fig. 1. For the flying goose formation (V shape), the robots are at (4, 1), (4, 7), (5, 2), (5, 6), (6, 3), (6, 5), and (7, 4). For the front vector formation (Q shape), the robots are at (2, 4), (3, 4), (4, 4), (5, 4), (6, 3), (6, 5), and (7, 4). For the T-shaped formation (T shape), the robots are at (2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (3, 3), and (4, 3).

The 7 robots form the flying goose formation initially. We would like to make the robots changed to the front vector formation, and then changed to the T-shaped formation. The assignment results are summarized in Table 2, and the routes for the robots are shown in Fig. 2.

Table 2. The assignment results.

Pts of V-shape	(4, 1)	(4, 7)	(5, 2)	(5, 6)	(6, 3)	(6, 5)	(7, 4)
Pts of Q-shape	(2, 4)	(4, 4)	(3, 4)	(5, 4)	(6, 3)	(6, 5)	(7, 4)
Distance for #1	5	3	4	2	0	0	0
Pts of T-shape	(2, 1)	(2, 2)	(2, 3)	(2, 4)	(3, 3)	(2, 5)	(4, 3)
Distance for #2	3	4	2	3	3	4	4
Pts of V-shape	(4, 1)	(5, 2)	(6, 3)	(4, 7)	(5, 6)	(6, 5)	(7, 4)
Distance for #3	2	3	4	5	5	4	4

6. Conclusion

This paper proposed a new algorithm to solve the task assignment problem for a multi-robot system on formation transformation. Each mobile robot is limited to horizontal and vertical movement and starts at the same time. We'd like the transformation to be completed in the shortest time. The proposed algorithm can be used to systematically find out how to assign the target points of each robot so that the robot's formation can be transformed in the shortest time. Unlike previous works, we use determinant of a matrix to find out the shortest time required to complete the formation change, and use the Munkres assignment algorithm to obtain the shortest time. This work can ensure the best assignment result.
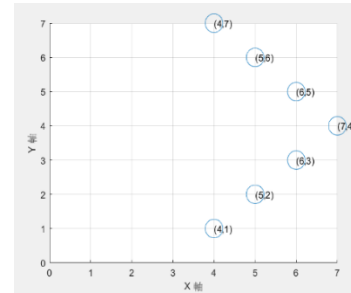
(a) flying goose formation
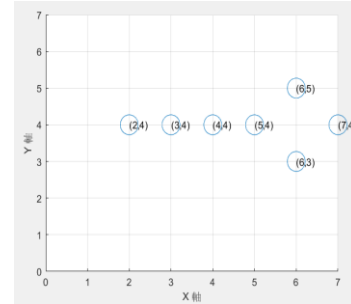
(b) front vector formation
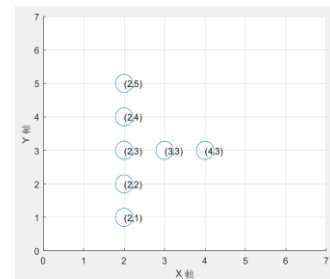
(c) T-shaped formation

Fig. 1. Formations considered in this paper

References

1. R. Burkard, M. Dell'Aamico and S. Martello, Assignment Problems, Society for Industrial and Applied Mathematics, 2012.
2. W. F. Chiang, Realization of the rectilinear shortest path planning problems, Master Thesis, National Taipei University, 2022.
3. F. Bourgeois and J. C. Lassalle, An extension of the Munkres Algorithm for the assignment problem to rectangular matrices, Communications of the ACM, 14(12), 802-804, 1971.
4. N. Metropolis and S. Ulam, The Monte Carlo Method, Journal of the American statistical association, 44(247), 335-341, 1949.
5. K. H. Hsia, B. Y. Li and K. L. Su, Task assignment on grid plane for multiple mobile robots, Tenth Int. Conf. on Innovative Computing, Information and Control (ICICIC2015) (Dalian, China, 2015), 193, 2015.
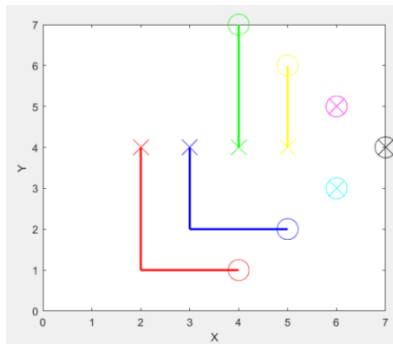
(a) flying goose formation -> front vector formation
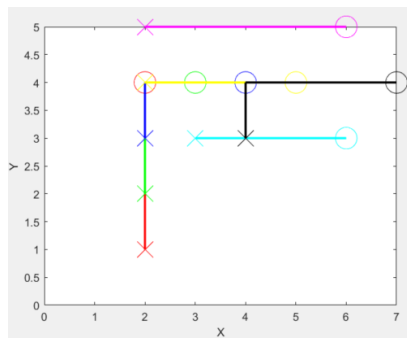
(b) front vector formation -> T-shaped formation
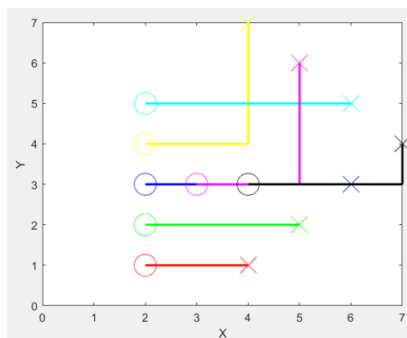
(c) T-shaped formation -> flying goose formation

Fig. 2 Formation transformation simulation

Authors Introduction

Dr. Kuo-Hsien Hsia

He received the Ph.D. degree in electrical engineering from National Sun Yat-Sen University, Taiwan. He is currently an Associate Pofessor of National Yunlin University of Science and Technology in Taiwan. His research interests are in the area of robotics, fuzzy systems, intelligent control. He is a member of IEEE.

Dr. Chun-Chi Lai

He received the Ph.D. degree in electrical engineering from National Chung Cheng University, Chiayi, Taiwan, in 2014. He is currently an Assistant Professor of National Yunlin University of Science and Technology in Taiwan. His research interests include multisensor fusion and intelligent robotics.

Mr. Yu-Le Chen

He received his master's degree in electrical engineering from National Yunlin University of Science and Technology, Taiwan in 2023. He is currently a student in the Government AI Career Training Institute, Taiwan.

Miss Yi-Ting Liu

She received her bachelor's degree in electrical engineering from National Chin-Yi University of Technology in Taiwan in 2022. Currently, she is a master's student at National Yunlin University of Science and Technology.

The Development of Utilization Rate and Energy Consumption Monitoring and Networking System for Old Machine

Chung-Wen Hung, Chun-Chieh Wang, Heng-En Chang
Nation Yunlin University of Science and Technology,
123 University Road, Section 3, Douliou, Yunlin 64002, Taiwan, R.O.C.

Chau-Chung Song
National Formosa University,
64 Wunhua Road, Huwei Township, Yunlin County 632, Taiwan, R.O.C.
Email: wenhung@yuntech.edu.tw, jasonccw@yuntech.edu.tw, m11212047@yuntech.edu.tw, ccsong@nfu.edu.tw

Abstract

In this paper, we present a system based on a microcontroller unit (MCU) for measuring the utilization rate of traditional, non-networked machinery. This equipment is designed for use with older machines equipped with andon lights. It employs optocoupler circuits to capture the status of these lights and current transformers to measure their operating and standby currents. Data is transmitted to the server using the Hypertext Transfer Protocol (HTTP) in JavaScript Object Notation (JSON) format. On the server side, a Hypertext Preprocessor (PHP) interprets the data, connects to a Structured Query Language (SQL) database, and stores the data using SQL commands. Users can access graphical data through a web-based interface, using it to refine production processes, reduce production costs, and minimize carbon emissions.

Keywords: Utilization, Non-networked machinery, Andon lights, Current transformers.

1. Introduction

With the widespread adoption of Industry 4.0 concepts and the concurrent advancement of relevant technologies such as networking and microchips, the Internet of Things (IoT) has emerged as a globally esteemed paradigm. The internet of machines refers to the ability of machines and devices to interact with the internet without human intervention. Information from devices, including operational progress, work duration, production history, and more, is transmitted via the Internet to servers for recording and analysis. The data facilitates analyses such as production process and utilization rate, allowing enterprises to adjust production processes based on this data. However, these actions are typically limited to new generations of IoT-enabled devices, as traditional machinery lacks such capabilities.

On the other hand, environmental protection has become a topic of global concern in recent years. Nations worldwide recognize the severity of greenhouse gas emissions and the importance of reducing carbon footprints. Governments have initiated carbon fees and tax policies, with the European Union passing a Carbon Border Tax set to take effect in 2026 [1]. This policy targets high carbon-emitting industries, imposing restrictions and charges. For enterprises, maintaining or reducing carbon emissions has become a matter of heightened concern. The determination of product carbon footprints and the subsequent implementation of carbon reduction labels are tasks to be addressed in the future.

This paper aims to assist factories in recording changes in andon light signals and corresponding time-related variations in electrical currents without intrusively altering traditional machinery's existing hardware and software. This is achieved by incorporating IoT devices to facilitate the documentation of machine operating states, thereby reducing downtime and enhancing overall operational efficiency. In addition, it also provides relevant information on the product's carbon footprint, aligning with the demands of Industry 4.0 and the implementation of carbon taxes.

2. Related Work
2.1. *Utilization*

The production or manufacturing process involves transforming raw materials or components into finished products. In the industrial context, production lines can be categorized into three types: automated production lines, semi-automated production lines, and manual production lines. The nature of a production line depends on the complexity of the finished product, production qunatity, and costs. Enterprises can plan and lay out production lines based on their requirements [2].

As highlighted by both [2], [3], inefficient working hours on the production line primarily stem from machine waiting times for personnel handling, waiting for material changes, troubleshooting, and maintenance. The efficiency metric used is referred to as "utilization rate." The utilization rate indicates the percentage of the total working hours where the machine is under load during effective working hours. Effective personnel management plays a crucial role in improving the factory's utilization rate.

2.2. *Industry 4.0 and the Internet of Things (IoT)*

The term Industry 4.0 emerged in Germany in 2011, referring to the technology used in the manufacturing process to monitor processes and employ data-based techniques for predicting, correcting, and adaptively adjusting production strategies [4]. Its objective is to integrate various data, tools, and processes to analyze and reduce costs, mitigate risks, and enhance efficiency [5], [6], [7]. Unlike previous industrial revolutions, Industry 4.0 is not based on a single technology. Instead, it is a convergence of modern technologies such as Big Data, the Internet of Things (IoT), Cloud Computing, and Artificial Intelligence (AI) [5].

The Internet of Things (IoT) applications have deeply penetrated our lives, with interconnected IoT devices sharing vast amounts of information without human intervention. According to [8], global IoT devices online in 2020 totaled 9.76 billion, with application areas illustrated in Fig. 1. Of these, 22% were applied in the industrial sector, 15% in transportation, and 14% in energy. By 2026, the number of online devices is projected to reach 21.09 billion; by 2030, it is estimated to approach nearly 30 billion devices.
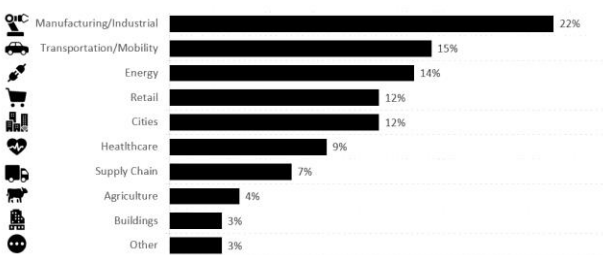
Fig. 1 The top 10 IoT application areas in 2020 [9]

This project aims to transform traditional machinery by further integrating it into the framework of Industry 4.0 to enhance machine utilization rates and production line efficiency. The referenced Internet of Machine Architecture model, as established in [10], [11], comprises five layers: the physical layer, sensor layer, connectivity layer, data layer, and application layer. The machinery transformation in line with Industry 4.0 primarily focuses on configuring the sensor, connectivity, and data layers. In the sensor layer, determining required machine states is crucial, involving establishing corresponding sensors and utilizing the production

information provided by the machine. The connectivity layer deals with how the collected information is transmitted to the server for processing. The data layer emphasizes planning the format of machine data and storing it in a database for subsequent applications and analysis.

2.3. *Carbon credit issues*

With global industrialization and increased energy consumption, there is a significant increase in carbon dioxide emissions, contributing to climate challenges. Addressing these issues requires international cooperation to reduce and control carbon emissions. The European Parliament has taken legislative action this year, entering a transitional phase from October. They are implementing the European Union Carbon Border Adjustment Mechanism (CBAM), which restricts the import of high-carbon products such as steel, cement, and electricity. It will be mandatory for imported products falling into supervised categories to purchase CBAM certificates [1] after 2026.

3. System Architecture

In this project, the collaborating partners are concerned about the machinery's operational status and energy consumption. However, due to the large number of machines, totaling 27 units with over ten different models, some of the actual machine appearances are depicted in Fig. 2. Embedded systems, coupled with sensors, will be utilized to extract information on machine status and energy consumption. This data will be wirelessly transmitted to a server for storage, facilitating subsequent analysis and processing.

Fig. 2 The actual appearance of the machinery

As illustrated in Fig. 3, the network architecture primarily consists of an MCU and a server. When the sensor reads values and transmits data to the MCU, the MCU packages the data in a predefined format. Subsequently, using the built-in Wi-Fi module, the data is transmitted to the server. Specific PHP (Hypertext Preprocessor) programs perform data analysis within the server and upload it to the MySQL database. Upon

completion of the upload, users can view all the data as of a webpage or export the data into reports through
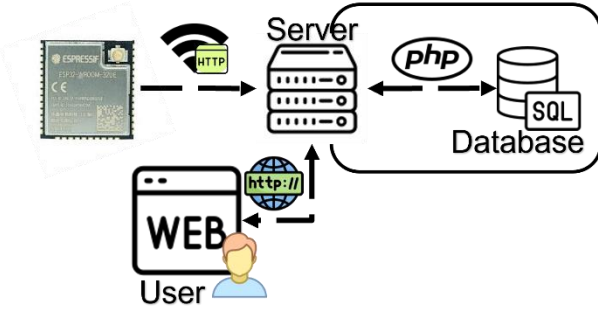
Fig. 3 The network architecture diagram

additional programming.

4. System Development
4.1. *Andon light status detection*

The core component of the Andon light status detection circuit is component U1 in Fig. 4, which is a optocoupler element. A optocoupler element is a device that uses light (such as visible light or infrared) as a medium to transmit electrical signals [12]. As previously mentioned, different machine models on the production line may use different electrical configurations for warning lights, distinguishing between AC and DC. Therefore, the PC814 bidirectional photocoupling element manufactured by Sharp, a Japanese electronics company, has been selected for development for this project. The main difference lies in the internal structure of the PC814, which includes a pair of light-emitting diodes — one forward and one reverse — enabling it to function in both directions of current flow. It elegantly resolved the compatibility issue between alternating current and direct current.
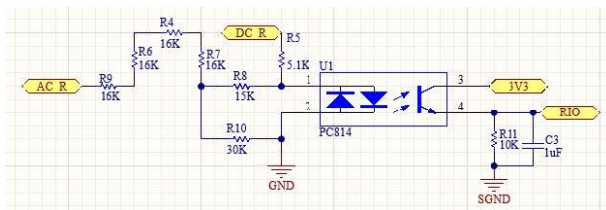
Fig. 4 Andon light status detection circuit diagram (detection part)

4.2. *Current Measurement*

This project utilized the SCT013 split-core current transformer manufactured by YHDC to avoid damaging the original power wiring. The circuit structure employed is illustrated in Fig. 5 with the calculation method for the burden resistor according to Eq. (1).

$$\text{Burden Resistor}(\Omega) = \frac{V_{AREF} * CT\ TURNS}{2\sqrt{2} * I_{pm}} \qquad (1)$$

Where V_{AREF} denotes analog reference voltage, CT TURNS represents the current transformer turns ratio,
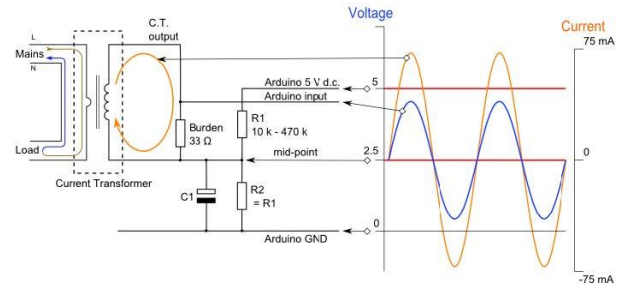
Fig. 5 Current measurement circuit diagram (measurement part) [13]

and I_{pm} denotes the maximum current on the primary side. By generating a voltage across the burden resistor through the secondary side current, the capacitor in Fig. 5 can accurately represent the direction and magnitude of the current. This information is acquired through the MCU's Analog-to-Digital Converter (ADC).

4.3. *Data format and transmission*

Upon receiving the status and data from peripheral devices, the MCU must analyze and upload the information to the server. The transmitted data should include details such as machine ID, data type, and timestamp. JSON (JavaScript Object Notation) can be employed as the data packaging format. JSON is a lightweight data interchange format supported by most programming languages, often with additional libraries for convenient packaging and parsing. The specific packaging format for this project is illustrated in Fig. 6.

Fig. 6 Data format

The data, transmitted through HTTP request by the MCU with Wi-Fi functionality, undergoes analysis on the server side using PHP. PHP also responds to the HTTP request and uploads the data to the database for storage. The server-side data transfer path is illustrated in Fig. 7.
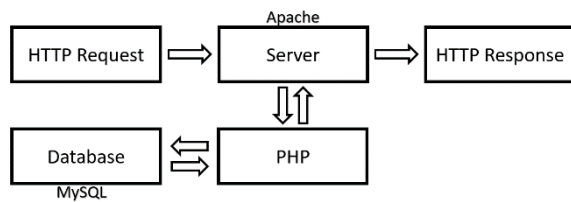
Fig. 7 The server-side data transfer path

5. Result

This paper uses the ESP32 series MCU, which Espressif Systems developed as the core for edge-side development. Its advantages lie in its cost-effectiveness and the integration of network modules with basic microprocessors. This MCU is commonly employed in the development of IoT applications.

After collecting data, present it to users in a graphical format for more straightforward interpretation and analysis. The system is developed as a Web App to avoid the need for different applications on various devices and the authentication thresholds of mobile applications. The advantages of a Web App lie in its accessibility—any internet-connected device with basic browser functionality can use it. The web page is coded using HTML (Hypertext Markup Language) for interface design. JavaScript is employed for real-time content updates and user interaction. PHP serves as an intermediary between JavaScript and the database for data requests.
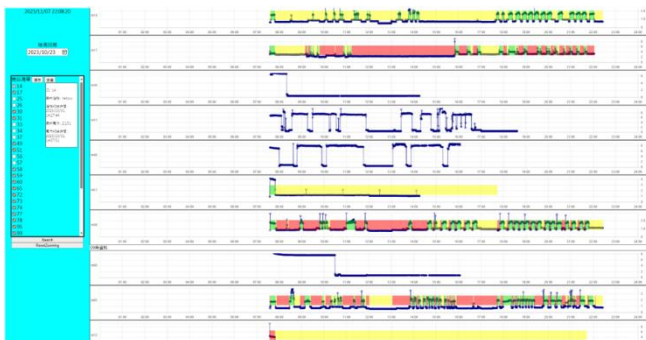
Fig. 8 Web App

Fig. 8 shows the interface displayed in the web app, divided into two main sections. Users can choose the machine number and time of interest on the left-hand side in the blue area. After selection, the data for the chosen machine will be displayed on the right-hand side in the white area. Each row in the chart on the right represents the data for a single machine on a given day. The various colored blocks represent the machine's status (indicator lights) during specific time intervals, while the lines depict the machine's current consumption. Analyzing this data assists factories in power consumption statistics, emission calculations, and formulating strategies for emission reduction. Aligns with national and international carbon policies, helping to minimize carbon-related costs and reduce unnecessary energy consumption during idle times.

6. Conclusion

This article utilizes a microcontroller with wireless connectivity, sensors, and electronic components to create a non-invasive device. Combined with a web app, factory administrators can swiftly gather production information to analyze and adjust production strategies.

References

[1] European Commission, Carbon Border Adjustment Mechanism, Taxation and Customs Union, https://taxation-customs.ec.europa.eu/carbon-border-adjustment-mechanism_en (accessed Nov. 15, 2023).

[2] SKAL Subramaniam, Siti Huzaimah Binti Husin, Yusmarnita Binti Yusop, and Abdul Hamid Bin Hamidon, Machine efficiency and man power utilization on production lines, in WSEAS International Conference. Proceedings. Mathematics and Computers in Science and Engineering, vol. 7, 2008, World Scientific and Engineering Academy and Society.

[3] Y. Y. Wang, J. Lionel, Case Study of the Machine Utilization on the Turning Process," in Industrial Science and Management Journal, Mar. 31, 2008.

[4] T. Ainsworth, J. Brake, P. Gonzalez, D. Toma, and A. F. Browne, A comprehensive survey of Industry 4.0, IIoT, and areas of implementation, SoutheastCon 2021*, pp. 1-6, 2021, IEEE.

[5] H. Khujamatov, E. Reypnazarov, D. Khasanov, and N. Akhmedov, IoT, IIoT, and cyber-physical systems integration, in Emergence of Cyber Physical System and IoT in Smart Automation and Robotics: Computer Engineering in Automation, Springer, 2021, pp. 31-50.

[6] K. Wójcicki, M. Biegańska, B. Paliwoda, and J. Górna, Internet of Things in Industry: Research Profiling, Application, Challenges and Opportunities—A Review, Energies, vol. 15, no. 5, p. 1806, 2022, MDPI.

[7] R. Y. Zhong, X. Xu, E. Klotz, and S. T. Newman, Intelligent manufacturing in the context of Industry 4.0: a review, Engineering, vol. 3, no. 5, pp. 616-630, 2017, Elsevier.

[8] L. S. Vailshery, IOT connected devices worldwide 2019-2030, Statista, https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/ (accessed Jun. 15, 2023).)

[9] M. H. Alsharif, A. Jahid, A. H. Kelechi, and R. Kannadasan, Green IoT: A Review and Future Research Directions, in Symmetry, 2023, vol. 15, p. 757.

[10] D. Jaspert, M. Ebel, A. Eckhardt, and J. Poeppelbuss, Smart retrofitting in manufacturing: A systematic review, Journal of Cleaner Production, vol. 312, p. 127555, 2021, Elsevier.

[11] P. Zheng, Z. Wang, C.-H. Chen, and L. P. Khoo, A survey of smart product-service systems: Key aspects, challenges and future perspectives, Advanced Engineering Informatics, vol. 42, p. 100973, 2019, Elsevier.

[12] Opto-Isolator, Wikipedia, https://en.wikipedia.org/wiki/Opto-isolator (accessed Dec. 4, 2023).

[13] CT sensors - interfacing with an arduino, CT Sensors - Interfacing with an Arduino - OpenEnergyMonitor 0.0.1 documentation, https://docs.openenergymonitor.org/electricity-monitoring/ct-sensors/interface-with-arduino.html (accessed Dec. 4, 2023).

Authors Introduction

Prof. Chung-Wen Hung

He received the PhD degree in Electrical Engineering from National Taiwan University in 2006. Currently, he is a professor at National Yunlin University of Science and Technology. His research interests include IoT, IIoT, power electronics, motor control, and AI application.

Chun-Chieh Wang

He received the PhD degree in The graduate school of engineering science and technology from Nation Yunlin University of Science and Technology. Currently, he is a project professor at National Yunlin University of Science and Technology.

Heng-En Chang

He received the B.S. degrees and now he is studying for the M.S. degree in electrical engineering from National Yunlin University of Science and Technology.

Chau-Chung Song

He received the Ph.D. degrees in Electrical Engineering from National chiao tung university. Currently he is a Distinguished professor and Chair in National Formosa University.

MCU Based Edge Computing Platform for Liquid Level Measurement

Chung-Wen Hung, Chun-Liang Liu, Tai-Hsuan Wu
Nation Yunlin University of Science and Technology, Taiwan
123 University Road, Section 3,Douliou, Yunlin 64002, Taiwan, R.O.C.

Chau-Chung Song
National Formosa University, Taiwan
64 Wenhua Road, Huwei Town, Yunlin County 632301, Taiwan, R.O.C.
Email: wenhung@yuntech.edu.tw, clliu@yuntech.edu.tw, m11112080@ yuntech.edu.tw, ccsong@nfu.edu.tw

Abstract

An edge computing system based on micro control unit (MCU) for liquid level measurement is proposed in this paper. The system includes a solenoid electromagnet for bottle hit and a microphone to capture sound waves. The signals are converted from time domain to frequency domain by Fast Fourier Transform (FFT), employing an artificial intelligence (AI) model to predict the water level. Artificial Neural Network (ANN) model is applied for classification on the MCU. When optimizing hyperparameters, the accuracy of each parameter combinations should be considered. Ensure the model size suits the limited MCU memory and computing capabilities. The experimental results can be up to 99% accurate under multiple tests.

Keywords: Artificial Neural Network, Hyperparameters, Edge Computing, Audio Process

1. Introduction

As artificial intelligence (AI) is more popular, there are many applications to be proposed no matter for the consumer or industrial market. Moreover, it is crucial to tackle the emerging challenges associated with edge computing [1], a topic of great significance and current prominence. The utilization of endpoints for model computation has attracted considerable attention, given its potential to reduce network latency and lower costs [2], [3]. Furthermore, the exploration of alternative AI models for measurement becomes pivotal in situations where sensor availability is constrained or cost prohibitive. When a container, such as bottle, experiences physical impact, it produces distinct tones depending on the liquid's varying levels inside, owing to vibration frequencies. Determining the water level through sound analysis, as presented in [4], relies on specific vessel geometries. However, employing an AI model or sound-based water level measurement becomes a viable option. Furthermore, given the limited resources of the endpoint platform, careful consideration of AI model selection and hyperparameter optimization for edge computing is crucial. Artificial neural networks (ANN) model, is simple yet effective for classification tasks, represent a viable candidate for this study [5], [6]. Moreover, Bayesian optimization emerges as a popular method for efficient hyperparameter optimization, offering convenient and efficient.

A micro control unit (MCU) based edge computing system for liquid level measurement is proposed in this paper, and all calculation and control including data acquisition are performed by a consumer MCU, Artery AT32F415 [7]. The proposed AI model utilizes one-dimensional sound information for classification, resulting in reduced memory requirements.

During the training stage, the proposed system offers a data sampling function, where the gathered information is transmitted to a personal computer (PC). The hyperparameters optimization and model training processes are conducted on the PC. Once completed, the trained model will be installed onto the MCU, enabling edge computing for measurement tasks performed by this MCU.

2. Methodology
2.1. SYSTEM STRUCTURE

The proposed system structure is shown in Fig. 1. Two micro switches are used to trigger procedure, which includes knocking, sampling, and edge computing. Moreover, these two switches also could make sure that the hitting situation, such as position and distance. MCU is the center of the system, and controls the full procedure, as mentioned previously, which also provide the edge computing. The UART (Universal Asynchronous Receiver/Transmitter) protocol is adopted for communication with PC, the sound sampling is sent to PC for training from MCU.
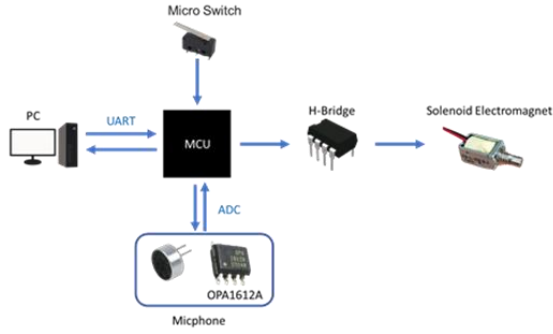
Fig. 1 The system diagram

A capacitive microphone is applied to convert sound to voltage, which amplitude is amplified and offset is shift by operational amplifier (OPA) [8] circuit for the input range of the Analog to Digital Converter (ADC), which is a built peripheral function of MCU. The knock component, a solenoid electromagnet, is driven by H-Bridge [9] to eject and attract.

2.2. DATA ANALYSIS AND REPROPCESSING

Before algorithm development, the data analysis is necessary. As mentioned previously, the converted sampling data are sent to PC for analysis and training via UART. The audio voltage is sampled at 10k Hz, and the number of sampling points is 2048 per trigger. The time domain waveforms of the different water levels are shown in Fig. 2 and there is not any saturation and truncation, it means that the sampling rate and length are suitable for this application. To develop the algorithm, the Fourier transform converts waveform data in the time domain into the frequency domain as shown in Fig. 3. Note, the frequency of the first spike, which decreases with increasing water volume, this phenomenon is consistent with the sound heard. Frequency information only from 83Hz to 1464.8Hz is adopted to identify water level. Furthermore, the frequency domain differences between different water levels are obvious, and the classification function could be performed by the artificial neural networks (ANN) model, which is simple and efficient for categorization. Compared to most other AI models, ANN requires less memory and calculation resource, in other words, it is suitable to be adopted into the simple edge-computing platform, MCU. So, Fourier transform (FFT) is selected for pre-processing, and ANN is applied to classify the water level.
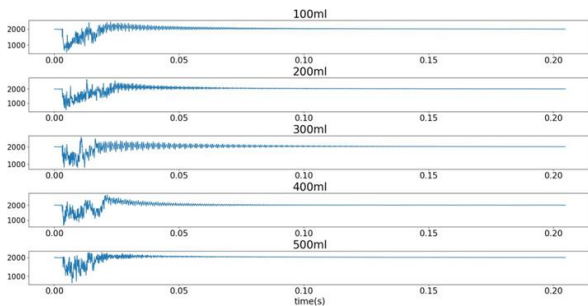
Fig. 2 Audio signals at different water level

In order to prevent the data from being affected by abnormal waveforms and to enhance the features, the system adopts Min-Max Normalization for each data. The Normalization finds the maximum and minimum values of the sound frequency, and then converts the maximum value is set to 1, and the minimum value is adjusted to 0. Finally, all audio data are scaled to the range between 0 and 1. The calculation formula is as (1) :

$$X_{nom} = \frac{X - X_{min}}{X_{max} - X_{min}} \qquad (1)$$

Here, X is the original data, Xmin and Xmax are the minimum and maximum value of the frequency domain data, and X_nom indicates the normalized data in frequency domain. The waveforms after data pre-processing are shown in the Fig. 3.
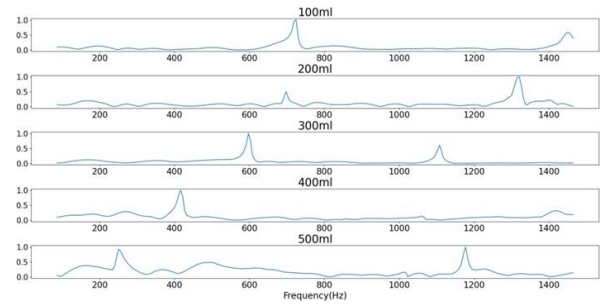
Fig. 3 Audio signals at different water level

2.3. HYPERPARAMETER OPTIMIZATION

Hyperparameter [10] is an important issue in deep learning algorithm as it determines the size and accuracy of AI model by selecting the number of hidden layers, the number of nodes and activation function in each layer. Typically, the hyperparameter combination that achieve the highest accuracy will be chosen to be implemented in the MCU. However, in this system, the implementation of edge computing is also a crucial consideration. Therefore, while ensuring the accuracy requirements are met, it is necessary to select a parameter combination that is suitable for the resource constraints.
Bayesian optimization [11] is a modelling method based on surrogate function that is used to optimizing the hyperparameter in the proposed system. It employs Gaussian process to model and updates the acquisition function based on the model's results, continually exploring the uncertain regions. As the number of searching increases, the surrogate function will be closer to the distribution of real function. Table 1 displays the results by hyperparameter search, sorted by the model size, number of hidden layers, number of nodes in each layer, and the corresponding model size and the accuracy rate Considering the memory limitation of the MCU, the first model which satisfies the accuracy requirements

while also relatively small memory requirements is selected as the training parameters for this system.

Table 1. HYPERPARAMETER BY BAYESIAN OPTIMIZATION RESULTS

Hidden Layer(s)	Node 1	Node 2	Node 3	Node 4	Size (kB)	acc
2	5	5	0	0	49.75	0.99
2	5	7	0	0	49.77	0.98
2	5	16	0	0	50.75	0.99
3	5	8	8	0	57.53	0.99
4	32	32	5	8	172.50	0.99
4	32	32	5	16	174.50	0.80

After confirming the hyperparameters, the collected data and hyperparameters are utilized to train the AI model. The MCU code includes a knocking driver, ADC sampling, and pre-processing functions. Subsequently, the execution program is installed on the MCU. The execution time of the program, as shown in Table 2 is only 0.71 seconds for the entire process, including data sampling, pre-processing, and prediction, as illustrated in Fig. 4. Regarding MCU memory usage, the proposed system requires only 125.8 KB of read-only memory (ROM) and 28.8 KB of random-access memory (RAM). This system can be implemented in most of the memory management systems available in the market.

Table 2. PROGRAM EXECUTION TIME

	Sampling	Pre-processing	Prediction	All
Time(s)	0.20	0.22	0.014	0.71

```
arm-none-eabi-size --format=berkeley "Water_AI.elf"
   text    data     bss     dec     hex filename
 128568     244   29520  158332   26a7c  Water_AI.elf
Finished building: Water_AI.siz
```

Fig. 4 Memory usage at MCU

3. Results and Discussion

The edge-computing platform proposed in this paper, which is include MCU, microphone, and solenoid electromagnet is shown in Fig. 5. In the picture, the micro switches on both sides of the solenoid valve ensure that the former can knock the middle of the bottle correctly
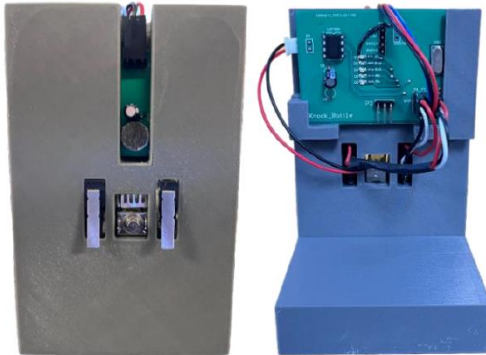
Fig. 5 Edge-computing platform

After installing the model into the MCU, the liquid level can be predicted. The performance of the AI model is often assessed using the confusion matrix, which categorizes predicted results into four classes: true positive (TP), false positive (FP), true negative (TN), and false negative (FN). These metrics are used to calculate various performance indicators, including accuracy as defined in formula (5), true positive rate (TPR), false positive rate (FPR), true negative rate (TNR), and false negative rate (FNR).

In this paper, a multicategory model is proposed as the basis of analysis. Fig. 6 illustrates the index calculation for the '100ml' category when there are five output categories. If the predicted result matches the true label ('100ml'), it is labeled as $C\_1,1$ (TP). Here, the symbol C represents a case, and C_(i,j) denotes the case in the i-th row and j-th column of a confusion matrix. Therefore, $C\_1,1$ represents the case in the 1st row (true label: '100ml') and 1st column (predicted label: '100ml') of a confusion matrix. TPR represents the probability that a positive sample is correctly identified as positive, while FPR is the probability that a negative sample is incorrectly classified as positive. TNR is the probability that an actual negative sample is correctly identified as negative, and FNR is the probability that a positive sample is incorrectly identified as negative. A complete list of formulas is provided in the following sections (2) to (6).

Fig. 6 Categories of Confusion matrix

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \tag{2}$$

$$TPR = \frac{TP}{TP + FN} = \frac{C_{100,100}}{C_{100,100} + \sum_{k=200}^{500} C_{100,k}} \tag{3}$$

$$FPR = \frac{FP}{FP + TN}$$
$$= \frac{\sum_{k=200}^{500} C_{k,100}}{\sum_{k=200}^{500} C_{k,100} + \sum_{i=200}^{500}(\sum_{j=200}^{500} C_{i,j})} \tag{4}$$

$$TNR = \frac{TN}{FP + TN}$$
$$= \frac{\sum_{i=200}^{500}(\sum_{j=200}^{500} C_{i,j})}{\sum_{k=200}^{500} C_{k,100} + \sum_{i=200}^{500}(\sum_{j=200}^{500} C_{i,j})} \tag{5}$$

$$FNR = \frac{FN}{TP + FN} = \frac{\sum_{k=200}^{500} C_{100,k}}{C_{100,100} + \sum_{k=200}^{500} C_{100,k}} \tag{6}$$

Chung-Wen Hung, Chun-Liang Liu, Tai-Hsuan Wu, Chau-Chung Song

The Fig. 7 confusion matrix summarizes the actual and predicted classifications of online testing, and the confidence values of each liquid level are in line with expectations. As a result, 95% accuracy when calculating the true negative value of each water level, and 99% accuracy rate when calculating the full water levels.

The results of this study can be used to verify the feasibility of the proposed system as well as the discussed artificial intelligence model.
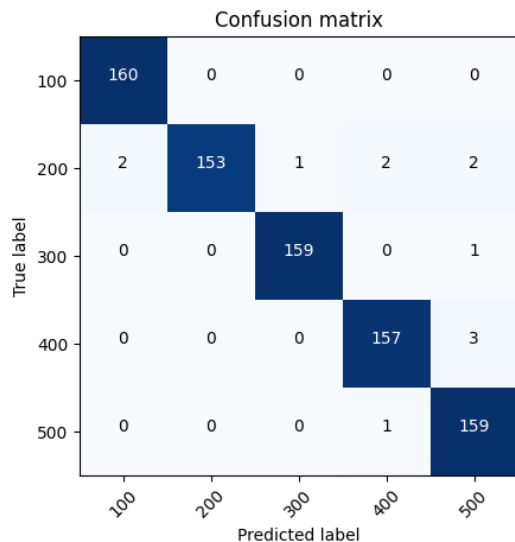
Fig. 7 Confusion matrix of the model

4. Conclusion

The paper proposes a MCU-based edge computing platform for liquid level measurement. This includes not only a knocking block, an audio signal preprocessing circuit, and MCU. but contains a detailed explanation of the development process, such as the optimization of hyperparameters, conversion of the model and installation for MCU. The utilization of glass bottles in the experiment led to more pronounced variations in the energy of the knocking sound. The experimental results can be up to 99% accurate under multiple tests. Experimental results indicate that the system is feasible and measurement accuracy is acceptable.

References

1. S. Deng, H. Zhao, W. Fang, J. Yin, S. Dustdar, and A.Y. Zomaya: Edge Intelligence: The Confluence of Edge Computing and Artificial Intelligence. IEEE Internet of Things Journal, vol. 7, no. 8, pp. 7457–7469(2020).
2. Z. Zhou, X. Chen, E. Li, L. Zeng, K. Luo and J. Zhang: Edge Intelligence: Paving the Last Mile of Artificial Intelligence with Edge Computing. in Proceedings of the IEEE, vol. 107, no. 8, pp. 1738–1762, 18884046 (2019).
3. C. W. Hung, S. X. Zeng, C. H. Lee, W. T. Li: End-to-End Deep Learning by MCU Implementation: An Intelligent Gripper for Shape Identification. Sensors. 21(3):891(2021).
4. H. H. P. Wu, C. W. Hung, S. H. Chang, Z. H. Yang: Analysis and Implementation of Noncontact Level Sensing for a Pressurized Cylinder. Journal of Sensors. 5027916 (2016).
5. J. J. Popoola, and R. V. Olst: Development and Comparative Study of Effects of Training Algorithms on Performance of Artificial Neural Network Based Analog and Digital Automatic Modulation Recognition. Journal of Engineering Science and Technology Review 8 (4), 135-144(2015).
6. J. J. Popoola, and R. V. Olst: Automatic recognition of analog modulated signals using artificial neural networks. Journal of Computer Technology and Applications, Vol. 2, No. 1, pp. 29-35(2011).
7. ARTERY Technology, AT32F415S DATASHEET, JUNE. 2022[V2.01].
8. Texas Instruments,: OPA1612A DATASHEET, March,2015[Rev.c].
9. ASIC, L9110H DATASHEET, L9110H Datasheet.
10. J. Bergstra, D. Yamins, and D. Cox: Making a Science of Model Search: Hyperparameter Optimization in Hundreds of Dimensions for Vision Architectures. Proceedings of the 30th International Conference on Machine Learning, Feb. 2013, pp. 115–123.
11. B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas: Taking the Human Out of the Loop: A Review of Bayesian Optimization. Proceedings of the IEEE, vol. 104, no. 1, pp. 148–175, Jan. 2016.

Authors Introduction

Prof. Chung-Wen Hung
He received the Ph.D. degrees in Electrical Engineering from National Taiwan University in 2006. Currently he is a Professor in National Yunlin University of Science &Technology. His research interests include the IoT, IIoT, and AI application.

Dr. Chun-Liang Liu
He received the Ph.D. degrees in Electrical Engineering from National Taiwan University of Science and Technology. Currently he is a Assistant Professor in National Yunlin University of Science &Technology.

Mr. Tai-Hsuan Wu
He received the B.S. degrees and now he is studying for the M.S. degree in electrical engineering from National Yunlin University of Science and Technology.

Dr. Chau-Chung Song

He received the Ph.D. degrees in Electrical Engineering from National chiao tung university. Currently he is a Distinguished professor and Chair in National Formosa University

Potential of genetic algorithms in multi-UAV coverage problem

Ramil Faizullin
Intelligent Robotics Department, Kazan Federal University, 420008, Kazan, Russian Federation

Tatyana Tsoy
Intelligent Robotics Department, Kazan Federal University, 420008, Kazan, Russian Federation

Edgar A. Martínez-García
Department of Industrial Engineering and Manufacturing, The Autonomous University of Ciudad Juarez, 32315 Cd Juárez, Chihuahua, Mexico

Evgeni Magid
Intelligent Robotics Department, Kazan Federal University, 420008, Kazan, Russian Federation
HSE University, Moscow, Russian Federation
E-mail: fzllnrml@gmail.com, tt@it.kfu.ru, edmartin@uacj.mx, magid@it.kfu.ru

Abstract

For a rapid area coverage multiple UAVs are often used simultaneously. However, a path planning for a UAVs group during an area coverage task is computationally challenging. In practice, heuristic algorithms are applied to solve this problem. This paper overviews approaches to an area coverage problem with a group of UAVs using genetic algorithms. The article explores modifications that may be useful for a genetic algorithm for solving the coverage problem as well as representation methods for chromosomes that reflect a path of multi-UAV. Additionally, UAV group collision avoidance strategies during area coverage are considered.

Keywords: Genetic algorithms, Coverage path, multi-UAV coverage

1. Introduction

A current rapid progress in robotics drives robotic systems into many areas of a human life including autonomous driving [1], manufacturing [2], search and rescue activities [3], agriculture [4], [5], and medical services [6]. In modern industrial applications of Industry 4.0 [7] a human–robot interaction in a shared workspace is emphasized [8], [9]. Service robotics [10] is a promising area for robots' integration covering multiple applications from education [11] and medicine [12] to entertainment [13] and advertisement [14]. A task of area coverage using mobile robots often arises in various practical applications of mobile robotics. Together with a proper scheduling, a determining of an optimal route for a mobile robot to ensure coverage of all points in a certain area while avoiding collisions with obstacles plays an important role [15].

Unmanned Aerial Vehicles (UAVs) provide an efficient solution for automating coverage and monitoring of large areas. UAVs are not constrained by obstacles on the ground and can quickly navigate through hard-to-reach areas while reducing human labor costs [16]. The use of UAVs for area coverage can be a key element in solving complex problems such as security [17], mapping [18], forest fire monitoring [19], search operations [20]. The task of covering an area with a group of UAVs attempts finding a path for each UAV such that the UAVs observe all locations of a given area without colliding with each other [21]. This task may have constraints, such as a time limit for an entire mission, safety constraints related to a minimum distance between UAVs, flight range constraints due to a battery capacity, and others.

The coverage path problem for a mobile robot is NP-hard [22] due to a need of analyzing an entire set of possible solutions, a number of which grows exponentially with increasing a complexity and a size of an area to be covered. In practice, heuristic algorithms are often employed for this task as they allow to get close to the optimal solution rather quickly, yet do not find a guaranteed optimal solution.

Genetic algorithms are popular heuristic algorithms that have a relatively simple implementation and could be easily parallelized, which is especially important when working with large solution spaces. Genetic algorithms can deal with multiple optimization criteria, which is often the case in practice. They provide a method for

exploring an entire solution space to find a global optimum, which is useful in problems with many local extrema. All these turn genetic algorithms into an effective solution of the multiagent coverage problem.

2. Genetic algorithms

Genetic algorithms are based on principles of biological evolution and often are applied to search problems [23]. They model a process of natural selection to find optimal solutions. These algorithms interact with a population of potential solutions, each called an *individual*. By successively evaluating a set of solutions a new generation is created. The best solutions have a higher chance of passing on their characteristics to a next generation. Over time, solutions within the population become better adapted to a particular task.

First a genetic algorithm creates an initial population, which in typically contains random *chromosomes*. Each chromosome is a possible solution of a problem and a fitness function assigns each chromosome a numerical value that reflects its effectiveness in solving the problem. To create a next generation, individuals are selected from the population in a fixed manner. Individuals with a higher fitness for an environment have a better chance of surviving and passing their traits on to the next generation. This leads to an improvement in the population over time. A crossover operation is performed to create new offspring chromosomes. This allows new individuals to inherit the best characteristics from their parents, thus preserving them for future generations. Mutations are random changes in chromosomes of individuals that maintain a genetic diversity and periodically contribute to significant leaps in a population development. After a crossover, there is a certain probability that a mutation occurs. This mutation slightly alters chromosomes of an offspring.

3. Features of genetic operators' application in area coverage

In the area coverage problem for UAVs, it is necessary to find an optimal trajectory employed by a UAV to visit all given points of an area. Each chromosome encodes a trajectory of a UAV and the population is a set of possible UAV trajectories. In practice, chromosomes are usually sequences of integers, yet a user decides on a way of trajectory encoding within a chromosome [24].

An important feature of applying a genetic algorithm to the coverage problem is that after crossover or mutation operations, a chromosome must encode some trajectory for a UAV. Therefore, crossover and mutation operations are subjects to certain constraints.

Fig. 1 shows a two-point classical crossover operation that produces two offspring. A chromosome represents an order of visiting obstacle-free cells while covering a

UAV's assigned area. If parents represent acceptable chromosomes, the offspring could contain the same genes and do not represent solutions of the coverage problem. For example, child A contains genes 2, 4, and 5, and according to the chromosome, the UAV trajectory passes through these regions twice, but it does not visit regions 1, 3, and 8.
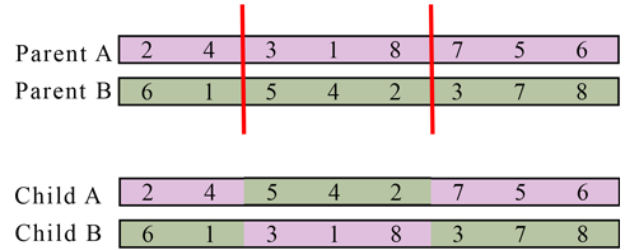
Fig. 1. An example of a crossover operation where an offspring no longer reflects a proper UAV trajectory

One solution is to modify the crossover operation. For example, using a two-point crossover from a single parent. With this method of the crossover from a single parent many children are formed and then two with a highest value of the fitness function are selected (Fig.2). In this case, each offspring reflects a proper UAV trajectory. In [25] the authors reported a four times computation acceleration of a genetic algorithm simulation using the proposed crossover method.
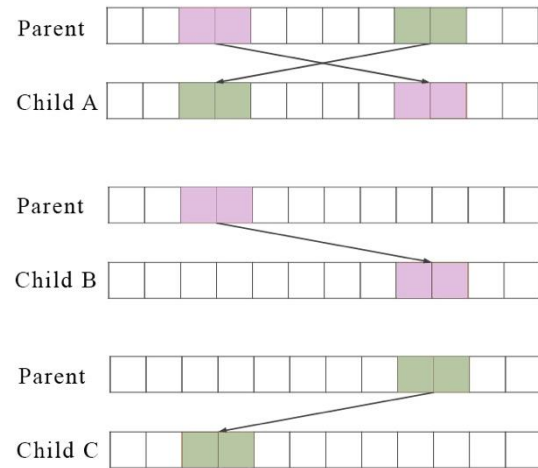
Fig. 2. An example of a crossover from one parent producing three offspring

4. Single UAV and group of UAV coverage tasks

In practice, using a group of UAVs for an area coverage increases a speed and a reliability of the procedure. Each UAV acts as a moving obstacle for other UAVs, and its route cannot be planned in isolation from the rest of the group [26]. Therefore, the genetic algorithm must consider routes of all UAVs together. A multiagent coverage problem solution is a path for each UAV in

which they cover a given area together and do not collide with each other, i.e., the solution combines paths of all UAVs into a single chromosome (Fig. 3) [27].
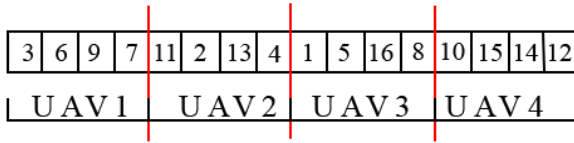
Fig. 3. A chromosome combines paths of four UAVs

5. UAV group collision avoidance strategies during area coverage

While using a group of UAVs reduces a time to complete the task, it also introduces potential risks of UAV collisions during the mission. Two strategies of UAVs' collision avoidance are discussed in this section in the context of genetic algorithms: assigning UAVs a different coverage areas and introducing a penalty for approaching other UAVs.

In the first approach, to prevent UAVs from colliding in the air, the area could be divided into equal subareas and each UAV is assigned a different subarea to cover [28]. This method allows a computation paralleling for each UAV with a genetic algorithm (as their trajectories are independent, (Fig. 4) and keeps a genetic algorithm simple. Disadvantages of this strategy include a high dependence of a resulting coverage effectiveness on a quality of an area distribution between UAVs, collision avoidance considerations while distributing UAVs (that takeoff in a common area) for the assigned subareas, and complicated splitting into subareas for a case of different survey priorities of the subareas.

In the second approach, a penalty function can be introduced to reduce a fitness function value if UAVs fly too close to each other at a coverage time [29]. This strategy is a flexible way to influence the genetic algorithm and allows using different penalty functions depending on task goals. Yet, it requires a pairwise
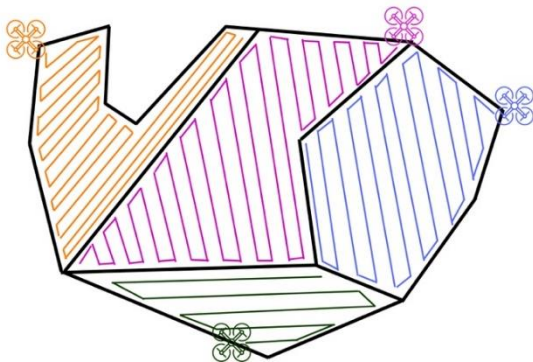
Fig. 4. Example of division of area to be covered into separate areas for each UAV

comparison of a minimum distance between each pair of UAVs and as a number of UAVs increases, a complexity increases exponentially.

6. Conclusions

Classical genetic algorithms require a task dependent adaptation in order to be employed for the problem of covering a known territory with a group of UAVs as the use of standard genetic algorithm operators in this context may destruct a chromosome that represents a trajectory of a UAVs' group. Genetic algorithms should optimize methods of representation and decomposition of a surveyed area and incorporate dynamic obstacles' collision avoidance strategies to increase the efficiency of their usage.

Acknowledgement

References

1. J. Cheng, L. Zhang, Q. Chen, X. Hu and J. Cai, A review of visual slam methods for autonomous driving vehicles, Engineering Applications of Artifcial Intelligence, vol. 114, 2022, pp. 104992.
2. A. Gribkov, M. Morozkin, V. Kuptsov, P. Pivkin and A. Zelenskii, Industry 4.0 concepts in the machine-tool industry, Russian Engineering Research, vol. 41(7), 2021, pp. 634–635.
3. E. Magid, Simulation Tools for Urban Search and Rescue Robotics, International Conference on Artificial Life and Robotics (ICAROB 2023), 2023, pp. 4-11.
4. V. Nguyen, Q. Vu, O. Solenaya and A. Ronzhin, Analysis of main tasks of precision farming solved with the use of robotic means, in MATEC Web of Conferences, EDP Sciences, vol. 113, 2017, pp. 02009.
5. R. Iakovlev, A. Saveliev, Approach to Implementation of Local Navigation of Mobile Robotic Systems in Agriculture with the Aid of Radio Modules, Telfor Journal. vol. 12(2), pp. 92-97. 2020.
6. D. Kolpashchikov, O. Gerget and R. Meshcheryakov, Robotics in healthcare, in Handbook of Artifcial Intelligence in Healthcare, Springer, 2022, pp. 281–306.
7. R. Galin and R. Meshcheryakov, Automation and robotics in the context of Industry 4.0: the shift to collaborative robots. In IOP Conference Series: Materials Science and Engineering, vol. 537(3), May 2019, pp. 032073.
8. R. Galin and R. Meshcheryakov, Review on human–robot interaction during collaboration in a shared workspace, in International Conference on Interactive Collaborative Robotics, Springer, 2019, pp. 63–74.
9. R. Sultanov, S. Sulaiman, H. Li, R. Meshcheryakov and E. Magid, A review on collaborative robots used in industrial and service sectors. Siberian Conference on Control and Communications, (SIBCON 2022), 2022, pp. 1-7.
10. R. Belk, Ethical issues in service robotics and artifcial intelligence, The Service Industries Journal, vol. 41(13-14), 2021, pp. 860–876.

11. L. Gavrilova, A. Kotik, T. Tsoy, E. A. Martinez-Garcia, M. Svinin and E. Magid, Facilitating a preparatory stage of real-world experiments in a humanoid robot assisted English language teaching using gazebo simulator, 2020 13th International Conference on Developments in eSystems Engineering (DeSE), IEEE, 2020, pp. 222–227.

12. A. Sagitov, T. Tsoy, H. Li and E. Magid, Automated open wound suturing: detection and planning algorithm, J. Robotics Netw. Artif. Life, vol. 5(2), 2018, pp. 144–148.

13. K. J. Morris, V. Samonin, J. Baltes, J. Anderson and M. C. Lau, A robust interactive entertainment robot for robot magic performances, Applied Intelligence, vol. 49(11), 2019, pp. 3834–3844.

14. P. Bustos, L. J. Manso, A. J. Bandera, J. P. Bandera, I. Garcia-Varea and J. Martinez-Gomez, The cortex cognitive robotics architecture: Use cases, Cognitive Systems Research, vol. 55,2019, pp. 107–123.

15. E. Galceran and M. Carreras, A survey on coverage path planning for robotics, Robotics and Autonomous systems, vol. 61(12), 2013, pp. 1258-1276.

16. E. Denisov, A. Sagitov, K. Yakovlev, K. Su, M. Svinin and E. Magid, Towards Total Coverage in Autonomous Exploration for UGV in 2.5D Dense Clutter Environment, In Proceedings of the 16th International Conference on Informatics in Control, Automation and Robotics, ICINCO; SciTePress, vol. 2, 2019, pp. 409-416.

17. T. M. Cabreira, L. B. Brisolara and P. R. Ferreira Jr., Survey on coverage path planning with unmanned aerial vehicles, Drones, vol. 3(1), 2019, p. 4.

18. D. Albani and D. Nardi, V. Trianni, Field coverage and weed mapping by UAV swarms, IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Ieee, 2017.

19. K. A. Ghamry and Y. Zhang, Cooperative control of multiple UAVs for forest fire monitoring and detection, 12th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA), IEEE, 2016.

20. I. Maza and A. Ollero, Multiple UAV cooperative searching operation using polygon area decomposition and efficient coverage algorithms, Distributed Autonomous Robotic Systems 6. Tokyo, Springer Japan, 2007, pp. 221-230.

21. J. Muñoz, B. López, F. Quevedo, C. A. Monje, S. Garrido and L. E. Moreno, Multi UAV coverage path planning in urban environments. Sensors, vol. 21(21), 2021, 7365.

22. E. M. Arkin, S. P. Fekete and J. S. Mitchell, Approximation algorithms for lawn mowing and milling, Computational Geometry, vol. 17(1-2), 2000, pp. 25-50.

23. S. Katoch, S. S. Chauhan and V. Kumar, A review on genetic algorithm: past, present, and future, Multimedia tools and applications, vol. 80, 2021, pp. 8091-8126.

24. A. Tuncer and M. Yildirim, Chromosome coding methods in genetic algorithm for path planning of mobile robots, Computer and Information Sciences II, 26th International Symposium on Computer and Information Sciences, Springer London, 2012.

25. T. H. Pham, Y. Bestaoui and S. Mammar, Aerial robot coverage path planning approach with concave obstacles in precision agriculture, Workshop on Research, Education and Development of Unmanned Aerial Systems (RED-UAS), IEEE, 2017.

26. Y. Bai, Y. Wang, M. Svinin, E. Magid and R. Sun, Adaptive multi-agent coverage control with obstacle avoidance, IEEE Control Systems Letters 6, 2021, pp. 944-949.

27. B. Li, B. Moridian, A. Kamal, S. Patankar and N. Mahmoudian, Multi-robot mission planning with static energy replenishment, Journal of Intelligent & Robotic Systems 95, 2019, pp. 745-759.

28. R. Almadhoun, T. Taha, L. Seneviratne and Y. Zweiri, A survey on multi-robot coverage path planning for model reconstruction and mapping, SN Applied Sciences 1, 2019, pp. 1-24.

29. M. Nazarahari, E. Khanmirza and S. Doostie, Multi-objective multi-robot path planning in continuous environment using an enhanced genetic algorithm, Expert Systems with Applications 115, 2019, pp. 106-120.

Authors Introduction

Mr. Ramil Faizullin
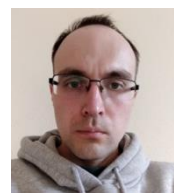
He is a Ph.D. student in Computer science and information processes at the Institute of Information Technology and Intelligent Systems (ITIS), Kazan Federal University (KFU), Russia.

Ms. Tatyana Tsoy

In 2012 she graduated from International Area Studies Master's Degree Program at the University of Tsukuba. In 2023 completed a PhD in Computer science and information processes at ITIS, KFU.

Professor Edgar A. Martínez-García

He is a full Professor at the Universidad Autónoma de Ciudad Juárez, Mexico; founder and Head of the Robotics Laboratory; leader of the Mechatronics academic body at the Institute of Engineering and Technology, since 2007. He obtained his Ph.D. degree in Robotics Engineering from the University of Tsukuba, Japan (2005).

Professor Evgeni Magid

A Professor, a Head of Intelligent Robotics Department and a Head of Laboratory of Intelligent Robotic Systems (LIRS) at KFU. Professor at HSE University, Russia. Senior IEEE member. Previously he worked at University of Bristol, UK; Carnegie Mellon University, USA; University of Tsukuba, Japan; National Institute of Advanced Industrial Science and Technology, Japan. He earned his Ph.D. degree from University of Tsukuba, Japan. He authors over 200 publications.

Construction of Anthropomorphic Grippers with Adaptive Control

Evgeny Dudorov
JSC 'SPA 'Android technics', 23 Grayvoronovskaya st., Moscow, 109518, Russia

Julia Zhdanova
Institute of Cybernetics, MIREA – Russian Technological University, 78 Vernadskij prosp., Moscow 119454, Russia

Ivan Zhidenko
JSC 'SPA 'Android technics', 23 Grayvoronovskaya st., Moscow, 109518, Russia

Vladimir Moshkin
Institute of Cybernetics, MIREA – Russian Technological University, 78 Vernadskij prosp., Moscow 119454, Russia

Alexander Eryomin
Intelligent Robotics Department, Kazan Federal University, Kazan, 420111, Russia

Evgeni Magid
Intelligent Robotics Department, Kazan Federal University, Kazan, 420111, Russia

Alexander Permyakov
JSC 'SPA 'Android technics', 23 Grayvoronovskaya st., Moscow, 109518, Russia
Email: rusandroid@mail.ru, aneremin@it.kfu.ru, magid@it.kfu.ru

Abstract

The functionality of industrial robots is provided primarily by the capabilities of their end effectors. The limited capabilities of robotic grippers determined the transition to the creation of anthropomorphic grippers. The number of degrees of freedom (DOF) of the end effector in the form of an anthropomorphic hand ensuring reliable grasping and holding the object should be at least nine, preferably twelve. The implementation of such a design is possible only when switching to the principle of construction of an underactuated grippers system. This paper presents the concept of constructing a group drive, which ensures the implementation of the movement of the output links of two or more executive groups from one motor. Technical solutions are based on the development of methods for analyzing complex mechanical systems using functional circuits.

Keywords: Anthropomorphic gripper, underactuated grippers, end effector, robotic grasp

1. Introduction

Today, robotic manipulators are actively used in industrial production [1]. They improve working conditions of employees and reduce enterprise costs ensuring a continuous production. Typical applications include assembly, casting, stamping, cutting, machine loading/unloading, welding and material handling [2]. Generally, a robotic arm functionality is determined by design and properties of an end effector. The most applied end tool is a robotic gripper. The choice of a grip depends on many characteristics, including weight, shape of an object, motion speed, permissible compression force, and a point of contact [3]. Values of each characteristics are often taken into account at a design stage of a robotic arm. Thus, problems arise with an inefficient operation of a robot in case of miscalculations in the design or/and an inability to adjust the arm to a new task. One of the solutions is to unify the hand [4].

The most widely used are universal vacuum suction grippers and multi-fingered hands [5]. The first one is a single mass of granular material that flows around a target object and takes its shape upon applying a pressure on the object. The disadvantage of this approach is a need to return to a neutral state after each grasping. Multi-fingered hands usually possess many independently actuated joints. This design of the end effector has sufficient softness and rigidity for a reliable and stable grip of an object [6]. [7]. MCU Based Edge Computing. The main difficulty of the approach is related to a computational complexity of tactile sensing and computer vision based algorithms [8]. Our research is aimed at developing a concept of a multi-fingered hand using both active and passive joints. Such design allows

Evgeny Dudorov, Ivan Zhidenko, Alexander Permyakov, Julia Zhdanova, Vladimir Moshkin, Alexander Eryomin, Evgeni Magid

the hand to envelop an object without complex calculations.

The most common anthropomorphic grippers are structure diagrams of grippers with three [9] or four parallel executive groups of links [10]. Each executive group includes, as a rule, three output links [11]. Collectively, the gripper has nine or more degrees of freedom (DOF). There are two typical layout schemes for installing individual drives for each link. In accordance with the first one, the motors are placed at the output links [10]. In the second case, the motors are located within a single link of the manipulator [12], [13].

Typically, dimensions of output links do not allow internal installation of motors with a sufficient power [14]. Consequently, forces generated at endpoints of the output links do not exceed 1.5 N [10]. Installing motors in a single link involves a significant complication of a design. For instance, a gripper by DLR [15] has 65 pulleys for laying flexible rods on a manipulator link and 38 pulleys directly on the gripper. At the same time, it is also impossible to provide a significant effort at the output link. In a modified DLR version, the maximum force is 9 N [16]. The use of individual drives complicates a control system. In this case, a number of controllers for drives equals to a number of output links.

A group drive solves the abovementioned problems of controllability and complexity of individual drives. In the ASIMO grippers [17] and JPL‐Nautilus Gripper [18] a single motor provides a movement of three output links in each executive group. An object is grasped according to the principle of kinematically dependent movements of the links. Thus, the manipulator reliably grips objects of only a certain size with all three links.

2. Related Work

A qualitative change in operational characteristics of a system with a group drive is achieved by using a principle of underactuated grippers. It is based on an implementation of additional passive elements into a structure of each executive group. These can be compression springs [19], [20] or tension springs [21], [22]. At the same time, an adaptive motion control is implemented by output links in the executive group. Links move sequentially, from a proximal link to a distal link. A change of a control object (an output link) is achieved when external conditions change, and a moving link stops after reaching an external object.

A further reduction of anthropomorphic gripper mass is possible by ensuring a motion of output links of adjacent executive groups from a single motor. Using the principle of underactuated grippers, flexible elements should be placed into a group drive system for an entire gripper. Similarly, for the executive group, this will ensure independent motion of the executive groups of links from

a single motor. Thus, inspired by [23] we designed a modifiable structure of an anthropomorphic gripper, which is presented in this paper.

3. A gripper design

The proposed design is based on an anthropomorphic gripper with two parallel executive groups of links (Fig. 1). From motor 1 that is installed at the basis of the gripper, the movement is transmitted by two parallel streams to two executive groups. A flexible twisting spring is integrated into a kinematic scheme of each stream, connecting coaxial shafts *a* and *b*. Links 2 and 3 transmit movement through transmission systems to the output links using the shafts. The installation of flexible elements adds a relative rotation of the coaxial shafts to a motion transmission system.
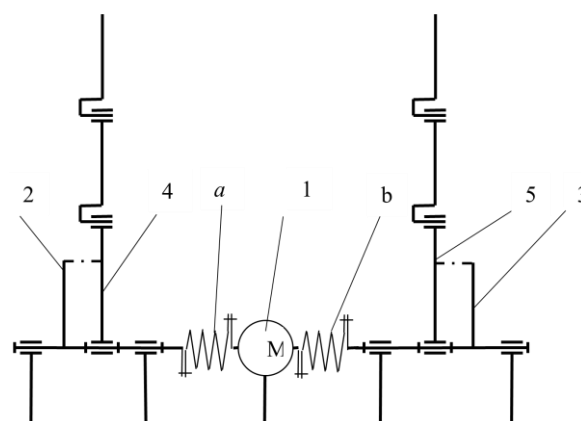
Fig. 1. Structure diagram of an anthropomorphic gripper with two parallel executive groups of the output links.
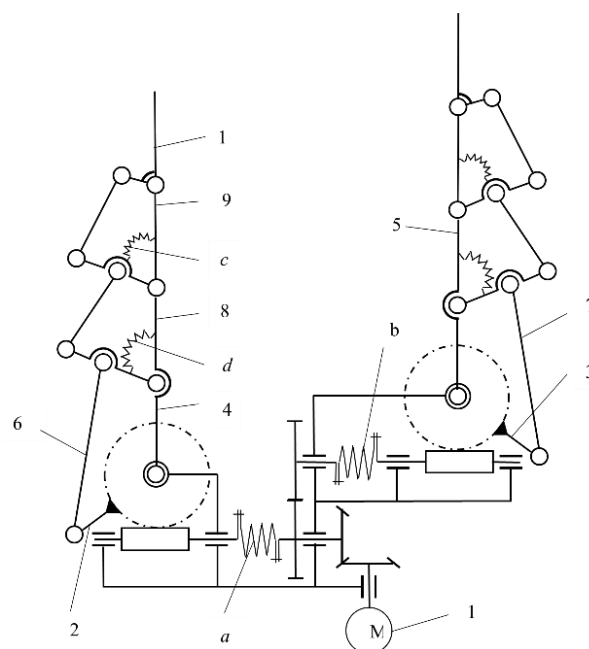
Fig. 2. Structure diagram of an anthropomorphic gripper with two opposite executive groups of the output links.

Fig. 2 shows a structure diagram of the anthropomorphic gripper with opposite executive groups of links 4, 5 and a lever system for motion transmitting [24]. The movement is transmitted from motor 1 in two parallel streams to driving links 2 and 3 involving twisting springs *a* and *b*. Then links 2 and 3 set into motion lever mechanisms 6 and 7. Executive groups 4 and 5 are constructed using the principle of underactuated grippers. Compression springs *d* and *c* are installed between the links of the motion transmission system and the driving links.

Executive groups 4 and 5 move from initial position synchronously in a direction of an object located between them, which should be grasped. A grasp can shift towards one of the executive groups, depending on a shape of the object. As the object is gripped by link 8, a rotational movement of leading link 2 decreases and spring *a* gets partially deformed. After grasping the object with all links 8-10, the transmission of rotation through spring *a* stops and the spring gets twisted. The movement of executive group 5 is maintained until its links completely encircle the object. The rotation of the motor stops after reaching a preset torque value. Deformed springs *a* and *b* retain a force corresponding to an end of the grasping process. This provides a given force effect on the object.

Fig. 3. The constructed anthropomorphic gripper with three executive groups of links.

Fig. 3 shows the constructed anthropomorphic gripper with three executive groups of links. The single motor provides the movement of the output links. The distal link possesses 10.5 N force for the entire length of the executive group links of 100 mm.

4. Conclusion

The paper proposed a new design an anthropomorphic hand with an underactuated gripper. A movement of output links of several executive groups of the gripper is performed using a single motor. A number of implemented flexible links equals to a number of executive groups. A shape of an object to be grasped determines a resulting motion of executive group links and a sequence of their motion. Experimental validation confirmed feasibility of the proposed construction of the anthropomorphic gripper.

References

1. L. Birglen and T. Schlicht, A statistical review of industrial robotic grippers. Robotics and Computer-Integrated Manufacturing, vol. 49, 2018, pp. 88-97.
2. H. A. F. Almurib, H. F. Al-Qrimli and N. Kumar, A review of application industrial robotic design. 2011 Ninth International Conference on ICT and Knowledge Engineering, IEEE, 2012, pp. 105-112.
3. Z. Samadikhoshkho, K. Zareinia and F. Janabi-Sharifi, A brief review on robotic grippers classifications. 2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE), IEEE, 2019, pp. 1-4.
4. M. S. Choi et al. Development of multi-purpose universal gripper. 2017 56th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE), IEEE, 2017, pp. 1421-1424.
5. P. V. P. Reddy and V. Suresh, A review on importance of universal gripper in industrial robot applications. Int. J. Mech. Eng. Robot. Res, vol. 2, 2013, pp. 255-264.
6. K. Khusnutdinov et al. Development and Implementation of Grasp Algorithm for Humanoid Robot AR-601M. ICINCO, vol. 2, 2019, pp. 379-386.
7. K. Khusnutdinov et al. Household objects pick and place task for AR-601M humanoid robot. Interactive Collaborative Robotics: 4th International Conference, ICR 2019, Proceedings 4, Springer International Publishing, August 2019, pp. 139-149.
8. E. Brown et al. Universal robotic gripper based on the jamming of granular material. Proceedings of the National Academy of Sciences, vol. 107, 2010, pp. 18809-18814.
9. L. U. Odhner et al. A compliant, underactuated hand for robust manipulation. The International Journal of Robotics Research, vol. 33, no. 5, 2014, pp. 736-752.
10. Y. H. Lee et al. Design of anthropomorphic robot hand with IMC joints. 2016 13th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI), IEEE, 2016, pp. 336-337.
11. N. Elangovan et al. Improving robotic manipulation without sacrificing grasping efficiency: A multi-modal, adaptive gripper with reconfigurable finger bases. IEEE Access, vol. 9, 2021, pp. 83298-83308.
12. M. Shahmohammadi and M. Liarokapis, A series elastic, compact differential mechanism: On the development of adaptive, lightweight robotic grippers and hands. 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2021, pp. 6110-6116.
13. K. Mizushima et al. Multi-fingered robotic hand based on hybrid mechanism of tendon-driven and jamming transition. 2018 IEEE International Conference on Soft Robotics (RoboSoft), IEEE, 2018, pp. 376-381.
14. U. Kim et al. Integrated linkage-driven dexterous anthropomorphic robotic hand. Nature communications, vol. 12, no. 1, 2021, p. 7177.
15. M. Grebenstein et al. The DLR hand arm system. 2011 IEEE International Conference on Robotics and Automation, IEEE, 2011, pp. 3175-3182.

16. F. Lange, G. Quere and A. Raffin, Decoupled Control of Position and/or Force of Tendon Driven Fingers. 2019 International Conference on Robotics and Automation (ICRA), IEEE, 2019, pp. 1176-1182.
17. Y. Sakagami et al. The intelligent ASIMO: System overview and integration. IEEE/RSJ international conference on intelligent robots and systems, IEEE, vol. 3, 2002, pp. 2478-2483.
18. S. B. Backus et al. Design and testing of the JPL‐Nautilus Gripper for deep‐ocean geological sampling. Journal of Field Robotics, vol. 37, no. 6, 2020, pp. 972-986.
19. D. Hirano, K. Nagaoka and K. Yoshida, Design of underactuated hand for caging-based grasping of free-flying object. Proceedings of the 2013 IEEE/SICE International Symposium on System Integration, IEEE, 2013, pp. 436-442.
20. Z. Ren et al. HERI hand: A quasi dexterous and powerful hand with asymmetrical finger dimensions and under actuation. 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2017, pp. 322-328.
21. S. Qiao et al. Self-adaptive grasp process and equilibrium configuration analysis of a 3-DOF UACT robotic finger. Mechanism and Machine Theory, vol. 133, 2019, pp. 250-266.
22. L. Kang, S. H. Kim and B. J. Yi, Modeling, Design, and Implementation of an Underactuated Gripper with Capability of Grasping Thin Objects. Machines, vol. 9, no. 12, 2021, p. 347.
23. T. Laliberte, L. Birglen and C. Gosselin, Underactuation in robotic grasping hands. Machine Intelligence & Robotic Control, vol. 4, no. 3, 2002, pp. 1-11.
24. A. Bogdanov, A. Permyakov and Y. Zhdanova, Synthesis of structural scheme of drive of adaptive multiple-link gripper. MATEC Web of Conferences, EDP Sciences, vol. 161, 2018, p. 03009.

Authors Introduction

Dr. Evgeny A. Dudorov

He is Candidate of Technical Sciences, Assistant Professor, Executive Director of JSC "Scientific production association "Android technics", Russia.

Mr. Ivan Zhidenko
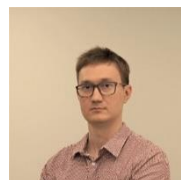
He is Head of the Advanced Projects Department, of JSC "Scientific production association "Android technics", Russia.

Ms. Julia Zhdanova

She is Deputy Director of the Institute of Cybernetics, MIREA — Russian Technological University, Russia.

Dr. Vladimir Moshkin

He is Candidate of Technical Sciences, assistant professor, MIREA — Russian Technological University, Russia.

Mr. Alexander Eryomin

He received a BSc degree from Siberian Federal University in 2022. Currently, he is a second-year student of Master degree program in Intelligent Robotics at the Institute of Information Technology and Intelligent Systems (ITIS) at the Kazan Federal University (KFU).

Prof. Evgeni Magid

He is Professor, Head of Intelligent Robotics Department and Head of Laboratory of Intelligent Robotic Systems (LIRS) at KFU, Russia. Professor at HSE University, Russia. Senior IEEE member. Previously he worked at University of Bristol, UK; Carnegie Mellon University, USA; University of Tsukuba, Japan; National Institute of Advanced Industrial Science and Technology, Japan. He earned his Ph.D. degree from University of Tsukuba, Japan. He authors over 200 publications.

Mr. Alexander Permyakov

He is General Director of JSC "Scientific production association "Android technics", Russia.

Simulation-Based Enhancement of SNR in Drone Communication through Uniform Linear Array Configurations

Gershom Phiri[1], Mastaneh Mokayef[1*], MHD Amen Summakieh[1], M.K.A Ahamed Khan[1], Sew Sun Tiang[1], Wei Hong Lim[1], Abdul Qayyum[2]

[1]*Faculty of Engineering Technology and Built Environment, UCSI University,*
Kuala Lumpur, 56000, Malaysia
[*]Email: mastaneh@ucsiuniversity.edu.my
https://www.ucsiuniversity.edu.my/

[2]*National Heart and Lung Institute, Imperial College London, UK*

Abstract

As drones navigate through shared airspace, they often encounter other drones, wireless devices, and communication systems. This coexistence creates potential sources of interference that can degrade the signal-to-noise ratio (SNR). To maintain reliable communication in drone systems, it is crucial to effectively manage and mitigate interference from other drones and wireless devices operating on the same frequency bands. By addressing these challenges, we can ensure a stable and dependable SNR for seamless communication among drones. This paper sheds light on the history of applications and challenges of utilizing flying base stations for wireless networks and analyzes different factors that affect signal-to-noise ratio (SNR) to enhance the performance of drone communication.

Keywords: SNR, UAV, drone, Matlab, IoT

1. Introduction

Unmanned aerial vehicles (UAVs) are set to play a crucial role in the next generation of wireless networks, offering a promising solution to meet the ever-increasing user demands. Their mobility, flexibility, improved line-of-sight capabilities, and ability to reach inaccessible areas make them ideal candidates to act as aerial base stations. Researchers are actively exploring various aspects of deploying, analyzing performance, managing resources, optimizing trajectories, and modeling channels in such networks. This survey article focuses on different applications and the algorithms involved in implementing aerial base stations, providing a comprehensive review of each research area. In summary, this article highlights the key applications, challenges, and technology employed in the design and analysis of UAVs as base stations.

In recent years, the widespread adoption of drones across various industries and applications has been remarkable. From capturing stunning aerial photographs to delivering packages, drones have become increasingly prevalent. One crucial factor that significantly influences the performance and efficiency of drone networks is the signal-to-noise ratio. The signal-to-noise ratio represents the ratio between the desired signal and unwanted noise in a communication system. A high signal-to-noise ratio is paramount for clear and reliable communication between drones and ground control stations, as well as for

accurate data transmission [1]. However, achieving a high signal-to-noise ratio in drone networks presents several challenges [2]. The use of multiple drones operating in close proximity can result in interference and signal degradation, affecting the overall quality of communication. Moreover, the noise generated by the drones themselves, commonly referred to as ego-noise, can further impact the signal-to-noise ratio. These challenges necessitate careful consideration and innovative solutions to optimize the signal-to-noise ratio in drone networks, ensuring efficient and reliable communication between drones and ground control stations [3]. The interference experienced by each drone at a particular location can be measured by its signal-to-noise ratio (SNR). The SNR is influenced by multiple factors, including the transmission power levels, antenna properties, drone altitude, path loss, and the power spectral density of the surrounding noise. These variables collectively contribute to the observed interference levels and play a crucial role in determining the quality of communication in drone networks [4]. The altitude at which drones fly significantly impacts the signal-to-noise ratio in their networks. Higher altitudes result in larger distances between the transmitter and receiver, leading to a decrease in signal quality. Atmospheric attenuation further contributes to signal loss as altitude increases. Transmission power levels and antenna characteristics of drones also play a vital role in determining signal strength. Optimal power levels ensure overcoming noise and interference, while antenna properties affect signal

Gershom Phiri, Mastaneh Mokayef, MHD Amen Summakieh, M. K. A Ahamed Khan, Sew Sun Tiang, Wei Hong Lim, Abdul Qayyum

coverage and directionality. Lastly, the noise power spectral density, originating from various sources, introduces background noise that affects the overall signal-to-noise ratio. Careful consideration of these factors is crucial for maintaining reliable communication in drone networks [5], [6], [7]. In [8], they delve into the realm of drone-assisted backscatter communication within an Internet of Things (IoT) sensor network. Their work focuses on developing a framework to evaluate the likelihood of ground-based sensor nodes being covered. Raja's study in 2021 explores how communication propagates between drones and base stations in indoor environments, taking into account factors like diffraction, frequency, and atmospheric attenuation [9]. While paper [10] proposes an innovative multi-UAV communication model that prioritizes security. This model employs a wireless mesh network and cryptographic techniques to ensure efficient and protected communication among multiple drones. The study underscores the significance of establishing reliable data communication security between drones and servers. Collectively, these papers offer valuable insights into coverage probability assessment, propagation mechanisms, multi-UAV coordination, and security measures within the realm of drone communication. To aid in understanding the communication requirements of UAVs, Table 1 concisely presents a summary of these requirements. This reference table offers a clear and accessible overview of the key aspects involved in UAV communication, making it easier for readers to grasp the necessary information.

Table 1: Essential Communication Requirements for UAV Systems

	Data Type	Data Rate
DL	Synchronization (PSS/SSS)	N/A
	Radio control (PDCCH)	
	Command and control (C&C)	60-100 kbps
UL	Command and Control (C &C)	60-100 kbps
	Application data	Up to 50 Mbps

A Comparative Study of the distinctive features of drone types are presented in [9], [10], [11] and the summary of these results is tabulated in Table 2.

Table 2: Distinctive Features of Drone Types [11], [12]

	Micro (weight6100g)	Very Small (100g<weight<2kg)	Small (2kg6weight<25kg)	Medium (25kg6weight6150kg)	Large (weight>150kg)
Model	Kogan nano drone	Parrot Disco	DJI Spreading Wings S900	Scout B-330 UAV helicopter	Predator B
Illustration	A 16g N/A Multi-rotor	B 750g N/A Fixed-wing	C 3.3kg 4.9 kg Multi-rotor	D 90kg 50kg Multi-rotor	E 2223kg 1700kg 1700kg Multi-rotor
Weight	16g	750g	3.3kg	90kg	2223gg
Payload	N/A	N/A	4.9kg	50kg	1700kg
Flying mechanism	Multi-rotor	Fixed-wing	Multi-rotor	Multi-rotor	Multi-rotor
Range	50-80 m	2km	N/A	N/A	1852km
Altitude	N/A	N/A	N/A	3km	15km
Flight time	6-8 min	45min	18min	180min	1800min
Speed	N/A	80 km/h	57.6 km/h	100km/h (horizontal)	482km/h
Power supply	3.7V/160mAh Li-battery	2700mAh/25A 3-cell LiPo Battery	LiPo Battery (6S, 10000mAh_15000mAh, 15C(Min))	Gasoline	950-shaft-horsepower Turboprop Engine

2. Methodology
2.1. Performance evaluation Through SNR strength

Unmanned aerial vehicles (UAVs) have gained significant popularity, due to their autonomous features, versatility, and wide range of applications. In this section, we will delve into a detailed evaluation of drone performance. Our main focus will revolve around analyzing a key metric of the SNR. This metric will be used to thoroughly assess and scrutinize the effectiveness of drone operations, providing valuable insights into their overall efficacy. SNR is commonly expressed in decibels and represents the ratio of signal power to noise power. An SNR greater than 1:1 (greater than 0 dB) indicates that the signal is stronger than the noise. When information is transmitted wirelessly from one point to another, it is

referred to as a data link. In the context of unmanned aviation, this wireless transmission is known as a radio link or radio modem. The data link enables the exchange of information between the aircraft's autopilot and the Ground Control Station (GCS). This communication involves two distinct links: the uplink, which transmits information from the GCS to the aircraft, and the downlink, which transmits information from the aircraft to the GCS. Eq. (1) is used to calculate the SNR level.

$$SNR\ (dB) = P_{signal}\ (dB) - P_{noise}\ (dB) \quad (1)$$

Where P is the power. The power levels can be expressed in decibels per dBm. The range at which a drone can fly in relation to a base station is determined by the minimum SNR required for the specific application. It is important to establish the maximum distance that can be achieved while maintaining the desired SNR. During flight, the SNR changes as the drone moves through different distances, adapting to meet the target SNR. To illustrate this relationship, Fig. 1 depicts a simulation graph created using MATLAB, showing how the SNR is influenced by the distance traveled by the drone.

a.

b.
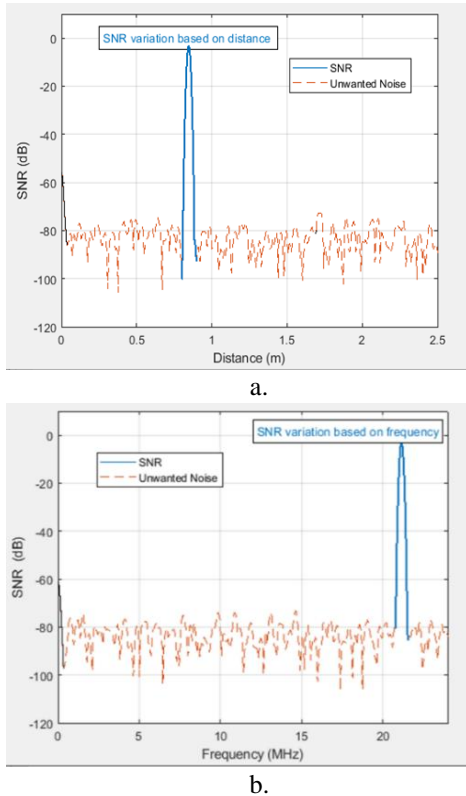
Fig. 1 SNR as a function of (a) distance traveled by drone, and (b) frequency of operation [13]

2.2. *SNR enhancement methods*

The analysis presented in Fig. 1 provides valuable insights into the significance of the signal-to-noise ratio (SNR) in determining the quality of output. The observed threshold value for SNR is -10 dB; however, to achieve

optimal output quality, it is recommended to have an SNR higher than -20 dB. The graph illustrates a gradual decline in signal quality as the range increases, with the highest signal quality observed in close proximity to the transmitter. Considering the fixed parameters of a frequency threshold of 2.4 GHz, a power threshold of 1W, and a specific distance, it becomes evident that a stronger signal is required at a distance of 1m and a frequency of 10 MHz, with a transmit power of 1W [11]. This is due to the reduced interference from various sources such as remote controls, phones, Bluetooth devices, plants, and buildings, which tend to degrade the signal quality. As the distance between the transmitter and receiver increases, the SNR decreases, indicating weaker signal quality at longer ranges.

3. SNR enhancement by Array Gain for Line-of-Sight Propagation

SNR enhancement by array gain is a technique used to improve the quality and reliability of wireless communication in line-of sight propagation scenarios. Modern wireless communication systems often implement large-scale antenna arrays to take advantage of the benefits offered by array gain [14]. To improve the SNR level and consequently wireless communication, we proposed the use of an array antenna instead of the single antenna into the drone side [15], [16]. The SNR is calculated and compared for both SISO and SIMO cases using Eq. (2) and Eq. (3) respectively.

$$SNR_{siso} = \frac{P_t G_t G_r \lambda^2}{4\pi d^2 N_0} \quad (2)$$

Where P_t, G_t, G_r, and λ are transmitted power, Transmitted gain, Received gain and wavelength of transmitted signal respectively. The d and N_0 also stands for distance between transmitter-receiver, and the noise power.

$$SNR_{siso} = \frac{P_t G_t \sum G_{ri} \lambda^2}{4\pi d^2 N_0} \quad (3)$$

Where G_{ri} is the gain of the i-th receiving antenna.

The result of these calculations has been presented in a comparison way showing the energy per bit to noise power spectral density ratio. This technique is commonly used as a metric to quantify the quality of a communication system in terms of the received signal power and the noise power affecting the transmission.
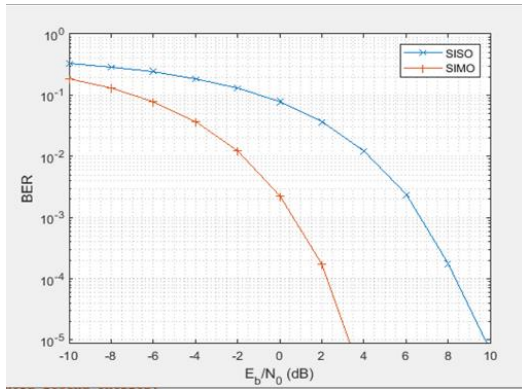
Fig. 2 Energy per Bit to Noise Power Spectral Density Ratio (E/N) Comparison for SISO and SIMO Systems

Due to the LOS communications, the transmitter and receiver are in direct communication. For the ease of calculation, a four-element Uniform Linear Array (ULA) with half-wavelength spacing, is considered in our simulations. The SISO and SIMO channel BER and the energy per bit to noise power spectral density ratio is depicted in Fig. 2. In a SIMO system, the BER curve demonstrates a notable improvement of 6 dB when a receive array is utilized. This gain is primarily attributed to the coherent nature of the received signals that interact with the components of the receive array. As a result, the receive array can be strategically oriented towards the transmitter, leading to an enhanced Signal-to-Noise Ratio (SNR). This phenomenon indicates that the receiver has knowledge of the incoming signal path, allowing for improved performance in terms of signal quality and error rate. The observed gain can be explained by the fact that the received signals around the receive array components exhibit coherence. This coherence arises due to the spatial diversity provided by multiple antennas in the receive array. Each antenna captures multiple instances of the transmitted signal, with slight variations in phase and amplitude. When these signals are combined, they undergo constructive interference, leading to an increase in the overall received signal power. This phenomenon highlights the advantage of utilizing a receive array with multiple antennas, as it enables the system to effectively capture and utilize the coherent signals, resulting in improved signal strength.

4. Conclusion

In summary, this research paper emphasizes the significant role of Unmanned Aerial Vehicles (UAVs) in wireless networks, offering a wide range of applications such as aerial base stations, wireless network user services, and mobile relays in flying ad-hoc networks. Deploying UAV base stations can greatly enhance wireless network coverage and power, enabling communication in various scenarios, including those where the dissemination of public safety information is critical. Furthermore, UAVs show promise in supporting secure and energy-efficient millimeter-wave communications and Internet of Things (IoT) communications. When equipped with cellular links, drones require secure and low-latency communication with ground base stations. The coherent nature of received signals through the Single-Input Multiple-Output (SIMO) receive array presents an opportunity to direct the array towards the transmitter, thereby increasing the Signal-to-Noise Ratio (SNR). Analysis of the Bit Error Rate (BER) curve reveals a gain of 6 dB achieved through the utilization of the receive array. This gain indicates that the receiver has knowledge of the incoming signal path. These findings highlight the advantages of incorporating UAVs into wireless networks, showcasing their potential to enhance coverage, power, and communication capabilities. The ability to steer the receive array towards the transmitter, coupled with the coherent nature of received signals, contributes to improved SNR and overall system performance. Future research can focus on optimizing UAV deployments, refining beamforming techniques, and exploring additional applications where UAVs can be effectively utilized in wireless networks. By harnessing the capabilities of UAVs in these contexts, advancements can be made towards more secure, reliable, and efficient wireless communication systems.

References

1. R. P. Fernandes, J. A. Apolinário Jr, A. L. Ramos, and J. M. de Seixas, "Applying the majority voting rule in acoustic detection and classification of drones."
2. M. Mokayef, T. A. Rahman, R. Ngah, and M. Y. Ahmed, "Spectrum sharing model for coexistence between high altitude platform system and fixed services at 5.8 GHz," Int. J. Multimed. Ubiquitous Eng, vol. 8, pp. 265-275, 2013.
3. M. A. Summakieh and M. Mokayef, "Single wideband microstrip patch antenna for 5G wireless communication application," Journal of Electrical andElectronics Engineering (IJR DO) Volume-1, Issue-4, Paper-2, 2016.
4. X. Ren, S. Vashisht, G. S. Aujla, and P. Zhang, "Drone-edge coalesce for energy-aware and sustainable service delivery for smart city applications," Sustainable Cities and Society, vol. 77, p. 103505, 2022.
5. I. Gómez-Arista, J. A. Dávila-Pintle, N. Montalvo-Montalvo, A. A. Rubin-Alvarado, Y. E. Bravo-García, and E. Reynoso-Lara, "Fourier Coefficients Applied to Improve Backscattered Signals in A Short-Range LIDAR System," Electronics, vol. 9, no. 3, p. 390, 2020.
6. D. Mukhutdinov, A. Alex, A. Cavallaro, and L. Wang, "Deep learning models for single-channel speech enhancement on drones," IEEE Access, vol. 11, pp. 22993-23007, 2023.
7. Z. A. Khan, S. Amjad, F. Ahmed, A. M. Almasoud, M. Imran, and N. Javaid, "A Blockchain-Based Deep-Learning-Driven Architecture for Quality Routing in Wireless Sensor Networks," IEEE Access, vol. 11, pp. 31036-31051, 2023.
8. A. C. Eska, "Cellular Communication Propagation at Drone around Building Environment with Single Knife Edge at 10 GHz," JURNAL INFOTEL, vol. 13, no. 1, pp. 25-30, 2021.

9. G. Raja, S. Anbalagan, A. Ganapathisubramaniyan, M. S. Selvakumar, A. K. Bashir, and S. Mumtaz, "Efficient and secured swarm pattern multi-UAV communication," IEEE Transactions on Vehicular Technology, vol. 70, no. 7, pp. 7050-7058, 2021.

10. F. Ronaldo, D. Pramadihanto, and A. Sudarsono, "Secure communication system of drone service using hybrid cryptography over 4g/lte network," in 2020 International Electronics Symposium (IES), 2020: IEEE, pp. 116-122.

11. B. Vergouw, H. Nagel, G. Bondt, and B. Custers, "Drone technology: Types, payloads, applications, frequency spectrum issues and future developments," The Future of Drone Use: Opportunities and Threats from Ethical and Legal Perspectives, pp. 21-45, 2016.

12. P. Garg, "Characterisation of Fixed-Wing Versus Multirotors UAVs/Drones," Journal of Geomatics, vol. 16, no. 2, pp. 152-159, 2022.

13. G. Phiri, M. Mokayef, S. S. Tiang, and W. C. Hong, "Drone Performance Analysis Based on SNR Factor," 2022.

14. L. N. Ribeiro, A. L. de Almeida, and J. C. Mota, "Low-rank tensor MMSE equalization," in 2019 16th International Symposium on Wireless Communication Systems (ISWCS), 2019: IEEE, pp. 511-516.

15. M. Khalily, M. R. Kamarudin, M. Mokayef, S. Danesh, and S. E. A. Ghahferokhi, "A new wideband circularly polarized dielectric resonator antenna," Radioengineering, vol. 23, no. 1, pp. 175-180, 2014.

16. A. A. Aldair, M. T. Rashid, A. F. Halihal, and M. Mokayef, "Design of pitch angle controller for wind turbine based on pi neurofuzzy model," Indonesian Journal of Electrical Engineering and Computer Science, vol. 15, no. 3, pp. 1664-1670, 2019

Authors Introduction

Dr. Mastaneh Mokayef

She has received her PhD from Wireless Communication Centre Faculty of Electrical Engineering in University Technology Malaysia (UTM) in 2014. She has also obtained her master's degree from the faculty of engineering in 2009 from the University Technology Malaysia. She has been working in UCSI University, Malaysia, since 2015 in which she currently serves as an Assistant Professor in the Faculty of Engineering and Built Environment. Her research interests include: Wireless communications, spectrum sharing method, spectrum management, etc.

Mr. MHD Amen Summakieh

He received the B.Eng. degree (Hons.) in communication and electronics engineering from UCSI University, Malaysia, in 2016, and the M.Eng.Sc. degree from Multimedia University, Malaysia, in 2020. His research interests include heterogeneous LTE-advanced cellular networks, user association, metaheuristic algorithms, and antennas design.

Dr. M. KA. Ahamed Khan

He completed his undergraduate degree in Electronics and Communication Engineering and postgraduate degree in Electronics and Control Engineering in India. Additionally, he obtained an advanced diploma in Power Electronics and Drives from Lucas-Nuelle in Germany in 2002, and a Diploma in Drives and Controls from Woo Sun in Korea in 2014. He pursued a PhD in Robotics, Power Electronics, and Controls in the United States and holds certifications. He is a Senior member of the IEEE in the USA and a member of MIET in the UK.

Dr. Sew Sun Tiang

She is an Assistant Professor in Faculty of Engineering at UCSI University in Malaysia. She received her PhD in Electrical and Electronic Engineering from Universiti Sains Malaysia in 2014. Her research interests are optimization and antenna design.

Dr. Wei Hong Lim

He is an Assistant Professor in Faculty of Engineering at UCSI University in Malaysia. He received his PhD in Computational Intelligence from Universiti Sains Malaysia in 2014. His research interests are optimization and artificial intelligence.

Dr. Abdul Qayyum

He received his PhD in Electrical and Electronic Engineering from Universiti Teknologi Petronas, Malaysia, in 2018. He developed the totype. Previously, he completed his bachelor's degree in computer engineering and master's degree in electronic engineering from Pakistan. He is a postdoctoral researcher at the University of Burgundy, France, where he studies cardiac MRI images using deep learning approaches. He is currently affiliated with the National Heart and Lung Institute at Imperial College London, UK.

Empowering Elderly Individuals through the Intelligent Shopping Trolley

Mastaneh Mokayef[1*], Muzaiyanah Binti Hidayab[1], MHD Amen Summakieh[1], M.K.A Ahamed Khan[1], Kim Soon Chong[1], Chin Hong Wong[2], Chua Huang Shen[3], Abdul Qayyum[4]

[1]*Department of Electrical and Electronic Engineering, Faculty of Engineering Technology and Built Environment, UCSI University, Kuala Lumpur, 56000, Malaysia*
*Email: mastaneh@ucsiuniversity.edu.my
https://www.ucsiuniversity.edu.my/

[2]*Maynooth International Engineering College, Fuzhou University, Fujian, China*

[3]*Department of Electrical & Electronics Engineering, UOW Malaysia University Utropolis Glenmarie, Shah Alam, Malaysia*

[4]*National Heart and Lung Institute, Imperial College London, UK*

Abstract

In this research, we have developed a prototype that aims to improve the weekly shopping experience for senior citizens. Our system tracks the movements of elderly individuals, eliminating the need for them to exert physical force in pushing or pulling the shopping trolley. To achieve this, we utilize a combination of sensors, including a gyroscope and magnetometer, to estimate the user's walking distance and direction. Additionally, we employ WiFi fingerprinting to accurately determine the user's position. Our experiments have yielded satisfactory results in terms of tracking accuracy and the overall functionality of the system. By addressing the specific challenges faced by senior citizens during the routine and essential process of grocery shopping, our smart shopping trolley concept seeks to enhance their experience. Through the integration of tracking technology and sensor-based solutions, we aim to make shopping more convenient and comfortable for elderly individuals. The positive outcomes observed in our experiments validate the effectiveness and feasibility of this approach.

Keywords: smart shopping trolley, elderly citizen, smart phone, UML

1. Introduction

Senior citizens often encounter challenges when it comes to shopping, particularly in the case of grocery shopping. While there are options available such as helpers or online shopping, many elderly individuals prefer to do their own shopping but still require assistance in carrying their bags back home. The advancement of technology has sparked interest in integrating different systems, resulting in the rise of the Internet of Things (IoT) and its widespread use in wireless communication.

The term "smart" has become increasingly common in various electronic devices and technologies, such as smartphones, smart cars, smartwatches, and even smart homes. In line with this trend, the concepts of "smart shopping" and "smart trolleys" have emerged to provide valuable support to senior citizens. Smart trolleys often incorporate Radio Frequency Identification (RFID) technology, enabling products to be scanned as they are placed in the trolley. By utilizing ZIGBEE communication, the collected data can be transmitted to the counter, reducing the time spent waiting in queues. Moreover, the use of RFID technology in smart trolleys can enhance security by facilitating cashless transactions.

Additionally, the successful integration of purchased items, payment systems, and data collection relies on the implementation of IoT standards. This ensures smooth connectivity and compatibility among the various components involved in the shopping process. The Internet of Things (IoT) field is experiencing rapid growth in the world of wireless telecommunications [1]. Its purpose is to enable computers to interact with their surroundings and access information about devices and objects without the need for direct human involvement [2]. According to a report [3], IoT is characterized by moderate data transmission rates, affordability, and operates within the frequency range of 100MHz to 5.8GHz. This advancement in wireless technology offers users a flexible and cost-effective means of communication. A recent report [4], [5] suggests that by the year 2020, around 50 billion devices will be

Mastaneh Mokayef, Muzaiyanah Binti Hidayab, MHD Amen Summakieh, M.K.A Ahamed Khan, Kim Soon Chong, Chin Hong Wong, Chua Huang Shen, Abdul Qayyum.

connected to the internet, opening up a plethora of functionalities. There is a rising demand for technology and a growing need for convenient shopping experiences, which has spurred increased research efforts in the development of smart shopping concepts. These trolleys are equipped with core functionalities such as following the user, navigating through the shopping paths, and avoiding collisions with obstacles. This technology aims to eliminate the physical strain and difficulties that elderly individuals often face while shopping, resulting in a more comfortable and enjoyable experience. Advancements in autonomous navigation have been a key focus in the development of smart shopping trolleys. Researchers have been exploring various approaches to enable autonomous navigation for smart shopping trolleys. For example, a new automated smart cart system for modern shopping centers was presented, along with Android application-controlled approaches [1]. These advancements in autonomous navigation provide the trolleys with the ability to move independently, allowing elderly individuals to navigate the store without physical exertion. The implementation of smart shopping trolleys for elderly care holds several key advancements and benefits. Other recent reports in [6], [7] describe a smart trolley system for mega malls that uses a microcontroller. It allows customers to automatically scan products with a barcode scanner and see running totals on an LCD display, without needing to queue for checkout.

In this paper, the novel smart shopping trolley has been introduced. The proposed smart trolley is designed to avoid the pushing the typical shopping trolley. It is also designed to follow the user to ease of the controlling process for the user.

2. Hardware Interfaces

The smart shopping trolley designed specifically for elderly care integrates various essential components to improve its functionality and user-friendliness. These components are crucial in creating a smooth and convenient shopping experience for elderly individuals. The following are the components utilized in this project:

2.1. *Micro Controller Unit (NodeMCU)*

The NodeMCU micro-controller is used in this project as the main controller which has the integrated WiFi module. NodeMCU is an open source Lua based firmware for the ESP8266 WiFi SOC from Espressif and uses an on-module flash-based SPIFFS file system. NodeMCU is implemented in C that ease the use of this component in IoT based projects.

2.2. *DC motor*

A utilized DC motor would have only two terminals. Since these terminals are connected together only through a coil, they do not have a polarity. Revering the connection will only reverse the direction of the motor.

2.3. *Inertial gyro sensor*

In our proposed design, we have included a special sensor called a gyroscope to help us keep track of how something is rotating. This sensor is pretty clever because it also has a built-in accelerometer, which measures how fast something is moving in a straight line. So, while the accelerometer tells us about the linear movement of the user, the gyroscope tells us about the angular rotation. By combining these two sets of information, we can accurately detect and understand how the user is moving. Another great thing about this sensor is that it's small and doesn't cost too much, which makes it a perfect fit for our project. We wanted to make sure we chose components that are both effective and affordable.

2.4. *H-Bridge dual motor controller L298N*

In order to control the rotation direction of the motor, the current flow through the motor must be inverted. The most common method of doing that is by using an H-Bridge circuit. An H-Bridge circuit contains four switching elements, transistors or MOSFETs, with the motor at the center forming an H-like configuration. By activating two switches at the same time, the direction of the current flow and as a result the rotation direction of the motor will be changed. In this project, the L298N is used as a dual H-Bridge motor driver which allows speed and direction control of two DC motors at the same time.

3. Design and fabrication

In the world of shopping trolleys, we've seen various types introduced over the years. However, our project takes things a step further with the development of a smart shopping trolley that brings convenience to a whole new level. Imagine being able to control your shopping trolley using your smartphone, thanks to Wi-Fi technology. Through a real-time Wi-Fi connection between the trolley's micro-controller and your smartphone, you can effortlessly guide the trolley to your desired destination. All it takes is a few taps on a dedicated app installed on your smartphone to send commands to the micro-controller via the Wi-Fi signal. It's like having a personal assistant for your shopping needs.

But convenience isn't the only thing we focused on. We also wanted to ensure the safety of both the user and the trolley. That's why we integrated an ultrasonic sensor right at the front of the trolley. This little hero plays a crucial role in preventing accidents and damage caused by sudden stops or collisions. It works by emitting sound waves and then listening for the echoes that bounce back. By analyzing the time it takes for the sound waves to travel and return, the ultrasonic sensor can accurately

measure the distance between the trolley and any obstacles in its path. This allows the trolley to adjust its movement and avoid potential collisions by utilizing the signals it sends and receives. So, in a nutshell, our smart shopping trolley combines the convenience of smartphone-controlled Wi-Fi technology with the added peace of mind provided by the ultrasonic sensor. With this integration, we aim to enhance your shopping experience by giving you seamless control and ensuring a safe journey through the store, effortlessly avoiding obstacles along the way. It's like having a reliable shopping companion right by your side. The operation of ultrasonic sensor has been shown in Fig. 1.
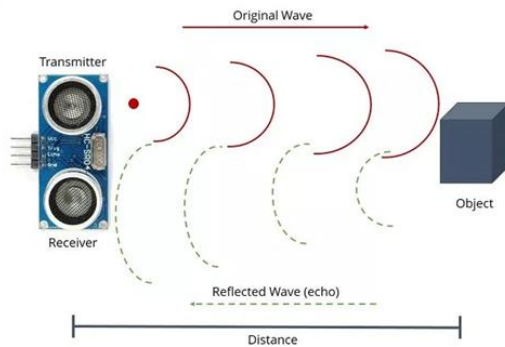
Fig.1 Obstacle detection via ultrasonic sensor

To establish the necessary connections for the ultrasonic sensor in the trolley, we connect the VCC (power supply) and GND (ground) pins to a 5V power source. Additionally, the trigger input (Trig) pin is linked to a digital output, while the echo (Echo) pin is connected to a digital input on the trolley's microcontroller.

To accurately measure the distance between the trolley and an obstacle, we employ a technique that involves pulsing the trigger pin to a high level for approximately 10 microseconds. After that, we wait for a high-level signal on the echo pin. The duration of this high-level signal corresponds to the time it takes for the ultrasonic sound to travel to the obstacle and back. By analyzing the response time and duration, we can determine the proximity of the trolley to the obstacle. This allows for precise distance measurement and enables the trolley to effectively detect and avoid obstacles. The distance can be calculated as follows:

$$\text{Distance}= \text{Time}\times\text{Speed}=\text{time}\times0.034\text{cm/µs} \quad (1)$$

As sound travels at approximately 340 meters per second. This corresponds to about 29.41 µs.

The circuit design has been shown in Fig. 2. As shown in this figure, the MCU, ultrasonic sensor, two motors and motor controller are connected together as the main components of our proposed design.
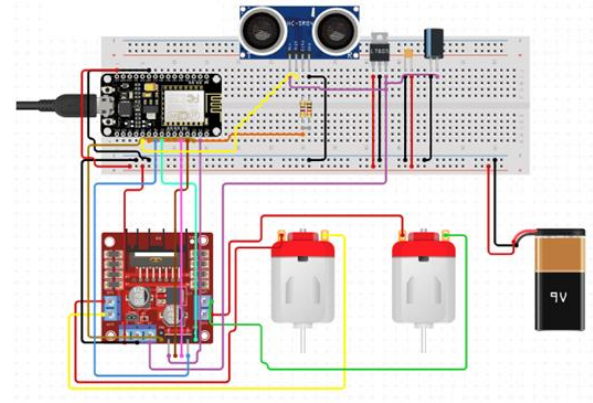
Fig.2 Circuit design for the proposed trolley

The Node MCU, known for its energy efficiency and 16-bit RSIC, is utilized as a controller because of its integrated ESP-12 and convenient WiFi module, which streamline the control process. The ESP-12 incorporates a CPU clock speed of 80MHz, with the ability to reach a maximum value of 160MHz. The L298N motor controller is employed to connect the DC motors to the microcontroller. Both the microcontroller and L298N motor controller are powered by a 9V battery in parallel. The controller has been programmed to define the movements of the trolley, as outlined in Table 1.

Table 1. The motor movement instruction

Movement	Left motor forward	Left motor backward	Right motor forward	Right motor backward
Forward	High	Low	High	Off
Backward	Low	High	Low	High
Left	-	-	High	Low
Right	High	Low	-	-

The controlling instructions are sent to the microcontroller through the app created by MIT app inventor. The block function of the created app is shown in Fig. 3.

Mastaneh Mokayef, Muzaiyanah Binti Hidayab, MHD Amen Summakieh, M.K.A Ahamed Khan, Kim Soon Chong, Chin Hong Wong, Chua Huang Shen, Abdul Qayyum.
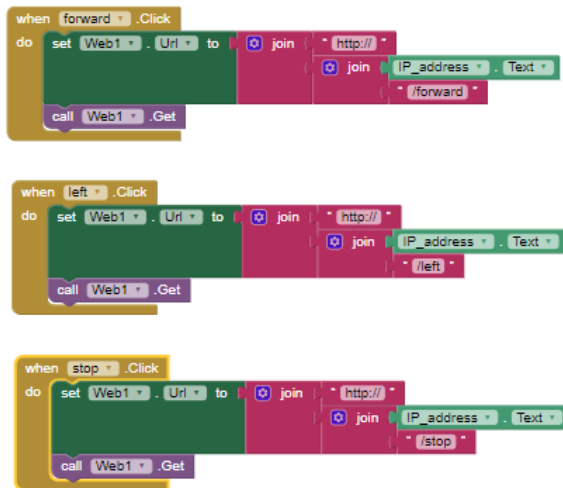
Fig.3 The controlling function block of the app

The fabrication of our proposed prototype is shown in Fig. 4.
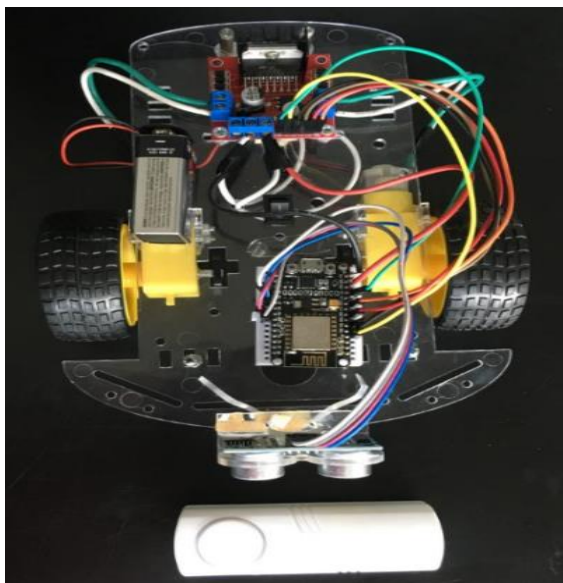
Fig.4 The fabricated prototype

4. Conclusion and Future Work

In conclusion, we have successfully designed, fabricated, and tested a smart shopping trolley specifically tailored for elderly care. The results have been highly satisfactory, indicating the effectiveness of our proposed solution. By implementing advanced technology, our smart trolley is designed to recognize the movement patterns of elderly users and autonomously follow them. This eliminates the need for users to exert physical effort in pushing or pulling the trolley, providing a more convenient and comfortable shopping experience. Looking ahead, there is potential for further improvement and expansion of our smart shopping trolley. One possible avenue for future work is the integration of GPS technology. By incorporating GPS, the trolley can accurately determine its indoor and outdoor location,

allowing users to easily navigate and even return the trolley to its designated location rather than being limited to indoor use only. This enhancement would offer even greater convenience and flexibility to the users, making their shopping trips more efficient and enjoyable. In summary, our smart shopping trolley for elderly care has shown promising results, offering a user-friendly and assistance-driven approach to shopping. The inclusion of GPS technology in future iterations has the potential to further enhance the functionality and versatility of the trolley, providing an improved and seamless shopping experience for elderly individuals.

References

1. V. Shete, A. Shaikh, R. Sawant, C. Singh and P. Srideviponmalar. "Smart Shopping Cart for Visually Impaired Individuals". Aug. 2022.
2. P. S. Yadav, V. Agrawal, J. C. Mohanta and M. F. Ahmed. "A Theoretical Review of Mobile Robot Locomotion based on Mecanum Wheels". Jun. 2022.
3. H. Bello-Salau, A. J. Onumanyi, M. David, R. Isa, C. O. Alenoghena and H. Ohize, "A new automated smart cart system for modern shopping centres".
4. Summakieh, MHD Amen, and Mastaneh Mokayef. "Single wideband microstrip patch antenna for 5G wireless communication application." Journal of Electrical and Electronics Engineering (IJR DO) Volume-1, Issue-4, Paper-2 (2016).
5. R. O'Neill. (Jun. 21, 2005). Smart Trolley Shops for You. Accessed: Jun. 21, 2017.
6. Aldair, Ammar A., et al. "Design of pitch angle controller for wind turbine based on pi neuro-fuzzy model." Indonesian Journal of Electrical Engineering and Computer Science 15.3 (2019): 1664-1670.
7. Patil, Shishir R., et al. "Smart Trolley with Automatic Billing System using Arduino." IAES International Journal of Artificial Intelligence (IJ-AI) 2.2 (2022): 2268-2273.

Authors Introduction

Dr. Mastaneh Mokayef

She has received her PhD from Wireless Communication Centre Faculty of Electrical Engineering in University Technology Malaysia (UTM) in 2014. She has also obtained her master's degree from the faculty of engineering in 2009 from the University Technology Malaysia. She has been working in UCSI University, Malaysia, since 2015 in which she currently serves as an Assistant Professor in the Faculty of Engineering and Built Environment. Her research interests include: Wireless communications, spectrum sharing method, etc

Mr. MHD Amen Summakieh

He received the B.Eng. degree (Hons.) in communication and electronics engineering from UCSI University, Malaysia, in 2016, and the M.Eng.Sc. degree from Multimedia University, Malaysia, in 2020. His research interests include heterogeneous LTE-advanced cellular networks, user association, metaheuristic algorithms, and antennas design.

Dr. M. KA. Ahamed Khan

He completed his undergraduate degree in Electronics and Communication Engineering and postgraduate degree in Electronics and Control Engineering in India. Additionally, he obtained an advanced diploma in Power Electronics and Drives from Lucas-Nuelle in Germany in 2002, and a Diploma in Drives and Controls from Woo Sun in Korea in 2014. He pursued a PhD in Robotics, Power Electronics, and Controls in the United States and holds certifications. He is a Senior member of the IEEE in the USA and a member of MIET in the UK.

Dr. Kim Soon Chong

He obtained his B.Eng., M.S., and Ph.D. degrees in electrical and electronic engineering from Universiti Kebangsaan Malaysia in 2012, 2015, and 2019, respectively. He currently serves as an Assistant Professor in the Department of Computer Engineering (Artificial Intelligence) within the Faculty of Engineering Technology and Built Environment at UCSI Kuala Lumpur. His research focuses on areas such as medical signal processing and instrumentation.

Dr. Chin Hong Wong

He is a Lecturer in Maynooth International Engineering College at Fuzhou University in China. He received his PhD in Electrical and Electronic Engineering from Universiti Sains Malaysia in 2017. His research interests are Energy harvesting and control system.

Mr. Chua Huang Shen (Mason)

He holds the position of Senior Lecturer at UOW Malaysia KDU. He received a Master of Science degree from the University of Belfast in 2004 and a Bachelor of Science degree (Hons) in Electronic Control Systems from the Moore School of Liverpool University in 2002. - Earned a bachelor's degree in engineering from Rahman University (UTAR). He is currently working on his Ph.D.

Dr. Abdul Qayyum

He received his PhD in Electrical and Electronic Engineering from Universiti Teknologi Petronas, Malaysia, in 2018. He developed the totype. Previously, he completed his bachelor's degree in computer engineering and master's degree in electronic engineering from Pakistan. He is a postdoctoral researcher at the University of Burgundy, France, where he studies cardiac MRI images using deep learning approaches. He is currently affiliated with the National Heart and Lung Institute at Imperial College London, UK.

Optimized Microstrip Slot UWB Patch Antenna for Medical Imaging

Maxime DUVACHER[2], **Mastaneh Mokayef**[1*], **MHD Amen Summakieh**[1], **M.K.A Ahamed Khan**[1], **Sew Sun Tiang**[1], **Wei Hong Lim**[1], **Abdul Qayyum**[3]

[1]*Faculty of Engineering Technology and Built Environment, UCSI University,*
Kuala Lumpur, 56000, Malaysia
[*]Email: mastaneh@ucsiuniversity.edu.my
https://www.ucsiuniversity.edu.my/

[2]*Polytech Nantes, Rue Christian Pauc, 44300 Nantes, France*

Abstract

This research work presents the development of an Ultra-Wideband (UWB) microstrip patch antenna (MPA) with the specific purpose of tissue characterization. The antenna was carefully designed and simulated to operate within a frequency range of 4.8 to 6.9 GHz, optimized for its intended application. To ensure the best performance, a series of simulations and comparisons were conducted using CST Microwave Studio. Various antenna shapes were tested and evaluated to determine the most effective design. The results of these simulations were highly promising, as they revealed a simulated return loss (S_{11}) of -33dB. This indicates excellent performance and demonstrates the suitability and acceptability of the proposed antenna for medical imaging such as breast imaging, tumor detection, or monitoring physiological changes.

Keywords: Microstrip antenna, medical imaging, UWB, CST

1. Introduction

In recent years, the use of microstrip antennas in medical imaging has attracted significant attention. These antennas have demonstrated immense potential in delivering high-resolution images and precise measurements in various medical imaging techniques like magnetic resonance imaging and positron emission tomography. According to Shimu et al., microstrip patch antennas have found wide-ranging applications in the medical field, as well as in other industries [1]. Microstrip antennas offer several advantages that make them well-suited for medical imaging applications. Firstly, microstrip antennas have a small size and low profile, which makes them ideal for integration into medical devices and equipment. Their lightweight nature allows for easy customization to meet specific requirements. Additionally, microstrip antennas can be designed with wide bandwidths, enabling the transmission and reception of diverse frequencies. Microstrip patch antennas also provide high gain and low return loss, ensuring efficient signal transmission and reception [1, 2]. Another advantage of microstrip patch antennas in medical imaging lies in their robust design and fabrication. These antennas are known for their durability and reliability, which are crucial factors in medical imaging where accuracy and consistency are of paramount importance. Furthermore, microstrip antennas exhibit excellent radiation characteristics, including low cross-polarization and high directivity. These characteristics are particularly valuable in medical imaging, as they contribute to clearer and more accurate imaging results. With these advantages, microstrip antennas have the potential to revolutionize medical imaging by providing high-quality images with improved resolution and accuracy. However, it is important to address the limitations associated with microstrip antennas in order to fully harness their potential in medical imaging.

One limitation of microstrip antennas is their narrow bandwidth, which restricts the range of frequencies that can be transmitted or received. Researchers are actively exploring different techniques, such as adding additional resonant elements or using multiband designs, to increase the bandwidth of microstrip antennas. Additionally, microstrip antennas often have relatively low gain compared to other types of antennas [3]. To address this, researchers are investigating techniques such as designing stacked or array configurations to enhance the gain of microstrip antennas. Impedance matching is another issue that microstrip antennas face, leading to signal degradation. Researchers are focusing on improving the manufacturing process and developing advanced impedance-matching techniques, such as using matching networks or adding dielectric layers, to optimize impedance matching and improve overall performance. Furthermore, efforts are being made to enhance the polarization characteristics of microstrip antennas, as certain imaging techniques require specific

polarization orientations for optimal image quality. Researchers are exploring various polarization techniques, such as circularly polarized microstrip antennas or the use of metasurfaces, to control the polarization of the antenna [4]. The study in [1] presents the design, fabrication, and analysis of a compact microstrip patch antenna operating in the Industrial, Scientific, and Medical (ISM) frequency range of 2.4-2.5 GHz for medical applications. To investigate the performance of this antenna structure through simulations, the high-frequency structure simulator (ANSYS HFSS) and Advanced Design System (ADS) has been used. Moreover, to measure the level of electromagnetic waves absorbed by the human head model the specific absorption rate (SAR) is simulated at different positions. A 2.45 GHz Rectangular Patch Microstrip Antenna (RMSA) was designed and analyzed for breast cancer detection in [5], using an insect feed technique for impedance matching. Simulations showed the RMSA had a VSWR<2 bandwidth of 70 MHz (2.8% of operating frequency), resonating at 2.45 GHz. Results indicate the RMSA outperforms existing designs, making it a promising option for improved breast tumor detection. While in [7, 8] a low-cost microstrip patch antenna was developed specifically for microwave imaging (MWI) applications across a wideband frequency range. The antenna design incorporates an artificial magnetic conductor (AMC) to enhance its performance. Computer simulations using CST demonstrated that the presence of the AMC significantly improved the frequency selectivity at 8.6 GHz, resulting in a peak realized gain of 9.90443 dBi (simulated) and 10.61 dBi (measured). To validate the simulation results, the proposed microstrip antenna was fabricated and experimentally tested.

2. Antenna Structure

In this paper, we introduce a newly proposed Ultra-Wideband microstrip antenna and provide an overview of its structure. The decision to utilize an FR-4 substrate is driven by its cost-effectiveness and wide accessibility in the market. The simplicity and convenience of implementing FR-4 as a substrate make it an ideal choice for the design of our antenna. In the upcoming sections, we will delve into the fabrication process and design specifics of the microstrip antenna, emphasizing its potential for Ultra-Wideband applications.

As depicted in Fig. 1, the antenna is mounted on the FR-4 substrate including the main and sub-fluidline for the impedance matching purpose. The detailed dimensions of the design can be found in Table 1.
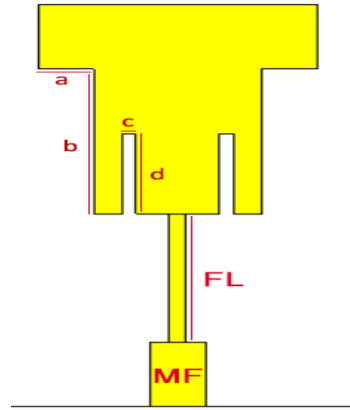
Fig.1 Front View of the Proposed Antenna

Table 1. Antenna Dimensions

Area	Parameter	Dimensions (mm)
PATCH (copper)	Width	10
	Length	13
	a	2
	b	9
	c	0.5
	d	5
	Thickness	0.035
FEED LINE (FL) (copper)	Width	0.6
	Length	8
	Thickness	0.035
MAIN LINE (ML) (copper)	Width	2
	Length	4
	Thickness	0.035
SUBSTRATE (FR-4)	Width	28
	Length	29
	Thickness	1.6
GROUND (copper)	Width	28
	Length	10.5
	Thickness	0.035

3. Results and Discussion

3.1. *Accuracy of numerical integration*

As defined in [4], the UWB has a bandwidth of greater than 500MHz and an operating frequency of between 3.1GHz and 10.6GHz. The return loss threshold of -10dB has been defined in [5]. A simulation of return loss S_{11} is shown in Fig. 2.
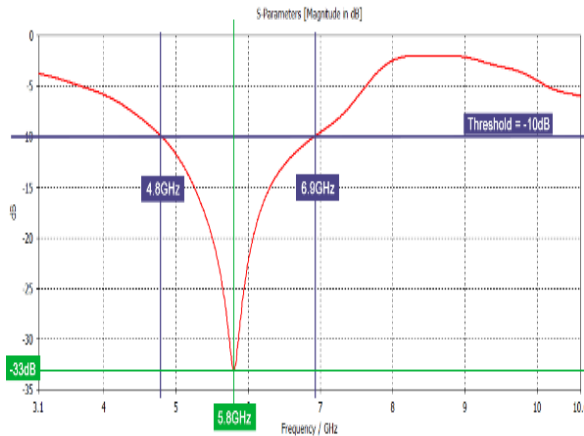
Fig.2 Return Loss of the Proposed Antenna

The simulation and expected results match well. The threshold is pointed at -10dB, so the bandwidth is 2.1GHz, in agreement with the 500MHz minimum required by UWB. Furthermore, the resonance frequency is 5.8GHz. This one is great, the signal propagation characteristics at 5.8GHz in free space are similar in rainy conditions. This issue is important to consider in tropical climates such as Malaysia [6]. The radiation pattern of the proposed antenna is illustrated in Fig. 3.
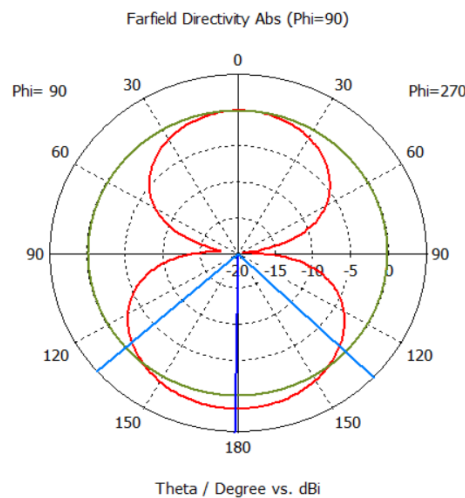
Fig. 3 Antenna's radiation pattern

As shown in Fig. 3, the main lobe magnitude is 1.68 dBi and the main lobe direction of 179 deg has been obtained. The Angular width (3dB) of 94.8 degrees and the side lobe level of-1.7 dB has been obtained. Finally, the results of the VSWR simulation are given in Fig. 4.
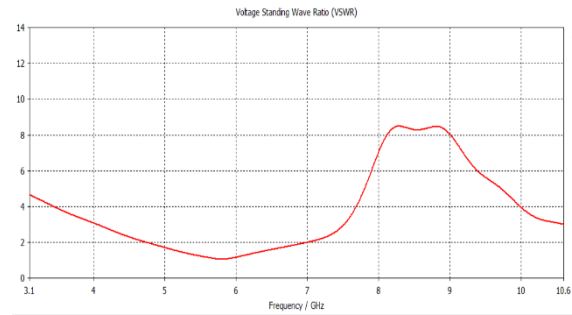
Fig. 4. Simulation of VSWR

The acceptable range of Voltage Standing Wave Radio (VSWR) is under 2V. So the acceptable frequency of between 4.7 and 7 GHz are the acceptable range which matches well with the previous results obtained in the operating frequency from 4.8 to 6.9 GHz.

4. Conclusion

In conclusion, this project successfully proposed an Ultra-wideband (UWB) microstrip patch antenna (MPA) using an FR-4 substrate. The achieved return loss of -33dB at a resonance frequency of 5.8GHz demonstrates the effectiveness of the antenna design. With dimensions of 29×28mm, this antenna offers a cost-effective, user-friendly, and portable solution for various applications, particularly in breast cancer detection. Through meticulous design and simulation using CST Microwave Studio, the antenna was optimized to operate within the desired frequency range of 4.8 to 6.9 GHz. The extensive simulations and comparisons performed confirmed the superior performance of the proposed design. The obtained results, including the impressive return loss of -33dB, highlight the antenna's suitability for medical imaging, such as breast imaging, tumor detection, or monitoring physiological changes.

References

1. Alhuwaidi, Sadiq, and Tanghid Rashid. "A novel compact wearable microstrip patch antenna for medical applications." 2020 International Conference on Communications, Signal Processing, and their Applications (ICCSPA). IEEE, 2021.
2. Borkar, Nishant Madhukar and P. K. Parlewar. "Antennas in Medical Applications." 2021 IEEE Indian Conference on Antennas and Propagation (InCAP) (2021): 152-154.
3. S. Alhuwaidi and T. Rashid, "A Novel Compact Wearable Microstrip Patch Antenna for Medical Applications," 2020 International Conference on Communications, Signal Processing, and their Applications (ICCSPA), Sharjah, United Arab Emirates, 2021, pp. 1-6, doi: 10.1109/ICCSPA49915.2021.9385726.
4. Khalily, Mohsen, et al. "A new wideband circularly polarized dielectric resonator antenna." Radioengineering 23.1 (2014): 175-180.
5. SM, Asha Banu, and Niresh Kumar. "Design of Microstrip Patch Antenna for the Detection of Malignant Tumor."

M. Mokayef et al.

2022 4th International Conference on Inventive Research in Computing Applications (ICIRCA). IEEE, 2022.

6. Ahmed, M. Firoz, and Md Hasnat Kabir. "Multislot Rectangular Microstrip Patch Antenna Design to Detect Breast Cancer." ECS Transactions 107.1 (2022): 10721.

7. Summakieh, MHD Amen, and Mastaneh Mokayef. "Single wideband microstrip patch antenna for 5G wireless communication application." Journal of Electrical andElectronics Engineering (IJR DO) Volume-1, Issue-4, Paper-2 (2016).

8. Hamza, Musa N., et al. "Low-Cost Antenna-Array-Based Metamaterials for Non-Invasive Early-Stage Breast Tumor Detection in the Human Body." Biosensors 12.10 (2022): 828.

Authors Introduction

Dr. Mastaneh Mokayef

She has received her PhD from Wireless Communication Centre Faculty of Electrical Engineering in University Technology Malaysia (UTM) in 2014. She has also obtained her master's degree from the faculty of engineering in 2009 from the University Technology Malaysia. She is a member of Board of Engineers Malaysia (BEM) since 2017, She has been working in UCSI University, Malaysia, since 2015 in which she currently serves as an Assistant Professor in the Faculty of Engineering and Built Environment (FETBE). Her research interests include: Wireless communications, spectrum sharing method, spectrum management, cellular communication systems and Antenna design. To date, he has been awarded with the qualifications of a Chartered Engineer (C.Eng.) from U.K. Engineering Council.

Mr. MHD Amen Summakieh

He received the B.Eng. degree (Hons.) in communication and electronics engineering from UCSI University, Malaysia, in 2016, and the M.Eng.Sc. degree from Multimedia University, Malaysia, in 2020. His research interests include heterogeneous LTE-advanced cellular networks, user association, metaheuristic algorithms, and antennas design.

Dr. M. KA. Ahamed Khan

He completed his undergraduate degree in Electronics and Communication Engineering and postgraduate degree in Electronics and Control Engineering in India. Additionally, he obtained an advanced diploma in Power Electronics and Drives from Lucas-Nuelle in Germany in 2002, and a Diploma in Drives and Controls from Woo Sun in Korea in 2014. He pursued a PhD in Robotics, Power Electronics, and Controls in the United States and holds certifications as a Professional Engineer (PEng) in the USA and a Chartered Engineer (CEng) in the UK. He is a Senior member of the IEEE in the USA and a member of MIET in the UK.

Dr. Sew Sun Tiang

She is an Assistant Professor in Faculty of Engineering at UCSI University in Malaysia. She received her PhD in Electrical and Electronic Engineering from Universiti Sains Malaysia in 2014. Her research interests are optimization and antenna design.

Dr. Wei Hong Lim

He is an Assistant Professor in Faculty of Engineering at UCSI University in Malaysia. He received his PhD in Computational Intelligence from Universiti Sains Malaysia in 2014. His research interests are optimization and artificial intelligence.

Dr. Abdul Qayyum

He did his PhD in Electrical and Electronics Engineering from Universiti Teknologi PETRONAS, Malaysia 2018. He developed deep learning based algorithms for depth estimation of vegetation, trees near power lines for Tenaga Nasional Berhad (TNB) and Sabah Electric Supply Berhad (SESB) under the ministry of Green , Water and Technology (KeTTHA) Malaysia. He also developed a prototype for Vital signs (heart rate, breathing rate, Sp02) estimation and assessment of stroke and Arterial fibrillation (AF) using face video analytic based on deep leaning models. Earlier, he had completed his bachelor's in computer engineering and Master in Electronic Engineering from Pakistan. Besides, he gained one-year industrial experience while working as a BSS engineer for Huawei, Pakistan. He had also taught several courses under electrical, specifically, signal processing domain for 7 years in various public and private universities in Pakistan. He was working as a research scientist in CISIR, UTP for less than one year and was developed deep learning algorithms for brain signal (EEG) classification and reconstruction, remote sensing image segmentation and biomedical image analysis. He is associated with burgundy university France as a Post Doc researcher and working on the cardiac MRI images using deep learning approach. He is also working as a consultant in various projects involving deep learning models in Big Data, Vital Sign Estimation, IoT and BCI applications. Currently he is associated with National Heart and Lung Institute, Imperial College London, UK.

Development of an Innovative Undergraduate Industrial Automation and Robotics Degree Program

M. K. A. Ahamed Khan[1*], Mastaneh Mokayef[1], Ridzuan, A[1], Irraivan Elamvazuthi[2], Badli Shah Yusoff[3], Abu Hassan Darusman[3]

[1]UCSI University, Faculty of Engineering, Taman Connaught, 56000 Malayisa
*Email: Mohamedkhan@ucsiuniversity.edu.my

[2]University Technology Petronos, Faculty of Engineering, Malaysia

[3]NIKL Malaysia France Institute, Malaysia

Abstract

In recent years, the need for integrated engineering courses has increased. Due to its multidisciplinary nature, Industrial Automation and Robotics degree course is an ideal example of curriculum integration. This paper discusses several issues such as course offerings, topical content, student profile, student performance and other pertinent matters related to the recent development of an Industrial Automation and Robotics undergraduate degree programme at the University of Kuala Lumpur, Malaysia.

Keywords: Industrial Automation, Robotics, Curriculum Development, Undergraduate

1. Introduction

Industrial Automation and Robotics is a multidisciplinary science that encompasses the areas of computer science, mathematics, physics, mechanical engineering, industrial engineering, electrical engineering, computer engineering, materials science, and manufacturing engineering among others. It provides an opportunity to break the barriers between all these disciplines in a single course and it offers an excellent example of multidisciplinary integration for engineering students. In recent years, several institutions, following the lead of the robot-building course at the Massachusetts Institute of Technology [1], now offer robot building classes and laboratories. Some graduate level courses in Artificial Intelligence also rely on robot building techniques to illustrate general concepts of decision-making and machine intelligence [2], [3].

In Malaysia colleges, public and private universities offer many programs which lead to the degree of Bachelor of Engineering and Bachelor of Engineering Technology. These programmes range in topics from Electrical, Mechanical, Chemical, etc. With increasing frequency, employers are requiring employees in their engineering or technical staff to have multi-skills. Naturally, aspiring employees want to take-up inter-disciplinary courses to enhance their marketability. To help meet these desires, a new degree was developed by University Kuala Lumpur with the first batch of students enrolled in July 2002.

2. Planning Process

A number of authors have considered the broad issues associated with current and future trends in multidisciplinary engineering education [4], [5]. It is clear that there is no one correct curricula or approach to treating multidisciplinary engineering education. Rather, one must take into careful consideration the regional employment demands, as well as the preparation necessary for graduate study. It is necessary to practice careful discernment regarding the rapid technological fluctuations that are prevalent today. Specifically, industry can be expected to make regular demands of undergraduate preparation to meet their perceived needs [6], [7], [8]. Sometimes these expectations will provide valuable insights into changing career paths, and at other times the demands will be inappropriate, and best understood as related to the "tools and toys" of the trade. Engineering students must have a fundamental concept base which allows agility and flexibility, for if a program is too specialised, the range of opportunities available for graduates will be compromised. In the course of investigating the content and structure of a new program to address the design and development of the 'industrial automation and robotics' programme, it became clear that close ties to the fundamental disciplines associated with electrical engineering was critical. The wide variety of distributed sensors, their associated electronic interfaces, the large-scale control systems, and the extreme environments within which they operate require careful treatment of analog and digital electronic circuit design, communication and control theory, mathematics, and physics. These topics traditionally constitute the core of the electrical engineering discipline [9], [10].

A curriculum planning process is necessary to investigate the individual program topics, their relationships, and the pedagogy necessary to present them in an integrated fashion at an undergraduate level. We utilised two methods of inquiry to meet these goals. The first method of inquiry was based on a series of

formal meetings (both group and individual) with representatives drawn from our industrial partners and industrial advisory board. This varied group of people has been asked to consider their educational needs (industrial automation and robotics) based on the large and clearly identified trends in their industries. The second method involved the exploration of the substantial educational literature available including such publications as the IEEE Transactions on Education and the Journal of Engineering Education. At the end of the planning period, the following outcomes were generated:

- a compilation of the topical content for courses, particularly for topics that are new in the sense

curriculum so that appropriate acquisition of laboratory equipment could be addressed and the attendant facilities set up; and
- a time-bound roadmap for introducing the new curriculum

3. The Development of Bachelor of Engineering Technology Programme

The development of Bachelor of Engineering Technology Programme in Industrial Automation and Robotics Technology (IART) will discussed in the following paras. Once admitted into IART degree programme, the student takes a minimum of 120 semester

Table 1: Subject Distribution of IART

NO	CATEOGORY	SUBJECTS	CH	%	TCH	%
1	General	LAN	9	7.31	25	20.32
		English	6	4.88		
		French	2	1.63		
		Mathematics	8	6.50		
2	Management	Unikl	4	3.25	15	12.19
		Mfi	11	8.94		
3	Technical	Automation	32	26.02	77	62.61
		Electrical	32	26.02		
		Mechanical	13	10.57		
4	Intra	Intra	6	4.88	06	4.88
	TOTAL CREDITS HOURS		123		123	100%

that they address emerging technologies and hence are non-traditional;
- a prioritisation of the topics since covering everything may not be feasible;
- a list of courses and their appropriate hierarchy in the curriculum;
- an identification of the necessary technologies that would be needed to offer this new

From Table 1, it can also be seen that the general studies is composed of 20.32%, management 12.19%, technical 62.61% and INTRA 4.88%. The total credits hours is 123. It is important to note that the department responsible for the technical concentration areas have specified the courses for each IART core area. Since concentration areas are attached to individual academic departments such department of automation, electrical and mechanical, this allows the individual departments a great deal of flexibility. There are currently 3 different technical concentrations available for the IART student, i.e., automation, electrical and mechanical.

Various subjects are offered in the 8 semesters. As an example, some of the subjects offered in semester 1 is shown in Table 2.

credit hours of coursework (over 8 semesters) divided among four general categories of classes. The four areas are general studies, management, technical and industri training (INTRA) as shown in Table 1, where CH is the credit hours and TCH is the total credit hours.

Table 2: Subjects of 1st Semester

No	Code	Subjects	Credit Hours	No. of Hrs Per Week	LEC	TD	P	No of Hrs Per Sem
1	AUT 3412	Introduction to Automation	2	3	1		2	54
2	FDE 3012	Engg. Drawing	2	3	2			54
3	ELE 3112	Electrical Fundamental	2	3	1		2	54
4	INM 3012	Safety Management	2	2	2			36
5	LAN 1002	Bahasa Malaysia	3	3	3			54
6	LAN 1004	Islamic (or Moral) Studies	3	3	3			54
7	MAT 3112	Mathematics 1	2	3	1	1.5		45
8	MAA 3012	Static and Dynamics	2	2.5	1	1.5		45
			18	22	14	3	4	396

From Table 2, it can seen that the subjects offered in semester 1 are 'Introduction to Automation', Engineering Drawing, Electrical Fundamental, Safety Management, Bahasa Malaysia, Islamic (Moral) Studies', 'Mathematics I' and 'Statics and Dynamics'. As an example, looking at 'Introduction to Automation', it can noted that it is 2 credits hours subject where the number of contact hours between the lecturers and students are 3 hours which is divided as 1 hour for lecture and 2 hours for practical. In one semester, 54 hours is allocated for this subject. A total of 396 'contact hours' takes place between the lecturers and students.

To look at the topical contents of the subjects offered, an example of a typical higher level subject is discussed below. The title of the subject title is known as 'Robotics I'. It is offered as a 3 credit hour core subject. Its main objective is to provide senior undergraduate students with broad knowledge on robots. The class consists of two and a half hours of lecture and tutorial, and a three-hour laboratory session weekly. The lecture and tutorial portion is used to teach fundamentals of robotics while the laboratory portion teaches 'hands on' robot programming techniques. The course is concerned with "classical" Robotics, which covers geometric models of robot manipulators, kinematics, dynamics, control, and path planning. Many excellent textbooks in this area are available. The authors have used and particularly recommends the texts by John J. Craig [6] or P. J. McKerrow [7].

The course contents are listed below:
- Review of Mathematical concepts. This part is a review and an introduction to trigonometric functions, vectors and matrices, and geometric transforms such as translations and rotations of objects in space. The use of Matlab in solving numerical problems is examined through lecture examples and homework assignments.
- Kinematics of robot manipulators. The use of the Denavit-Hartenberg model for robot manipulators is introduced and the study of direct and inverse kinematics is presented.
- Robot path planning and generation. Student are taught to use cubic splines to determine a path for the robot between two end points.

- Dynamics and Control of a Two-link robot. The students are introduced to the concepts of Newton-Euler dynamics only in the case of a two-link robot. Classical joint-position and velocity control is also examined without delving too deeply into this subject.
- Robot Manipulator Programming. A six-axis industrial robot, ABB and Staubli provides students with a test bed for robot manipulator programming. As a laboratory exercise, students program the ABB and Staubli, for a pick and place task or a simple assembly task.

4. Student Data

This section is divided into three sections, i.e., student profile in first section which encompasses the student intake and graduation, whilst the second section covers the student performance and the third section discusses about student employment.

4.1 Student Profile

The student profile which encompasses the student intake and graduation is given in Table 3.

Table 3: Student Profile

Intake	M	F	TOTAL
Batch July 2002	40	18	58
Batch July 2003	19	4	23

Final Semester	M	F	TOTAL
Batch July 2002	31	13	44
Batch July 2003	12	3	15

Graduate	M	F	TOTAL
Batch July 2002	30	11	41
Batch July 2003	12	3	15

Fail to Graduate	M	F	TOTAL
Batch July 2002	1	2	3
Batch July 2003	0	0	0

It can be seen from the Table 3, for the batch July 2002 out of 44 students, 41 manage to get the graduation and 3 fail to graduate. For the batch July 2003 all 15 students graduated.

The age group of batch July 2002 and batch July 2003 is shown in Table 4.

Table 4: Age Group

Age	28	27	26	25	24	23	22	21	Total
Batch July 2002	1	2	3	1	6	8	11	26	58
Batch July 2003			1			4	6	12	23

From Table 4, it can be seen that the age of the students enrolled in first semester varies from 21 to 28 years. The first and second batches had initial enrolment of 58 and 23 students. The higher age group is an evidence of candidates enrolement with some Industrial experience.

4.2   Student Performance

The final semester results for the batch July 2002 and July 2003 is shown in Table 5.

Table 5: Final Semester Results

Category	DL	KB	Total
Batch July 2002	3	38	41
Batch July 2003	4	11	15

DL	>3.5
KB	2.00-3.49
KP1	<2.00 1st time
KP2	<2.00 2nd time
GB	<1.00 or <2.00 3rd time
INC	Incomplete semester

From Table 5, it can be seen that in batch July 2002, 3 students fall under the category of DL and 38 students comes under the category of KB. Total number of students passed is 41.

In batch July 2003, 4 students fall under the category of DL and 11 students come under KB. Total number of students passed is 15.

The final CGPA results for the batch July 2002 and batch July 2003 is shown in the Table 6.

Table 6: Final Semester Results

CGPA	<2.00	2.00-2.50	2.50-3.00	3.0-3.50	>3.5	Total
Batch July 2002		3	26	9	3	41
Batch July 2003			10	5		15

From Table 6, it can be seen that in batch July 2002, 3 students have got final CGPA ranging from 2.00 to 2.50, 26 students have got CGPA ranging from 2.50 to 3.00, 9 students have got CGPA ranging from 3.00 to 3.50 and finally 3 students rated above 3.5 CGPA. Total passed is 41.

In batch July 2003, 10 students have got CGPA ranging from 2.50 to 3.00 and 5 students have got CGPA ranging from 3.00 to 3.50.Total passed is 15.

The trend of GPA for the batch July 2003 is shown in the Table 7.

Table 7: Trend of GPA

	KB	KP1	GB	INC	Total
Sem 2003/2	17		2	3	22
Sem 2004/1	13	3			16
Sem 2004/2	15				15
Sem 2005/1	15				15
Sem 2005/2	15				15
Sem 2006/1	15				15

The trend of GPA is shown in Table 7, it can be seen that for the batch July 2003 [Sem 2003/2] out of total 22 students, 17 got KB, 2 got GB and 3 doesn't complete the semester. In Sem 2004/1, 13 students come under KB and 3 come under KP1. In Sem 2004/2, 15 come under KB.In Sem 2005/1, 15 come under KB.In Sem 2005/2, 15 come under KB. In Sem 2006/1, 15 come under KB.

4.3 Student Employment

The majority of the students enrolled in the IART degree program are currently employed. In the first batch, 60-70% of students were able to find employment within 6 months of their graduation whereas 50-60% of the students were employed from the second batch which graduated in July 2006. They are employed in multinationals and as well as locally owned small and medium enterprises. Amongst the companies that there are attached with are Texas Instruments, Solectron, GlaxoSmith Kline, B Braun, Hicom, to name a few. At the time of writing this paper, it is already four months since the graduation of the second batch. As mentioned above, although 60% of them have found employment

within this period, it is expected that over time, the majority of graduates will find employment.

5.0 Conclusion

This paper has discussed the development of an multidisciplinary course in industrial automation and robotics for undergraduate engineering students. The IART degree offers a student with a completed diploma the opportunity to complete a bachelor's degree in a timely fashion, with little lose in transfer credits. The average CGPA of IART students are competitive, and employers recognise the degree and employ the graduates because of the academic content and the graduates' ability to perform, both academically and professionally.

Acknowledgements

References

1. Martin, F., Oberoi, P., Sargent, R., "The 6.270 RobotBuilder's Guide" , MIT Media Lab, Cambridge, MA, 1992.
2. Turner, C., Ford, K., Dobbs, S., Suri, N., & Hayes, P., "Robots in the Classroom." In J. Stewman (Ed.),Proceedings of the Ninth Florida AI Research Symposium (pp. 497-500). Key West, FL:FLAIRS.
3. Turner, C., Ford, K., Hayes, P., Shamma, D., Manseur, R., "Robots in Education", Proceedings of the US-Japan Graduate Student Forum in Robotics, Osaka, Japan, November 1996.
4. W. Wolf and J. Madsen, "Embedded Systems Education for the Future," Proceeding of the IEEE, January 2000, pp. 23-30.
5. A. McGettrick et al, "Computer Engineering Curriculum in the New Millennium," IEEE Transactions on Education, November 2003, pp. 456-462.
6. John J. Craig, "Introduction to Robotics, Mechanics and Control", 2nd Edition, Addison-Wesley. 1989
7. McKerrow, P. J., "Introduction to Robotics," Addison-Wesley, 1991
8. Sergeyev and Alaraje ., "Promoting Robotics Education: Curriculum and State-of-the-Art Robotics Laboratory Development'. The Technology Interface Journal/Spring 2010
9. Olga I. Vaganova, Zhanna V. Smirnova., "Development of training content for master students in course "mechatronics and robotics" at the University" .Amazonia Investiga. Volume 8, No 22
10. C. Li, L. Fu and L. Wang, "Innovate engineering education by using virtual laboratory platform based industrial robot," 2018 Chinese Control And Decision Conference (CCDC), Shenyang, China, 2018, pp. 3467-3472.

Authors Introduction

Dr M. K. A. Ahamed Khan

He is currently working at UCSI University, Malaysia. He received his Ph.D in Robotics and controls from USA. His area of research is robotics, AI and controls. He has published more than100 papers. He is also an IEEE Senior member. He is also the past chair for IEEE RAS Malaysia chapter.

Dr. Mastaneh Mokayef

She has received her PhD from Wireless Communication Centre Faculty of Electrical Engineering in University Technology Malaysia (UTM). She is a member of Board of Engineers Malaysia (BEM). She has been working in UCSI University, Malaysia, in which she currently serves as an Assistant Professor in the Faculty of Engineering and Built Environment (FETBE). To date, he has been awarded the qualifications of a Chartered Engineer (C.Eng.) from the U.K. Engineering Council.

Ts Amar Ridzuan Bin Abd Hamid

He is a lecturer of Mechanical and Mechatronic programmes from Department of Mechanical Engineering, UCSI University, Malaysia. He has completed his Master Degree from Universiti Putra Malaysia, Postgraduate Diploma from UCSI, and a Bachelor Degree with Hounours in Mechanical Engineering (Automotive) from Universiti Teknikal Malaysia Melaka (UTeM), Malaysia.

Dr. Irraivan Elamvazuthi

He is a senior Member, IEEE. He received the Ph.D. degree from the Department of Automatic Control and Systems Engineering, The University of Sheffield, U.K., in 2002. He is currently an Associate Professor at the Department of Electrical and Electronic Engineering, Universiti Teknologi PETRONAS (UTP), Malaysia. His research interests include control, robotics, mechatronics, power systems, and bio-medical applications. He is also the Chair of the IEEE Robotics and Automation Society (Malaysia Chapter).

Dr Abu Hassan Darusman

He is currently an Associate Professor at the Department of Industrial automation at UNIKL MFI, Malaysia. His research interests include control, robotics, mechatronics.

M. K. A. Ahamed Khan1, Mastaneh Mokayef1, Ridzuan, A1, Irraivan Elamvazuthi, Badli Shah Yusoff, Abu Hassan Darusman

Dr Badli Shah Yusoff

 He is currently an Associate Professor at the Department of Industrial automation at UNIKL MFI, Malaysia. His research interests include control, robotics, mechatronics.

Smart Assistive Trolley for Elderly Care and Independence

Dina Koushek[1], Mastaneh Mokayef[1*], MHD Amen Summakieh[1], M.K.A Ahamed Khan[1], Abdul Qayyum[2], Sivajothi A/L Paramasivam[3]

[1]Faculty of Engineering Technology and Built Environment, UCSI University,
Kuala Lumpur, 56000, Malaysia
[*]Email: mastaneh@ucsiuniversity.edu.my
https://www.ucsiuniversity.edu.my/

[2]National Heart and Lung Institute, Imperial College London, UK

[3]University of Wollongong Malaysia, Shah Alam, Slenagor, 40150, Malaysia

Abstract

As people get older their shopping experience gets harder, as they have to keep pushing a trolley that is made of steel, as well as they keep on adding items through their shopping journey the trolley gets heavier. As a result, this project aims to help these senior citizens by providing a robotic trolley that follows them during their journey (through face detection) without the need of any physical interrreference, in addition to calculating the walking distance providing the property of showing how many meters they have walked. The trolley has been fabricated with the walking distance estimation feature that has been accomplished by the ultrasonic sensor after the camera detects the user and tracks him. This strategy is achieved by using an OpenCV library that is particularly could be used in Python programming language. The results have shown a great improvement in the elderly individuals' lives, as it supports them by giving more comfort and extra liberty to the shopping experience.

Keywords: Smart following trolley, OpenCV, Face detection, Walking distance estimation

1. Introduction

As we live in a time where technology is advancing daily, people occasionally come up with contemporary ideas to update or develop specific technologies as elaborated by Ng [1]. One of these technologies is the shopping trolleys which are available in all supermarkets to make it easier for customers to choose and store products. Customers need to drop the products they want to purchase and then proceed to checkout. Currently, when entering the shopping mall or a grocery store the customer has to choose a shopping cart (trolley), select the items they want to buy, and then put them in the shopping trolley which is very challenging this to do for senior citizens.

To get a robot to move and follow you, a brain of the prototype is in order, a brain will move everything in sequence and give orders to all the components chosen by the engineer (designer), and change this sequence if needed to and all of this is accomplished by the algorithm which is the software used to put everything in to place, the software means the language we people choose to communicate with device in this case an automated trolley we are creating as explained by Li [2].

The brain will keep reading the messages from the sensors the engineer chose to use as ultrasonic sensors which is very good when it comes to objects detection as demonstrated by Sanghavi [3], an infrared sensor is efficient to use for people detection even at night because it follows the temperature mapping technique as analyzed by Nagarajan [4], as well as the proximity sensor can be used for the reasons as the ultrasonic sensors used for however with better efficiency, on account of the proximity's higher efficiency it is used for accidents or collision prevention, which is why it's a perfectly suitable option as elaborated by Lee [5].

In order to get the trolley moving automatically, the presence of a motor is a must, as it will receive a signal from the brain (microcontroller) to move, however to get the motors to move in different directions forward, backward, right, and left motor controllers are needed, many motors can be used to get a robot to move it depends on the speed required, the place the robot will be used in, as DC motors or a servo motor.

Before the microcontroller sends a message to the motors to move in any direction first the microcontroller should locate the person who is supposed to follow and then send the message, and that is accomplished by adding a camera to the trolley that would track the person

Dina Koushek, Mastaneh Mokayef, MHD Amen Summakieh, M. K. A Ahamed Khan, Abdul Qayyum, Sivajothi A/L Paramasivam.

based on his/her features, or even the camera could follow something that the trolley user is holding on to, or even by using a transceiver

2. Methodology and Experimental Setup

In this chapter, the methodology applied for this project is described, as well as the mathematical approaches and correlations related to the smart follower trolley and the components selection shall be mentioned in Fig. 1.
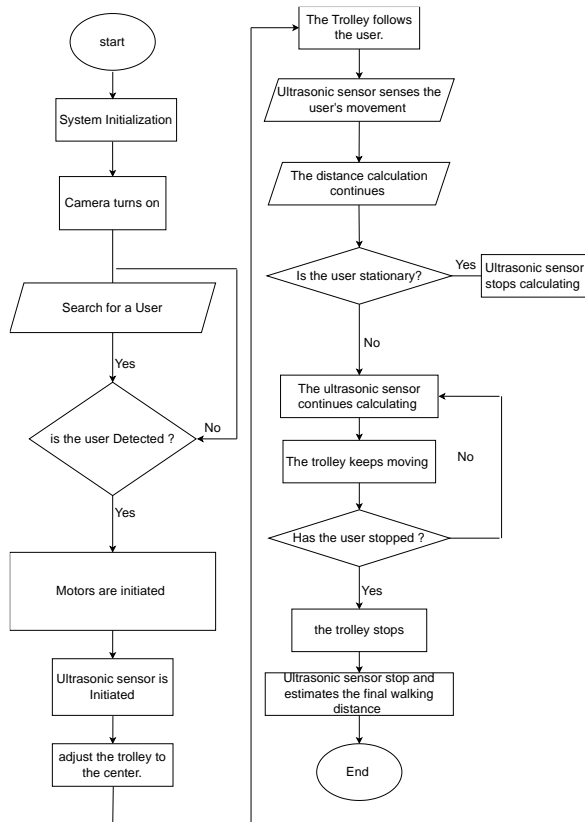
Fig. 1. Project flow chart

In this proposed project the methodology thesis will be implemented into three parts. To begin with, how will the detection and tracking of the user's movements will be implemented, following on from that, walking distance besides the user's directions estimation, finally the whole project's (the smart follower trolley) fabrication specification in terms of connecting all the components chosen due to their quality and cost.

The components used to bring this project's idea to life are a raspberry pi 4 model B as the main system core (brain of the system), a permanent magnet DC motor, an Arduino UNO is used to control these motors while being connected to the Raspberry pi 4 model B through an USB as a communication interface, the sensors chosen are the ultrasonic sensor used for obstacle detection and avoidance, as for the user's detection the infrared sensor is used along with a camera settled on the trolley.

2.1. *Hardware of the trolley*

Raspberry pi 4 Model B

The most recent model of the Raspberry Pi computer line is the Raspberry Pi 4 Model B. Comparing it to the Raspberry Pi 3 Model B+ of the previous generation, it provides revolutionary improvements in processing speed, multimedia performance, memory, and connection while maintaining backward compatibility and a similar level of power consumption. Performance on the desktop is comparable to that of entry-level x86 PC systems for the end-user on the Raspberry Pi 4 Model B as denoted by Raspberry Pi [6].

This fourth-generation raspberry pi model is chosen for in project as the main system core in other words the brain of the robot, due to its high-performance 64-bit quad-core processor, a dual display support at a resolution up to 4K (Kilo) through the micro-HDMI (High Definition Multimedia Interface) ports as modular compliance certification for Bluetooth and dual-band wireless LAN (Local Area Network) enables the board to be integrated into finalized devices with a great reduction in compliance testing, reducing cost and speeding up time to market, as well as the raspberry pi 4 model B contains a higher memory (2/4/8GB RAM option), its connectivity is better than the other raspberry pi computer types, as it has 2.4 GHz and 5.0 GHz, the data transfer is done at a very high rate as it contains 2 USB 3.0 ports 2 USB 2.0 ports. It has 40 GPIO pins.

Arduino UNO

An ATmega328P-based microcontroller board is the Arduino UNO. It contains 6 analogue inputs, a 16 MHz ceramic resonator, a USB port, a power jack, an ICSP(In Circuit Serial Programming) header, and a reset button. It also has 14 digital input/output pins, six of which can be used as PWM outputs. It comes with everything required to support the microcontroller; to get started, just use a USB cable to connect it to a computer, or an AC-to-DC adapter or battery to power it. You can experiment with your UNO without being very concerned about making a mistake as demonstrated by Arduino memory guide [7].

An ATmega328P-based microcontroller board is the Arduino UNO. It contains 6 analogue inputs, a 16 MHz ceramic resonator, a USB port, a power jack, an ICSP header, and a reset button. It also has 14 digital input/output pins, six of which can be used as PWM outputs. It comes with everything required to support the microcontroller; to get started, just use a USB cable to connect it to a computer, or an AC-to-DC adapter or battery to power it. You can experiment with your UNO without being very concerned about making a mistake as demonstrated by Arduino Memory Guide [7].

Using an Arduino UNO to control the PMDC motors is very effective and why is that, due to its simplicity, the Arduino UNO is designed in a way to work efficiently with real-time projects, as well as the Arduino UNO can work easily with motor drivers as an L293D motor driver, L298N motor driver, in addition to it way to work effectively with components that consume low power, as well as changing the speed of a DC motor is fairly simple. The speed of a DC motor can change when PWM is applied to the analogue output pin of an Arduino. Therefore, it is a straightforward task. Unlike the raspberry pi 4 model B, it works efficiently with components that have high power consumption, it will work with the PMDC motor, but not as effectively as the Arduino UNO as explained by DC motor speed [8].

High precision ultrasonic sensor range finder

Three ultrasonic sensors will be used, each one of them will be directly connected to the raspberry pi 4 model B (brain), they will be used for obstacle avoidance, ultrasonic sensor ESP32 is chosen as its better than the ultrasonic sensor ESP8266 in terms of performance as well as it provides BLE.

PMDC motor

A PMDC motor is chosen for this project dues to its high torque and low power consumption, the kind of PMDC motor is chosen based on its torque speed, and this parameter is chosen after choosing the trolley itself and deciding which size is preferred to bring this idea to life.

43A H-Bridge driver

If any motor would work on its own it would work in only one direction (forward), however in order for us to control how the motor we are using, in this case, a PMDC motor needs a driver to make the motor move forward or backward, this driver will be receiving the orders from the main system core (raspberry pi 4) to tell what should be the motor's next move, in addition to what velocity should it move with, as well as the direction it should be headed to upon the logic added to this core.

LM2596 3A Buck module with display

A converter is in order as its very important when it comes to prototyping, as different components are chosen for this project and each work at a very specific level of voltage, so the converter's job is to step down the voltage in order not to ruin the components.

2.2. *Software*

The most important library used for this project is OpenCV python has the option of image processing that is chosen in this project to use in the software part for user's detection. Processing a video means executing operations frame-by-frame to the video. Frames are nothing more than a single discrete instance of the video at a given point in time. Even in one second, there could be several frames. It is possible to treat frames similarly to an image. Therefore, we can execute all actions on frames that we can on images precisely explained by [9].

3. Results and Discussion

In order for the brain to do its job which is to control all other components it needs to be connected to a power source which in this case is the laptop, after the connection is settled the set-up is made by VNC viewer which is a way to access and control the raspberry pi (any kind) desktop remotely from another computer, using a USB C-type cable to activate the brain itself, for this minicomputer to work it needs internet to function based on this the raspberry pi 4 model B has the feature to connect it to the internet without using an ethernet cable. In this project the raspberry pi has a wireless connection with internet.

The connection between the Raspberry Pi and the Arduino UNO is done using a USB B type cable, this is only done for energizing it and it's the way in which the Arduino UNO will take the orders from the raspberry pi, as for the connections between the Arduino UNO and the PMDC motor, first the motor works at a 250 watts, and it needs to be stepped down this is the main reason why a converter is used. The motors are designed in a way that when it works it works only in one direction to make this motor move in all directions a motor driver is used. The connections are made as follows starting with connecting the Arduino UNO to the motor driver ENA pins and IN1, afterward connecting the motor driver to the PMDC motor through positive and negative terminals, and the step-down converter as well in the same way. This converter will be connected to the power supply. And all of the components are connected to the same ground on the used breadboard.

For connecting the ultrasonic sensor with the raspberry pi to achieve the second objective, first the ultrasonic sensor is set at the front of the trolley to be able to detect the person and calculate his walked distance, there are four main wires for this connection, one from the VCC (ultrasonic sensor) to the positive railway, ground wire to the negative railway on breadboard. As for the raspberry pi the GPIO pin 5 volts which is pin number two connected to the positive railway of the breadboard, as well as pin number six is connected to the negative railway where I plugged in the ultrasonic sensor. Then connecting the trig connection of ultrasonic sensor to GPIO 23 which is pin number 16 through a blank rail on the breadboard, and finally connecting ECHO in the ultrasonic sensor with GPIO 24 which is pin number eighteen through blank rail on the breadboard. Fig. 2

Dina Koushek, Mastaneh Mokayef, MHD Amen Summakieh, M. K. A Ahamed Khan, Abdul Qayyum, Sivajothi A/L Paramasivam.

shows the primary fabrication of the proposed smart trolley.

Fig. 2 Automated Trolley

3.1. *Software Results*
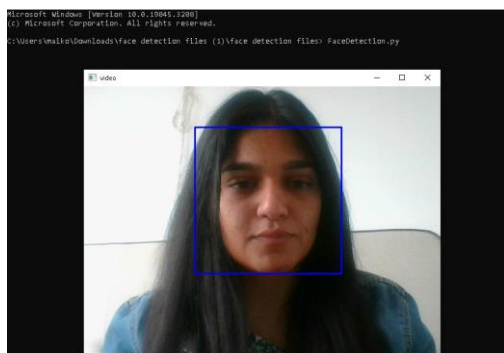
Fig. 3. Face Detection

The performance or real-time face detection is one of the goals of this project first the numpy and OpenCV libraries were added to the code, and then adding the XML file that contains all of the important knowledge for accomplishing the face detection for us to enable the camera to recognize the person in front of it, afterward a video of size 640 cm x 480 cm starts that its only functionality is to keep capturing images, then a while loop is created for the continuous photo capturing. Inside the while loop, the photos (frames) captured will be turned into a grayscale for a better quality formation. The next step for all the faces detected in each frame is a blue box will appear marking all of them that will keep appearing even if the people move around in the camera's region the system will wait for the person using the trolley to press the "ESC" button if they would like to end the shopping journey. If the person presses this button all data will be destroyed, and the system will be ready for the next customer. The result of face detection is illustrated in Fig. 3.

Supported Sustainable Development Goals

The Sustainable Development Goals ("SDGs") are intended to direct international development efforts during the ensuing fifteen years, from 2015 to 2030. They are planned to be endorsed by the United Nations General Assembly (UNGA) in 2015 as part of its development agenda. The Millennium Development Goals ("MDGs"), which were in effect from 2000 to 2015, have been replaced. The sustainable development means to make the world a better place without destroying the possibilities for the next generations keeping three life aspects in mind which are social progress, economic development and climate and environment. The smart follower trolley for elderly care project supports as explained by Pogge [10].

4. Conclusion

The fabrication of a smart assistive trolley specifically designed for elderly care has brought about a remarkable advancement in enhancing the shopping experience for senior citizens. The incorporation of face detection technology and walking distance estimation capabilities in this robotic trolley has resulted in increased comfort, convenience, and a greater sense of independence for the elderly. These improvements ultimately contribute to enhancing the overall quality of life for older individuals, making their shopping experiences more enjoyable and empowering.

References

1. Y. L. Ng, C. S. Lim, K. A. Danapalasingam, M. L. P. Tan, and C. W. Tan, 'Automatic human guided shopping trolley with smart shopping system', *J Teknol*, vol. 73, no. 3, pp. 49–56, 2015, doi: 10.11113/jt.v73.4246.
2. E. Li, L. Bi, and W. Chi, 'Brain-controlled leader-follower robot formation based on model predictive control', in *IEEE/ASME International Conference on Advanced Intelligent Mechatronics, AIM*, Institute of Electrical and Electronics Engineers Inc., Jul. 2020, pp. 290–295. doi: 10.1109/AIM43001.2020.9158815.
3. S. Sanghavi, P. Rathod, P. Shah, and N. Shekokar, 'Design and implementation of a human following smart cart', in *Proceedings of the 3rd International Conference on Smart Systems and Inventive Technology, ICSSIT 2020*, Institute

of Electrical and Electronics Engineers Inc., Aug. 2020, pp. 1142–1149. doi: 10.1109/ICSSIT48917.2020.9214217.

4. S. Nagarajan, P. Banerjee, W. Chen, and B. A. Chin, 'Control of the Welding Process Using Infrared Sensors', 1992.

5. H. K. Lee, S. Il Chang, and E. Yoon, 'Dual-Mode Capacitive Proximity Sensor for Robot Application: Implementation of Tactile and Proximity Sensing Capability on a Single Polymer Platform Using Shared Electrodes', *IEEE Sens J*, vol. 9, no. 12, pp. 1748–1755, 2009, doi: 10.1109/JSEN.2009.2030660.

6. Raspberry Pi Trading Ltd., 'Raspberry Pi 4 Computer Model B', 2019.

7. 'Arduino Memory Guide | Arduino Documentation'. Accessed: Jul. 28, 2023.

8. DC Motor speed control and measurement | Arduino Project Hub'.

9. 'Python - Process images of a video using OpenCV - GeeksforGeeks'.

10. T. Pogge and M. Sengupta, 'The Sustainable Development Goals (SDGs) as Drafted: Nice Idea, Poor Execution', Washington International Law Journal, vol. 24, 2015.

Authors Introduction

Ms. Dina Ashraf Mohamed Talaat Koushek

She is currently pursuing a Bachelor's degree in Electronics Engineering (Communication) with Honors at UCSI University Malaysia.

Dr. Mastaneh Mokayef

She has received her PhD from Wireless Communication Centre Faculty of Electrical Engineering in University Technology Malaysia (UTM) in 2014. She has also obtained her master's degree from the faculty of engineering in 2009 from the University Technology Malaysia. She has been working in UCSI University, Malaysia, since 2015 in which she currently serves as an Assistant Professor in the Faculty of Engineering and Built Environment. Her research interests include: Wireless communications, spectrum sharing method, spectrum management, etc.

Mr. MHD Amen Summakieh

He received the B.Eng. degree (Hons.) in communication and electronics engineering from UCSI University, Malaysia, in 2016, and the M.Eng.Sc. degree from Multimedia University, Malaysia, in 2020. His research interests include heterogeneous LTE-advanced cellular networks, user association, metaheuristic algorithms, and antennas design.

Dr. M. KA. Ahamed Khan

He completed his undergraduate degree in Electronics and Communication Engineering and postgraduate degree in Electronics and Control Engineering in India. Additionally, he obtained an advanced diploma in Power Electronics and Drives from Lucas-Nuelle in Germany in 2002, and a Diploma in Drives and Controls from Woo Sun in Korea in 2014. He pursued a PhD in Robotics, Power Electronics, and Controls in the United States and holds certifications. He is a Senior member of the IEEE in the USA and a member of MIET in the UK.

Dr. Abdul Qayyum

He received his PhD in Electrical and Electronic Engineering from Universiti Teknologi Petronas, Malaysia, in 2018. He developed the totype. Previously, he completed his bachelor's degree in computer engineering and master's degree in electronic engineering from Pakistan. He is a postdoctoral researcher at the University of Burgundy, France, where he studies cardiac MRI images using deep learning approaches. He is currently affiliated with the National Heart and Lung Institute at Imperial College London, UK.

Dr. Sivajothi A/L Paramasivam

He completed MEng in Manufacturing Management at the University of South Australia and then received his PhD in Mechanical Engineering from University of Northumbria, Newcastle, UK. He is a Senior Lecturer, currently Head of Department for school of Engineering in the same university. He is interested in discipline-based and pedagogic research related to teaching and its associated framework, etc.

Parallel Cross Window Attention Transformer and CNN Model for Segmentation of Instrument during Surgery

Abdul Qayyum
Imperial College, London, United Kingdom, Email: a.qayyum@imperial.ac.uk

Steven Niederer
The Alan Turning Institute, London, United Kingdom

M. K. A. Ahamed Khan
Department of Mechanical and Mechatronics Engineering, Faculty of Engineering, Technology and Built Environment, UCSI University, Malaysia.
Email: mohamedkhan@ucsiuniversity.edu.my

Moona Mazher
Centre for Medical Image Computing, Department of Computer Science, University College London, UK
Email: m.mazher@ucl.ac.uk

Imran Razzak
University of New South Wales, Sydney, Australia

Mastaneh Mokayef
UCSI University, Faculty of Engineering, Malaysia

C. S. Hassan
UCSI University, Faculty of Engineering, Malaysia

Ridzuan, A
UCSI University, Faculty of Engineering, Malaysia

Abstract

Precise segmentation of surgical instruments is a fundamental component in the development of computer-aided surgery systems by assisting the surgeons to navigate the patient's body aiming to enhance the surgical precision and patient safety. Though real-time tracking of surgical instruments is critically important in invasive computer-assisted surgeries, it is challenging to achieve a highly sensitive and accurate system in complex surgical environment. Recently, synthetic data for instrument segmentation in surgery (Syn-ISS) challenge using synthetic datasets is organized to develop high performance methods for instrument segmentation. In this work, we present encoder and decoder-based hybrid parallel cross window attention-based transformer during the feature extraction, which consists of the multi-scale channel attention, convolutional layers, and Transformer layers, forming a unified block. Syn-ISS challenge dataset comprised of two tasks. In first task1, they need to develop deep learning-based method for binary instrument segmentation and in second task multiclass instrument segmentation is required. Experiments conducted on Syn-ISS dataset achieved 0.993 F-score for task 1 and 0.993, 0.975, and 0.951 F-score for shaft, wrist, and jaw segmentation respectively for Task 2.

Keywords: Deep Learning, Parallel Cross Window Attention, Transformer, 2D Instrument Surgery segmentation, Dense Net.

1. Introduction

Minimally invasive segmentation using optical imaging systems have gained popularity in modern healthcare due to their advantages, including reduced patient recovery time and lower mortality rates. Optical imaging has enabled the use of robotic platforms such as the da Vinci surgical system by Intuitive Surgery for complex minimally invasive surgeries [1]. Nevertheless, during endoscopic surgical suturing procedures, the presence of surgical instruments can impede surgeons' dexterity due to the confined working space and limited visual field-of-view. These visual obstructions elevate the risk of tissue scars and tears. Therefore, the crucial task is to transparently remove or mask the surgical instruments from the background and subsequently fill the masked region with appropriate background content. Automated segmentation of surgical instruments in MIS is currently a focal point of research due to its significant practical applications [2].

Abdul Qayyum, M. K. A. Ahamed Khan, Moona Mazher, Imran Razzak, Mastaneh Mokayef, C. S. Hassan, Ridzuan, A

The challenges associated with surgical instrument segmentation are diverse and contingent upon factors such as the source of dataset acquisition, the type of surgical procedure, the specific instruments or tools in use, image resolution, dataset scale, tool characteristics, and challenging conditions like occlusions, rapid appearance alterations, specular reflections, smoke, blur, and blood spatter. Segmentation of surgical instruments has been formulated using both instance segmentation [2] and semantic and segmentation [2].

Recently, Syn- ISS (synthetic data for instrument segmentation in surgery) challenge using synthetic datasets is organized to develop high performance methods for instrument segmentation [3]. Recently, there is a different method has been proposed using medical imaging and signals [4], [5], [6], [7], [8] for classification and segmentation. Based on our previous work on segmentation [6], [7], [8], [9]. We presented encoder-decoder based hybrid transformer and CNN model for instrument segmentation with parallel cross window attention-based transformer block in encoder and 2D Dense Net layer CNN blocks in decoder. The main contribution in this work is:

i. Developed Parallel Cross Window Attention Transformer encoder block for 2D segmentation task.
ii. Proposed 2D Dense Net block at decoder side of proposed model using transformer-based encoder features.
iii. Compare performance on synthetic data for instrument segmentation in surgery for binary and multiclass surgery instrument segmentation.

2. Methodology
2.1. Parallel Cross Window Attention Transformer and CNN model for instrument segmentation

Due to the intrinsic locality of convolution, which is incapable of modelling long-range dependencies. In addition, Transformer generates single scale features with only token wise attention, and it ignores the relationship among channels, thus subpar to tackle situations such as segmenting multi scale lesion regions in medical images. Considering these issues, we introduce the hybrid Transformer block during the feature extraction, which consists of the multi-scale channel attention, convolutional layers, and Transformer layers, forming a unified block. Our proposed model is based on encoder and decoder layers, and we have proposed a hybrid transformer and CNN model for instrument segmentation. Our proposed model consisted of Parallel Cross window attention-based transformer block on the encoder side and 2D Dense Net layer CNN blocks on the decoder side. We have used Dense Net 201 based layers on the decoder side. The proposed model is shown in Figure 1 (a) and Figure 1 (b).

2.2. Parallel Cross Window Attention Transformer Block:

An efficient hybrid segmentation framework consisting of integration of convolutional neural network and learnable global attention heads using Efficient Parallel-Cross Attention module. We use the depth-wise separable convolution as an efficient version of convolution implemented by depth wise conv and pointwise conv, where the depth wise convolution gathers the spatial information while the pointwise convolution gathers along the channel dimension. Furthermore, we have concatenated features from multi-window transformer block with depth wise convolutional layer. In encoder side, we have used transformer-based block aided with cross attention window-based mechanism, however, we have used normal 2DCNN based module at decoder-side. The features are concatenated from window area partition and window partition and further these features are concatenated with depth wise convolutional layer features and then pass these features to next layer of the encoder block.

Unlike the vision transformer (ViT) that computes relationship between tokens at each step of self-attention module, swin transformer is based on computation of attention within partition of non-overlapping local windows of lower resolution feature map and original image. In contrast to the original swin transformer that uses patch merging layer to empower it for pixel level tasks, we used rectangular-paralleled-piped windows to accommodate non-square images using Parallel-Cross Attention approach (window area partition and window partition). To extract different feature maps from each convolutional block with parallel window-based transformer in encoder, each block consists of Parallel Cross Window Attention Transformer block, patch merging and depth wise [7] convolutional layer. The proposed block is shown in Figure 1 (b).

2.3 Convolutional Neural Network (CNN) block:

To better encode spatial location information and inject strong inductive bias, we adopt convolutional block to extract local spatial features. Specifically, given an input feature F_(i-1), we adopt the convolutional block to model local spatial features, which are shown as follows:

$$F_c^i = ConvBlocks_i(F_{i-1}), i \in \{1,2,3,4\} \qquad (1)$$

$where\ F_c^i \in R^{\frac{H}{2^{i+1}} \times \frac{W}{2^{i+1}} \times C_i}$ is the local features, which contains 2D spatial location information, making it possible to encode position information. The convolutional blocks in each decoder stage consisted of DensNet-201 based architecture is shown in Figure 1 (a).
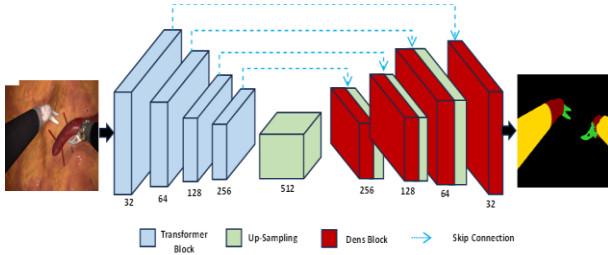
Fig. 1 (a) The proposed model based on hybrid transformer and Dense Net CNN for instrument and instruments parts segmentation
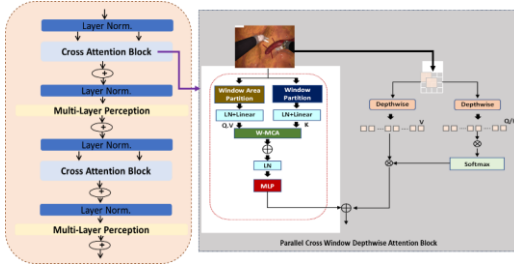
Fig. 1 (b) The proposed Transformer block based on parallel cross window depth wise attention module

In the encoder block, the spatial input size has been reduced with an increasing number of feature maps and on the decoder side, the input image spatial size has been increased using a 2D Conv-Transpose layer. The input features' maps that are obtained from every encoder block are concatenated with every decoder block feature map to reconstruct the semantic information. The spatial size doubled at every encoder block and feature maps are halved at each decoder stage of the proposed model. The feature concatenation has been done at every encoder and decoder block except the last 1x1 convolutional layer. The three-level deep-supervision techniques are applied to get the aggregated loss between ground truth and prediction.

3.0 Experiment
3.1 Dataset

Recently, Syn-ISS challenge [3] is introduce in MICCAI-2023 to develop high performance methods for instrument segmentation. Syn-ISS dataset is synthetic instrument segmentation datasets consist of two main tasks. The task-1 is a binary segmentation to annotate all pixels that contain an instrument and consist of 1200 instances of simulated scene along with computer generated corresponding masks. Task-2 further focuses on segmentation of pixels belonging to different parts of the instrument and consists of 1800 instances.

3.2 Network Setting

We adopted data augmentation of horizontal flips, vertical flips, and random rescales. The network is trained for 200 epochs using the Adam optimizer and the weight decay is 0.0001. We have used binary cross-entropy and dice loss used for training and optimization [10], [11]. [12]. The dataset has different spatial size, hence, we resized each image and label sample to 512x512, whereas we resized each sample to original input 2D image size by bilinear interpolation during inference time. The 25-batch size is used during training. The model is training on A6000 GPU machine with 4 GPUs and all model codes are developed from scratch using Pytorch Library.

4.0 Result

The training dataset has been divided into 80 percent training and 20 percent validation. We have trained and validated our proposed model using 5-fold cross validation and based on the best validation score, the proposed model has been submitted for task 1 binary segmentation and task 2 multiclass segmentation. We have evaluated the performance using IOU, F-Score, Recall, Precision, and HD. The best score produced by our proposed model based on validation dataset is shown in Table.1 for Task 1 and Task 2.

Table 1. The performance analysis of proposed solution for Task1 and Task2.

Algorithms	Tasks	classes	IOU	F-Score	Recall	Precision	HD
Proposed model 1	Task 2	shaft	0.9868	0.9930	0.9923	0.9920	14.23
		wrist	0.9497	0.9753	0.9753	0.9695	10.67
		jaw	0.9086	0.9506	0.9523	0.9357	12.88
Proposed model 1	Task 1	instrument	0.9868	0.9933	0.9934	0.9909	0

The visualization of validation input image for binary and multiclass segmentation is shown in Figure 2. Our proposed produced similar prediction masks as compared to ground-truth segmentation masks.

input   input

Abdul Qayyum, M. K. A. Ahamed Khan, Moona Mazher, Imran Razzak, Mastaneh Mokayef, C. S. Hassan, Ridzuan, A
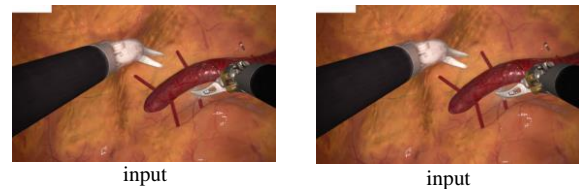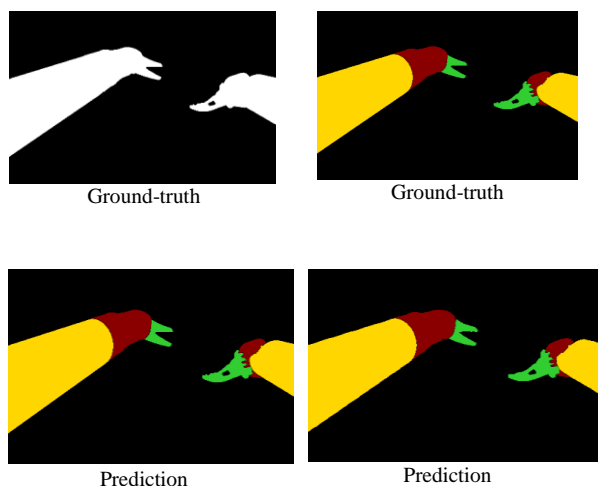
Fig. 2 The input, ground-truth, and predicted segmentation masks. shaft, wrist, and jaws are shown using a yellow, red, and green segmentation mask.

5.0 Conclusion

Ideally, surgical instrument segmentation is used in real time, to identify tools being used as a surgery is being performed. The integration of surgical instrument segmentation into computer-aided surgery systems offers numerous benefits, including real-time guidance, instrument tracking, and improved surgical outcomes. In this work, we presented an encoder and decoder-based hybrid parallel cross window attention-based transformer which consists of the multi-scale channel attention, convolutional layers, and Transformer layers. Experiments conducted on Syn-ISS challenge dataset achieved 0.993 F-score for task-1 and 0.993, 0.975, and 0.951 F-score for shaft, wrist, and jaw segmentation respectively for Task 2.

6. References

1. Cristina Gonz´alez, Laura Bravo-S´anchez, and Pablo Arbelaez. Isinet: an instancebased approach for surgical instrument segmentation. In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 595–605. Springer, 2020.

2. Mobarakol Islam, VS Vibashan, and Hongliang Ren. Ap-mtl: Attention pruned multi-task learning model for real-time instrument detection and segmentation in robot-assisted surgery. In 2020 IEEE international conference on robotics and automation (ICRA), pages 8433–8439. IEEE, 2020.

3. https://www.synapse.org/#!Synapse:syn50908388/wiki/620516

4. Payette, Kelly, Hongwei Bran Li, Priscille de Dumast, Roxane Licandro, Hui Ji, Md Mahfuzur Rahman Siddiquee, Daguang Xu et al. "Fetal brain tissue annotation and segmentation challenge results." Medical Image Analysis 88 (2023): 102833.

5. De Vente, Coen, Koenraad A. Vermeer, Nicolas Jaccard, He Wang, Hongyi Sun, Firas Khader, Daniel Truhn et al. "AIROGS: Artificial Intelligence for robust glaucoma screening challenge." IEEE Transactions on Medical Imaging (2023).

6. Chen, Zhihao, Alain Lalande, Michel Salomon, Thomas Decourselle, Thibaut Pommier, Abdul Qayyum, Jixi Shi, Gilles Perrot, and Raphaël Couturier. "Automatic deep learning-based myocardial infarction segmentation from delayed enhancement MRI." Computerized Medical Imaging and Graphics 95 (2022): 102014.

7. Qayyum, Abdul, Mona Mazhar, Imran Razzak, and Mohamed Reda Bouadjenek. "Multilevel depth-wise context attention network with atrous mechanism for segmentation of COVID19 affected regions." Neural Computing and Applications (2021): 1-13.

8. Ahmad, Iftikhar, Abdul Qayyum, Brij B. Gupta, Madini O. Alassafi, and Rayed A. AlGhamdi. "Ensemble of 2D residual neural networks integrated with atrous spatial pyramid pooling module for myocardium segmentation of left ventricle cardiac MRI." Mathematics 10, no. 4 (2022): 627.

9. Ahmad, R. F., Malik, A. S., Kamel, N., Amin, H., Zafar, R., Qayyum, A., & Reza, F. (2014, November). Discriminating the different human brain states with EEG signals using Fractal dimension: A nonlinear approach. In 2014 IEEE International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA) (pp. 1-5). IEEE.

10. Qayyum, Abdul, Moona Mazher, Tariq Khan, and Imran Razzak. "Semi-supervised 3D-InceptionNet for segmentation and survival prediction of head and neck primary cancers." Engineering Applications of Artificial Intelligence 117 (2023): 105590.

11. Eisenmann, Matthias, Annika Reinke, Vivienn Weru, Minu Dietlinde Tizabi, Fabian Isensee, Tim J. Adler, Patrick Godau et al. "Biomedical image analysis competitions: The state of current participation practice." arXiv preprint arXiv:2212.08568 (2022).

12. Qayyum, Abdul, Aamir Malik, Naufal M Saad, and Moona Mazher. "Designing deep CNN models based on sparse coding for aerial imagery: a deep-features reduction approach." European Journal of Remote Sensing 52, no. 1 (2019): 221-239.

Authors Introduction

Dr. Abdul Qayyum

He is currently working at National Heart and Lung Institute Imperial College London, UK. Previously, he was joined as lecturer at University of Bourgogne Franche-Comté France. He received his Ph.D in electrical & electronics engineering with specialization in deep learning and image processing in 2017 from Universiti Teknologi Petronas Malaysia. His area of interest is machine learning, deep learning and quantum machine learning for signal processing and bioemdical imaging.

Prof. Steven Niederer

He completed his undergraduate degree in Engineering Science at the University of Auckland in 2003 and his DPhil at the University of Oxford in 2008. In 2023, he moved to Imperial College London as the Chair in Biomedical Engineering at the National Heart and Lung Institute. His current work is focused on reducing barriers to adopting digital twin technology, developing virtual patient cohorts for in-silico trials, mapping organ scale function through to cellular and molecular physiology, and using modelling and simulation to personalize and guide therapies.

Dr M. K. A. Ahamed Khan

He is currently working at UCSI University, Malaysia. He received his Ph.D in Robotics and controls from USA. His area of research is robotics, AI and controls. He has published more than100 papers. He is also an IEEE Senior member. He is also the past chair for IEEE RAS Malaysia chapter

Dr. Moona Mazher

She is a senior postdoc research fellow at Department of Computer Science, University College London. She received her Ph.D. from the University of Rovira i Virgili, Spain, 212 a specialization in Neuroscience from Universiti Teknologi PETRONAS, Malaysia in 2017.. Her areas of interest are machine learning, deep learning, medical imaging, signal processing, computer vision, and explainable AI.

Dr. Imran Razzak

He is a Senior Lecturer in Human-Centered Machine Learning in the School of Computer Science and Engineering at University of New South Wales, Sydney, Australia. Previously, he was as a Senior Lecturer in Computer Science at School of IT, Deakin University, Victoria. His area of research focuses on connecting language and vision for better interpretation of multidimensional data and spans over three broad areas: Machine Learning, Computer Vision, and Natural Language Processing.

Dr. Mastaneh Mokayef

He is currently working at National Heart and Lung Institute Imperial College London, UK. Previously, he was joined as lecturer at University of Bourgogne Franche-Comté France. He received his Ph.D in electrical & electronics engineering with specialization in deep learning and image processing in 2017 from Universiti Teknologi Petronas Malaysia. His area of interest is machine learning, deep learning and quantum machine learning for signal processing and bioemdical imaging.

Dr Cik Suhana Hassan

She currently works at UCSI University. She received her PhD from UTP Malaysia. She has been doing research in Mechanical Engineering and Materials Engineering. She is passionate about turning environmental waste into value-added products as part of her quest to live a more environmentally friendly life.

Ts Amar Ridzuan Bin Abd Hamid

He is a lecturer of Mechanical and Mechatronic programmes from Department of Mechanical Engineering, UCSI University, Malaysia. He has completed his Master Degree from Universiti Putra Malaysia, Postgraduate Diploma from UCSI, and a Bachelor Degree with Hounours in Mechanical Engineering (Automotive) from Universiti Teknikal Malaysia Melaka (UTeM), Malaysia.

Magnetic Resonance Spectroscopy (MRS) Reconstruction Using Style Transfer Deep Depth wise Framework

Abdul Qayyum
National Heart & Lung Institute, Imperial College, London, United Kingdom,
Email: a.qayyum@imperial.ac.uk

Steven Niederer
National Heart & Lung Institute, Imperial College, London, United Kingdom
Alan Turning Institute, London, United Kingdom

M. K. A. Ahamed Khan
UCSI University, Faculty of Engineering, Malaysia

Moona Mazher
Centre for Medical Image Computing, Department of Computer Science, University College London, UK

Imran Razzak
University of New South Wales, Sydney, Australia

Mastaneh Mokayef
UCSI University, Faculty of Engineering, Malaysia

C. S. Hassan
UCSI University, Faculty of Engineering, Malaysia

M. Ridzuan, A
UCSI University, Faculty of Engineering, Malaysia

Abstract

The human brain is a complex and heterogeneous organ composed of distinct compartments such as cerebral cortex, the cerebellum, the brainstem, and the subcortical regions. To analyze the chemical composition of tissues in brain, in vivo magnetic resonance spectroscopy allows non-invasive measurements of neurochemicals in either single voxel or multiple voxels. The reconstruction spectra using 1/3rd of original data than current Edited-MRS scans will not only result in four times faster edited-MRS scans but also extensively reduction in radiations. In this work, we present a deep depth-wise channel attention module (DCAM) based fine-tuned network for magnetic resonance spectroscopy image reconstruction. Besides, we have used channel-wise convolutions and average pooling without dimensionality reduction. We have trained the initial network from scratch on track-1 simulated dataset, however due to the limited dataset, we finetune the network on track-2 and track-3. Experiments are conducted on Edited-MRS-Rec-Challenge dataset1 that showed significantly better performance.

Keywords: Deep Learning, Magnetic Resonance Spectroscopy, Style Transfer, depth-wise channel attention, Reconstruction, Huber loss.

1. Introduction

The brain is a highly complex and heterogeneous organ that consists of many different types of cells, including neurons, glial cells, and endothelial cells. These cells interact with each other in complex ways to perform a wide range of functions, including perception, thought, movement, and emotion. Due to its complexity and heterogeneity, studying the brain is a challenging task that requires a wide range of techniques and approaches. These include anatomical imaging techniques, such as magnetic resonance imaging (MRI), magnetic resonance spectroscopy (MRS) and computed tomography (CT), functional MRI (fMRI) and positron emission tomography (PET). The biochemical processes that occur in the brain are highly complex and dynamic, and involve the activity of numerous neurotransmitters, enzymes, receptors, and other signaling molecules. Magnetic Resonance Spectroscopy (MRS) is a non-invasive imaging technique identical with nuclear magnetic resonance (NMR) spectroscopy. It uses magnetic fields and radio waves to analyze the chemical composition of tissues in the human body such as quantification of Gamma Aminobutyric Acid (GABA), which is overlapped by creatine and glutamate.

MRS provides information about the concentrations of various metabolites within tissues, which can be used to diagnose and monitor diseases. Besides, it can also be

Abdul Qayyum, Steven Niederer, M. K. A. Ahamed Khan, Moona Mazher, Imran Razzak, Mastaneh Mokayef, C. S. Hassan, M. Ridzuan, A

used to detect tissue changes in stroke and epilepsy. MRS data is acquired as a series of spectra, each representing the concentration of a different metabolite [2], [3]. The spectra are typically acquired in a 3D volume, with each voxel (3D pixel) in the volume containing a spectrum and can provide estimates of GABA levels in the brain, however, it has limitations and may not always be precise or accurate. Figure 1 shows the different chemical peaks of a suspected brain tumor. Besides, its high-quality data is needed to efficiently estimate GABA which requires long scan times. Hence, Edited-MRS reconstruction challenge is organized to investigate the machine learning models for spectra reconstruct using 1/3rd of data than current Edited-MRS scans, which will not only result in to four times faster edited-MRS scans but also extensively reduction in radiations. In this work, we present a deep depth-wise channel attention module (DCAM) for magnetic resonance spectroscopy image reconstruction. Besides, we have used channel-wise convolutions and average pooling without dimensionality reduction.
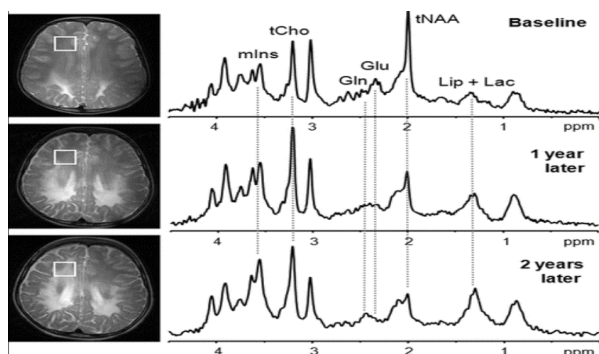
Fig. 1 MRS graph shows the different chemical peaks of a suspected brain tumor [1]

Recently deep learning models have been used in medical image analysis for segmentation, classification, and object recognition [4], [5], [6], [7]. The 2D and 3D segmentation models have been proposed for segmentation of medical imaging [8], [9], [10].

2. Methodology
2.1. Proposed Depth-Wise Edited-Mrs Reconstruction

In this work, we presented UNet based model for MSR reconstruction tasks. Figure 2 and Figure 3 show the proposed reconstruction framework that consists of convolutional blocks, Max-Pooling blocks, Up-Sampling blocks, 1x1 Conv, depth wise channel attention module (DCAM) blocks (see figure 4. The DCAM block is introduced in each encoder block after the convolutional layer block. The convolutional block consists of convolutional layers with Batch-Normalization and ReLU activation function to extract the different feature maps from each block in the encoder side. The 2D max-pooling layer has been used to reduce the input spectral size. The 1x1 convolutional layer with SoftMax function

has been used at the end of the proposed model. Spectral input size is reduced with increasing the number of layers in the encoder block and the spectral input size is recovered via a 2D up sampling layer using a bilinear up sampling method in the decoder side. The DCAM block handles the input feature maps extracted from each convolutional block in the encoder side and further passes these feature maps into the decoder block. The output of each encoder block is fed into the proposed DCAM. Later, the outputs of DCAMs are concatenating them with the corresponding decoder blocks. The output of the last encoder block is passed to the DCAM, and the resulting feature maps are directly concatenated with the corresponding decoder block. Each max-pooling layer receives the output from DCAM. In addition to DCAM, we have implemented different other modules such as channel attention module (CAM), spatial attention module (SAM), channel spatial attention Module (CSAM). As we have larger dataset available for task 1, we trained the proposed model from scratch on track 1. However, there are only few samples available for track 2 and track 3, thus we have developed the style-transfer based framework by finetuning the pre-trained model and freezing the parameters of the encoder and decoder layers except the bottom and last two layers where the weights or parameters are finetuned using vivo dataset for track-2 and track-3. Figure 3 shows the block diagram of finetune network for track-2 and track-3.
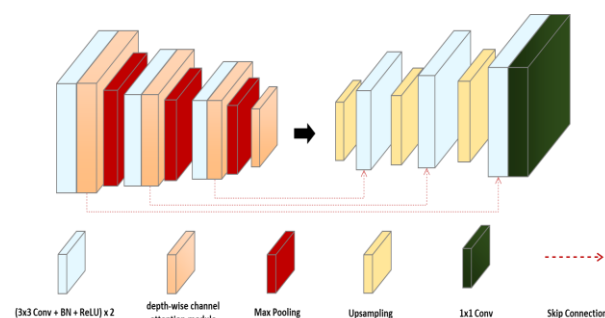
Fig. 2 The proposed model used for track1 simulated dataset
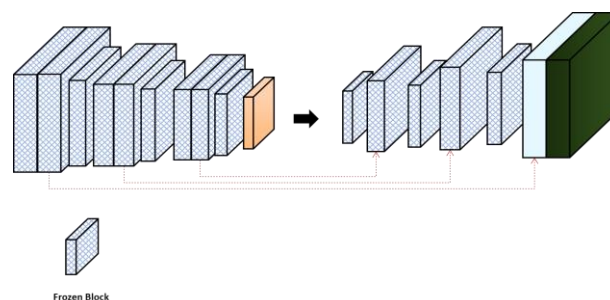
Fig. 3 The proposed model used for track2 and track3 vivo dataset
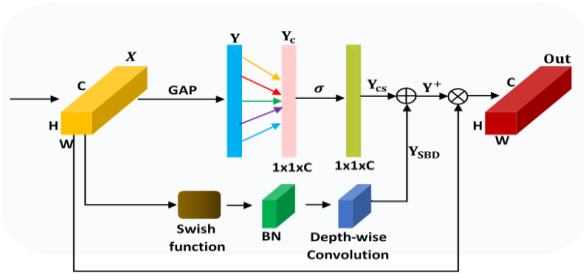
Fig. 4 Diagram of the proposed depth-wise channel attention module (DCAM)

Figure 4 is proposed depth-wise channel attention module (DCAM) used in encoder side of the proposed model. In this experiment, we have used two loss functions: Mean Squared Error (MSE) and Dynamic Huber loss (LHuber). MSE is sensitive towards outliers whereas L Huber is less sensitive to outliers than MSE loss. For N predictions, MSE loss function is defined by;

$$L_{MSE} = \frac{1}{N}\sum_i^N (X_i - \breve{X}_i)^2 \qquad (1)$$

N and $X = (x_1, x_2, \ldots, x_N)$ are the ground truth MRI spectrum and $\breve{X} = (\widetilde{x_1}, \widetilde{x_2}, \ldots, \widetilde{x_N})$ are their corresponding estimated MRI spectrum. The huber loss function is defined;

$$L_{Huber} = \frac{1}{N}\sum_i^N z_i \qquad (2)$$

Where N is the batch size and z_i is defined by;

$$z_i = \begin{cases} 0.5(x_i - \breve{x}_i)^2, & if \; |x_i - \breve{x}_i| \le \beta \\ \beta|x_i - \breve{x}_i| - 0.5\beta^2 & otherwise \end{cases} \qquad (3)$$

Where β is controlling hyperparameter. In our experiment, β decreases from 15 to 1 during the training for each epoch.

2.2. *Simulation Results*

In this section, we briefly first describe the dataset, followed by evaluation parameters, experimental setting, and evaluation. We have used preprocessing step provided by challenge organizer to convert dataset into Frequency domain spectra and added some noises (random Gaussian amplitude, random Gaussian frequency and random Gaussian phase noise, combined).

2.3 Dataset and Evaluation Matrix

The test set may differ in traditional features, such as by vendor and echo time and it will be lower quality data i.e., lower main SNR or larger frequency and phase shifts.

The dataset consists of 5500 spectra (5000 ground-truth spectra and 500 corrupted spectra). The evaluation of the proposed model is performed on an evaluation set which has similar features to that of train set. To evaluate the performance, we have used mean squared error (MSE), signal to noise ratio (SNR), linewidth (FWHM), peak shapes GABA, GLX (PS). Evaluation of model is based on task (simulated, in vivo homogeneous, in vivo heterogeneous) and each metric is ranked individually and well as combined (MSE, SNR, FWHM, PS with ratio of 40%,20%,2 0%, 20%, 20% respectively).

2.4 Parameter Setting

In this work, we β decreases from 15 to 1 during the training for each epoch. We have trained proposed models using Adam optimizer and dynamics Huber loss function using 0.0001 learning rate and 200 epochs. The dataset is normalized using clip [0,1] window. All models use PyTorch and trained on Tesla V100 GPU machine [11], [12]. We have converted all dataset time series into frequency spectrum. The challenge organizers provided simulated and in vivo data training sets representing GABA-edited MEGA-PRESS scans composed of two sub spectra (ON and OFF).

3. Results

We have used scripts provided by organizers to add varying noise, frequency, and phase shifts, that were used for data augmentation purposes. We have developed different variations of proposed framework (depth-wise channel attention module (DCAM), channel attention module (CAM), spatial attention module (SAM), channel spatial attention Module (CSAM)) and compared the performance. The frequency spectrum has been used as input to train our proposed model for task 1, 2 and 3. We have trained proposed model parameters from scratch using 80 percent simulated data in frequency domain and validated using 20 percent of the frequency spectra of simulated dataset. In task two, vivo data (i.e., single-vendor data) in frequency domain has been used to fine-tune the trained model that was trained using simulated dataset. Similarly, for task3, the heterogeneous in vivo data with different frequency spectra has been used to train the proposed model and later fine-tuned the proposed model using different number of epochs. We have used dynamic Huber and Dynamic Tukey loss function for training and optimizing the parameters of our proposed model. The combined loss (Huber and Tukey) loss function provided better performance on validation dataset. Figure 5 shows the comparative analysis. Table 1 describes the evaluation results of proposed edited magnetic resonance spectroscopy reconstruction framework. We can notice that our proposed framework achieved significantly better mean squared error (8.89e-03), signal to noise ratio (4.61e+01), linewidth (9.17e-02) and peak shapes score (9.44e-01) respectively in comparison to baseline method and other methods.

Abdul Qayyum, Steven Niederer, M. K. A. Ahamed Khan, Moona Mazher, Imran Razzak, Mastaneh Mokayef, C. S. Hassan, M. Ridzuan, A

Table 1 Comparative evaluation of proposed framework (depth-wise channel attention module (dcam), channel attention module(cam), spatial attention module(sam), channel spatial attention module (csam)) with baseline results on task 1, task 2 and task 3

	MSE	SNR	Linewidth	Shape Score
Comparative evaluation for track 1				
Baseline	4.37e-02	1.43e+01	2.70e-01	6.89e-01
Proposed DCAM	5.79e-03	3.66e+01	5.50e-02	8.52e-01
Proposed CAM	7.33e-03	4.11e+01	6.44e-02	9.33e-01
Proposed SAM	7.88e-03	5.26e+01	7.37e-02	9.683e-01
Proposed CSAM	6.33e-03	4.01e+01	6.11e-02	8.10e-01
Comparative evaluation for track 2				
Baseline	4.09e-02	1.46e+01	2.56e-01	7.09e-01
Proposed DCAM	5.79e-03	3.66e+01	5.50e-02	8.52e-01
Proposed CAM	7.33e-03	4.11e+01	6.44e-02	9.33e-01
Proposed SAM	7.88e-03	5.26e+01	7.34e-02	9.68e-01
Proposed CSAM	6.33e-03	4.26e+01	6.11e-02	8.12e-01
Comparative evaluation for Track3 with 2048 spectra				
Baseline	9.30e-03	3.77e+05	1.25e-01	9.21e-01
Proposed DCAM	8.26e-04	5.11e+02	8.18e-02	9.92e-01
Proposed CAM	7.12e-04	7.22e+02	5.27e-02	5.12e-01
Proposed SAM	5.32e-04	6.59e+02	4.22e-02	7.12e-01
Proposed CSAM	7.88e-04	6.44e+02	6.22e-02	6.33e-01
Comparative evaluation for Track3 with 4896 spectra				
Proposed DCAM	6.92e-04	5.95e+05	7.68e-02	9.94e-01
Proposed CAM	7.09e-04	2.29e+02	7.63e-02	9.97e-01
Proposed SAM	6.10e-04	3.11e+02	6.41e-02	9.89e-01
Proposed CSAM	6.88e-04	3.89e+02	6.11e-02	8.66e-01
Proposed DCAM	6.33e-04	3.22e+02	6.83e-02	8.98e-01
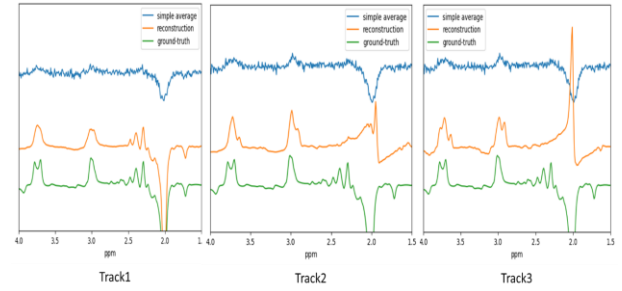
Fig. 5 Comparative analysis: ground truth vs proposes depth-wise channel attention module (DCAM) based reconstruction

4. Conclusion

In this paper, we presented deep depth-wise channel attention module (DCAM) based fine-tuned UNet architecture for magnetic resonance spectroscopy image reconstruction. Due to the availability of large dataset for task 1, we have trained network from scratch on track-1 simulated dataset followed by finetuning the network on track-2 and track-3. Experiments are conducted on Edited-MRS-Rec-Challenge dataset that showed significantly better performance for all parameters i.e. mean squared error (8.89e-03), signal to noise ratio (4.61e+01), linewidth (9.17e-02) and peak shapes score (9.44e-01) respectively in comparison to baseline method and other methods.

5. References

1. G. O¨ z, J. R. Alger, P. B. Barker, R. Bartha, A. Bizzi, C. Boesch, P. J. Bolan, K. M. Brindle, C. Cudalbu, A. Dinc， er et al., "Clinical proton mr spectroscopy in central nervous system disorders," Radiology, vol. 270, no. 3, pp. 658–679, 2014.

2. Z. Su, Z. Liu, M. Wang, S. Li, L. Lin, Z. Yuan, S. Pang, Q. Feng, T. Chen, and H. Lu, "Three-dimensional reconstruction of kambin's triangle based on automated magnetic resonance image segmentation," Journal of Orthopaedic ResearchR, vol. 40, no. 12, pp. 2914–2923, 2022.

3. B. Fu, Y. Dong, S. Fu, Y. Mao, and D. N. Thanh, "Learning domain transfer for unsupervised magnetic resonance imaging restoration and edge enhancement," International Journal of Imaging Systems and Technology, vol. 32, no. 1, pp. 144–154, 2022.

4. Payette, Kelly, Hongwei Bran Li, Priscille de Dumast, Roxane Licandro, Hui Ji, Md Mahfuzur Rahman Siddiquee, Daguang Xu et al. "Fetal brain tissue annotation and segmentation challenge results." Medical Image Analysis 88 (2023): 102833.

5. Qayyum, Abdul, Mona Mazhar, Imran Razzak, and Mohamed Reda Bouadjenek. "Multilevel depth-wise context attention network with atrous mechanism for segmentation of COVID19 affected regions." Neural Computing and Applications (2021): 1-13.

6. De Vente, Coen, Koenraad A. Vermeer, Nicolas Jaccard, He Wang, Hongyi Sun, Firas Khader, Daniel Truhn et al. "AIROGS: Artificial Intelligence for robust glaucoma screening challenge." IEEE Transactions on Medical Imaging (2023).

7. Qayyum, Abdul, Moona Mazher, Tariq Khan, and Imran Razzak. "Semi-supervised 3D-InceptionNet for segmentation and survival prediction of head and neck primary cancers." Engineering Applications of Artificial Intelligence 117 (2023): 105590.

8. Chen, Zhihao, Alain Lalande, Michel Salomon, Thomas Decourselle, Thibaut Pommier, Abdul Qayyum, Jixi Shi, Gilles Perrot, and Raphaël Couturier. "Automatic deep learning-based myocardial infarction segmentation from delayed enhancement MRI." Computerized Medical Imaging and Graphics 95 (2022): 102014.

9. Ahmad, Iftikhar, Abdul Qayyum, Brij B. Gupta, Madini O. Alassafi, and Rayed A. AlGhamdi. "Ensemble of 2D residual neural networks integrated with atrous spatial pyramid pooling module for myocardium segmentation of left ventricle cardiac MRI." Mathematics 10, no. 4 (2022): 627.

10. Qayyum, Abdul, Chun Kit Ang, S. Sridevi, MKA Ahamed Khan, Lim Wei Hong, Moona Mazher, and Tran Duc Chung. "Hybrid 3D-ResNet deep learning model for automatic segmentation of thoracic organs at risk in CT images." In 2020 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), pp. 1-5. IEEE, 2020.

11. Eisenmann, Matthias, Annika Reinke, Vivienn Weru, Minu Dietlinde Tizabi, Fabian Isensee, Tim J. Adler, Patrick Godau et al. "Biomedical image analysis competitions: The state of current participation practice." arXiv preprint arXiv:2212.08568 (2022).

12. Qayyum, Abdul, Iftikhar Ahmad, Mohsin Iftikhar, and Moona Mazher. "Object Detection and Fuzzy-Based Classification Using UAV Data." Intelligent Automation & Soft Computing 26, no. 4 (2020).

Authors Introduction

Dr. Abdul Qayyum

He is currently working at National Heart and Lung Institute Imperial College London, UK. Previously, he was joined as lecturer at University of Bourgogne Franche-Comté France. He received his Ph.D in electrical & electronics engineering with specialization in deep learning and image processing in 2017 from Universiti Teknologi Petronas Malaysia. His area of interest is machine learning, deep learning and quantum machine learning for signal processing and bioemdical imaging.

Prof. Steven Niederer

He completed his undergraduate degree in Engineering Science at the University of Auckland in 2003 and his DPhil at the University of Oxford in 2008. In 2023, he moved to Imperial College London as the Chair in Biomedical Engineering at the National Heart and Lung Institute. His current work is focused on reducing barriers to adopting digital twin technology, developing virtual patient cohorts for in-silico trials, mapping organ scale function through to cellular and molecular physiology, and using modelling and simulation to personalize and guide therapies.

Dr M. K. A. Ahamed Khan

He is currently working at UCSI University, Malaysia. He received his Ph.D in Robotics and controls from USA. His area of research is robotics, AI and controls. He has published more than100 papers. He is also an IEEE Senior member. He is also the past chair for IEEE RAS Malaysia chapter

Dr. Moona Mazher

She is a senior postdoc research fellow at Department of Computer Science, University College London. She received her Ph.D. from the University of Rovira i Virgili, Spain, 212 a specialization in Neuroscience from Universiti Teknologi PETRONAS, Malaysia in 2017.. Her areas of interest are machine learning, deep learning, medical imaging, signal processing, computer vision, and explainable AI.

Dr. Imran Razzak

He is a Senior Lecturer in Human-Centered Machine Learning in the School of Computer Science and Engineering at University of New South Wales, Sydney, Australia. Previously, he was as a Senior Lecturer in Computer Science at School of IT, Deakin University, Victoria. His area of research focuses on connecting language and vision for better interpretation of multidimensional data and spans over three broad areas: Machine Learning, Computer Vision, and Natural Language Processing.

Dr . Mastaneh Mokayef

He is currently working at National Heart and Lung Institute Imperial College London, UK. Previously, he was joined as lecturer at University of Bourgogne Franche-Comté France. He received his Ph.D in electrical & electronics engineering with specialization in deep learning and image processing in 2017 from Universiti Teknologi Petronas Malaysia. His area of interest is machine learning, deep learning and quantum machine learning for signal processing and bioemdical imaging.

Dr Cik Suhana Hassan

She currently works at the Faculty of Engineering, Technology and Built Environment, UCSI University. She received her PhD from UTP Malaysia. Suhana does research in Mechanical Engineering and Materials Engineering. Dr. Suhana is passionate about turning environmental waste into value-added products as part of her quest to live a more environmentally friendly life.

Ts Amar Ridzuan Bin Abd Hamid

He is a lecturer of Mechanical and Mechatronic programmes from Department of Mechanical Engineering, UCSI University, Malaysia. He has completed his Master Degree from Universiti Putra Malaysia, Postgraduate Diploma from UCSI, and a Bachelor Degree with Hounours in Mechanical Engineering (Automotive) from Universiti Teknikal Malaysia Melaka (UTeM), Malaysia.

Federated Learning on Brain Disease Research: Segmentation of Cerebral Small Vessel Diseases (CSVD) using Multi-scale Hybrid Spatial Deep Learning Approach

Moona Mazher
Centre for Medical Image Computing, Department of Computer Science, University College London, UK
Email: m.mazher@ucl.ac.uk

Abdul Qayyum
National Heart & Lung Institute, Imperial College, London, United Kingdom

M. K. A. Ahamed Khan
UCSI University, Faculty of Engineering, Cheras, Malaysia

Steven Niederer
National Heart & Lung Institute, Imperial College, London, United Kingdom
Alan Turning Institute, London, United Kingdom

Mastaneh Mokayef
UCSI University, Faculty of Engineering, Cheras, Malaysia

Ridzuan, A.
UCSI University, Faculty of Engineering, Cheras, Malaysia

C. S. Hassan
UCSI University, Faculty of Engineering, Cheras, Malaysia

Abstract

Federated learning is an emerging approach that enables large-scale decentralized learning without the need to share data among different data owners. This approach is particularly valuable in addressing data privacy concerns in medical image analysis. However, existing methods often impose a strict requirement for label consistency across clients, which significantly limits its applicability. Various clinical sites may only provide annotations for specific organs of interest, and there may be limited or no overlap in the labeled data among different sites. The human brain receives nutrients and oxygen through blood vessels in the brain. The pathology of small vessels, i.e. mesoscopic scale, is a vulnerable component of the cerebral blood supply and can result in major complications such as Cerebral Small Vessel Diseases (CSVD). In this paper, we propose a hybrid architecture for medical image segmentation to produce efficient representations from global and local features and adaptively aggregate them, aiming to fully exploit their strengths to obtain better segmentation performance in federated learning. Furthermore, we propose a multi-scale feature extraction module embedded at the bottom of the proposed model, which can efficiently extract hidden multi-scale contextual information and aggregate multi-scale features. Experiments on segmentation over three-dimensional rotational angiography of internal Carotid Artery with aneurysm (SHINY-ICARUS) challenge dataset show the effectiveness of the proposed multiscale framework.

Keywords: Federated Learning, Deep Learning, Medical Image Analysis, 3D volume Segmentation, Cerebral Small Vessel Diseases, Brain angiography.

1. Introduction

Distributed big data and digital healthcare technologies hold immense potential for advancing medical services. However, challenges arise when attempting to develop predictive models from diverse and intricate e-health datasets. Federated learning (FL) [1] has emerged as a decentralized learning paradigm that enables multiple data owners to collectively train deep learning (DL) models without sharing the raw data. Previous studies [1], [2] have showcased the viability of FL, particularly the federated averaging (Fe-dAvg) algorithm in the context of medical image segmentation. Federated Learning (FL), as a collaborative machine learning technique, aims to tackle these challenges by creating a joint predictive model across clients at multiple sites, especially within distributed medical institutions or hospitals. FL allows

for the collective training of deep learning models using distinct patient data from various hospitals for various clinical applications, including medical image segmentation. Given the non-convex nature of the training objective in deep neural networks, averaging locally trained models can result in suboptimal solutions within the parameter space, potentially leading to performance degradation [1]. However, a significant issue with FL is the performance degradation it may experience when dealing with data that are not independently and identically distributed (non-iid), a situation commonly encountered in medical image datasets. Recent publications by various authors [2], [3], [4], [5] have focused on federated deep learning, especially for segmentation and classification tasks. Wicaksana et al. [3] introduced "FedMix: Mixed Supervised Federated Learning for Medical Image Segmentation," while Xu et al. [4] presented "Federated Cross Learning for Medical Image Segmentation." Additionally, Qiu et al. [5] proposed "Federated Semi-Supervised Learning for Medical Image Segmentation via Pseudo-Label Denoising."

As the expert knowledge usually required for annotating medical images is much more demanding and difficult to obtain, various medical institutions have very limited strong pixel-level annotated images and most available images are unlabeled or weakly annotated. Therefore, a realistic clinical mechanism that utilizes every available supervision for cross-institutional collaboration without data sharing is highly desirable. We propose to develop a label-agnostic Mixed Supervised Federated Learning approach that efficiently uses data labeled in any form for medical image segmentation. Specifically, in the absence of pixel-level labels, it will effectively utilize unlabeled images as well as useful information contained in the weakly labeled for producing and selecting high-quality pseudo labels. We will devise an effective adaptive weight assignment across clients, where each client can learn an aggregation weight. Adaptive weight assignment is essential to handle inter-client variations in supervision availability. Recently, the SHINY-ICARUS challenge was organized to analyze vascular morphology and topology efficiently to provide a platform and benchmark for brain vascular segmentation [6]. Recent advancements in deep learning and imaging technology have enabled the development of advanced deep learning models and large-scale datasets [6], [7], [8]. Recently a lot of work proposed deep learning-based segmentation methods for medical imaging and signals [8], [9], [10], [11], [12]. In this work, we present an efficient hybrid attention and multi-scale feature aggregation 3D deep learning framework for automatic brain vascular segmentation. Experiments are conducted on the SHINY-ICARUS challenge which shows the effectiveness of the proposed framework.

The key contributions of this work are:

1. Present a 3D UNet-based segmentation framework aided with hybrid attention and multi-scale features for automatic brain vascular segmentation in the federated learning environment.
2. The attention-guided feature fusion module exploits the most useful features (both high-level features and low-level features) between two adjacent layers and the multiscale feature module assists the model in extracting precise features at different scales and training on each client and communicating with the server model using the weighting aggregation approach.
3. Extensive experiments were conducted on the SHINY-ICARUS challenge dataset which showed efficient segmentation of brain vascular.

2. Methodology

3D angiographies of brain vasculature segmentation especially segmentation of visible vasculature connected to one or more of the main feeding arteries of the brain is a challenging problem due to its complex nature. To effectively analyze the vascular morphology and topology, we present an efficient hybrid attention and multi-scale feature aggregation 3D deep learning framework for automatic brain vascular segmentation. In the following section, we first present the overall framework followed by details of each component. The proposed vascular segmentation architecture is based on a classical U-shape encoder–decoder structure [13], [14], [15]. Besides, we have also integrated two core modules (hybrid spatial and channel attention module, multi-scale feature extractor module) subtly and seamlessly which helps to select important multiscale contextual spatial information and semantic information adaptively. The hybrid spatial and channel attention module (HSCA) suppresses the low-level background noise and retains local semantic information of vessel structure, whereas the multiscale feature extractor module (MSFE) assists in the effective extraction of concealed multi-scale contextual information as well as aggregating multi-scale features. As a results, it helps to enhance the capability of the network to deal with complex cases where vascular heavily varies in shape and size, and many are intertwined. Figure 1 illustrates the proposed framework. To extract different feature maps from each convolutional block in the encoder, each block consists of 3D convolutional layers with batch-normalization and ReLU activation functions. We have used the 3D max-pooling layer to reduce the input spatial size of the image. Notice that we have reduced the size of spatial input with the increase in several layers and we have recovered the spatial resolution in the decoder. To recover the spatial resolution in the decoder, we have applied 3D up-sampling using bi-linear up-sampling. In this experiment, we have used $3 \times 3 \times 3$ kernel size in both the encoder and decoder and the number of feature map numbers to 16, 32, 64, 128, and 256 for each encoder. To down-sample the spatial resolution. we have set the kernel size to $2 \times 2 \times 2$ for the 3D-MaxPool layer on encoder side.

Finally, we have used a transpose3D convolutional layer with stride 2 and $2 \times 2 \times 2$ kernel size for up-sampling the size of each decoder. At the end, we have concatenated the output of each encoder block to the corresponding

decoder block. To generate the final output segmentation map, we have used a 1×1 convolutional layer with softmax function.
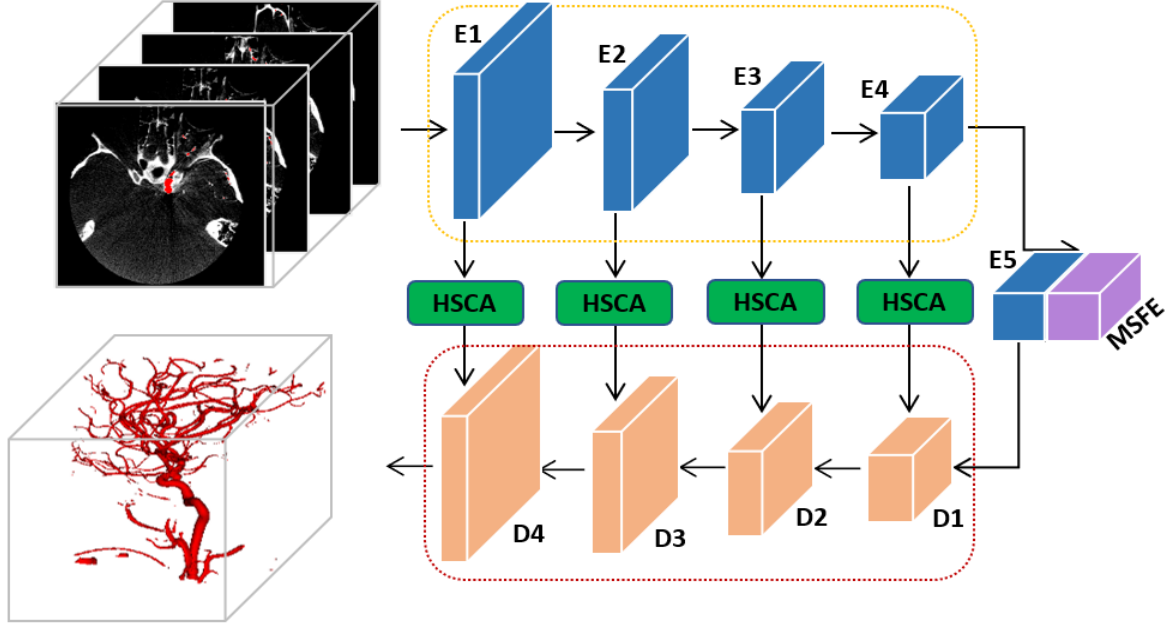
Fig. 1. Proposed Multi-scale Feature extraction Module (MSFE) and hybrid spatial and Channel attention module (HSCA) aided brain vascular segmentation framework. E1, E2, E3, E4 are encoders blocks and D1, D2, D3 and D4 are decoders blocks respectively.

Our approach follows the structure of the FedAvg [1] algorithm, a widely recognized benchmark in the field of Federated Learning (FL). This approach involves a central server node and K client nodes. The central server's main role is to manage the communication and computation processes among the various client nodes. Meanwhile, the client nodes are primarily responsible for training the model using their local data and computing devices. With a universal label set encompassing M organs, each client node possesses a local dataset D_k, which is labeled with a subset of $N_k (\leq M)$ organs. Our objective is to develop a segmentation network F(.,θ) for all M organs by leveraging the labeled datasets $\{D_k\}_{k=1}^{K}$, which are distributed across distinct client nodes and cannot be joined or centralized for training purposes. The training process within the Federated Learning (FL) framework is comprised of Tcommunication rounds between the server and client nodes. In each communication round, denoted as the t_{th} round, each client node k initially downloads the parameters of the current segmentation network, represented as $F(.,\theta^t)$ from the server (referred to as the global model). This process results in a local copy of the model, denoted as the local model.

Subsequently, the client node proceeds to train its local model using its local dataset, Dk, for a specified number of E epochs. Following the local training, the server collects the trained local models, $F(.,\theta_k^t)$, from all Kclient nodes and combines them into a new global model through a parameter-wise averaging process.

$$\theta^{t+1} = \sum_{k}^{K} \frac{|D_k|}{\sum_{j}^{K}|D_j|} \cdot \theta_t \qquad (1)$$

Since the local models are trained separately on the client nodes, the server node is only required to transmit the model parameters, as opposed to the raw data, from the clients. This approach allows the Federated Learning (FL) model to gain insights from distinct client datasets without compromising data privacy. Figure 2 showed the process of federated learning environment for brain tumor segmentation.

As labels for regions in brain vascular are sparse, hence, we have used binary cross-entropy and weighted binary cross-entropy loss [14] which can adjust the learning bias between vessels and background. Besides, we also used dice similarity coefficient loss [15] to ensure the

segmentation of small vessels. Finally, we define the 3D optimization loss function.

$$L = BCE + \alpha WBCE + (1 - \alpha)LDSC \qquad (2)$$

where α is the weight balance parameter between weighted binary cross entropy (WBCE) and LDSC (Dice loss). We have empirically set $\alpha = 0.6$.
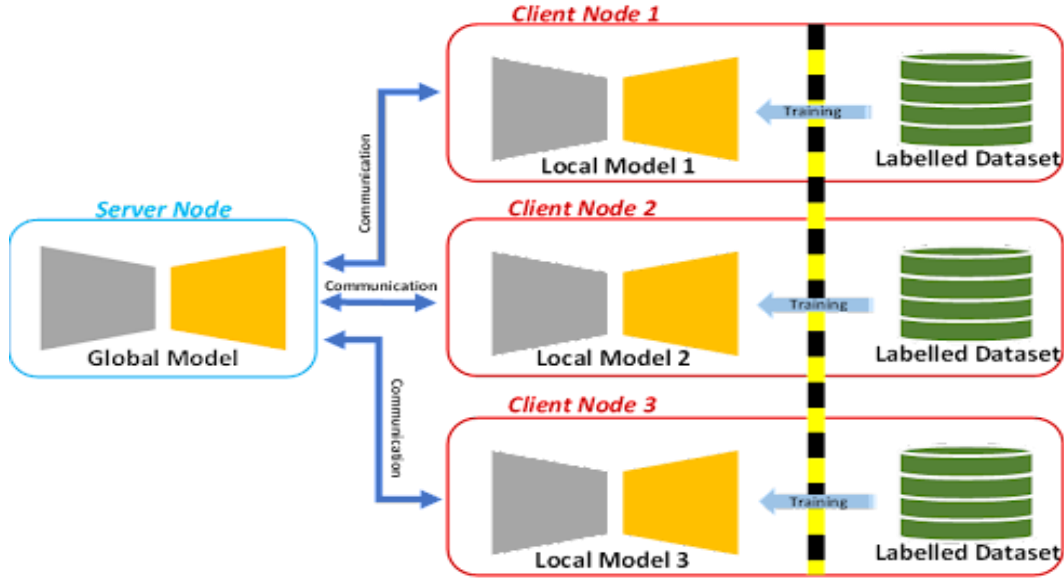
Fig. 2 Federated Learning environment for brain tumor segmentation

3. Results and Discussion

In this section, we first present details of the dataset followed by the experiment and results. We have evaluated and compared the performance using the Dice, Jaccard index, volumetric similarity coefficient, and balanced average Hausdorff distance (HD). SHINY-ICARUS challenge evaluation is based on two overlapping matrices such as clDice5 and Dice. ClDices measure the similarity while favoring both connectivity and topology. Both clDice5 and Dice carry the same weight for the evaluation.

3.1. Dataset

SHINY-ICARUS dataset consists of manually segmented cerebral arteries NIfTI volumes with 16-bit representation. The images (resolution $0.227 \times 0.227 \times 0.227$mm) are acquired by interventional neuroradiologists from patients presenting intracranial aneurysms, using a GE INNOVA 3D. The dataset consists of contract images of most of the patient's head. In this experiment, we have divided the dataset into five-fold cross-validation based on the best validation score, the proposed model uploaded in a docker container. The challenge organized provided 35 volumes for training the proposed model. A detailed description of the dataset can be found [6]. A sample dataset with manual annotation is shown in Figure 3.

(a) image   (b) mask
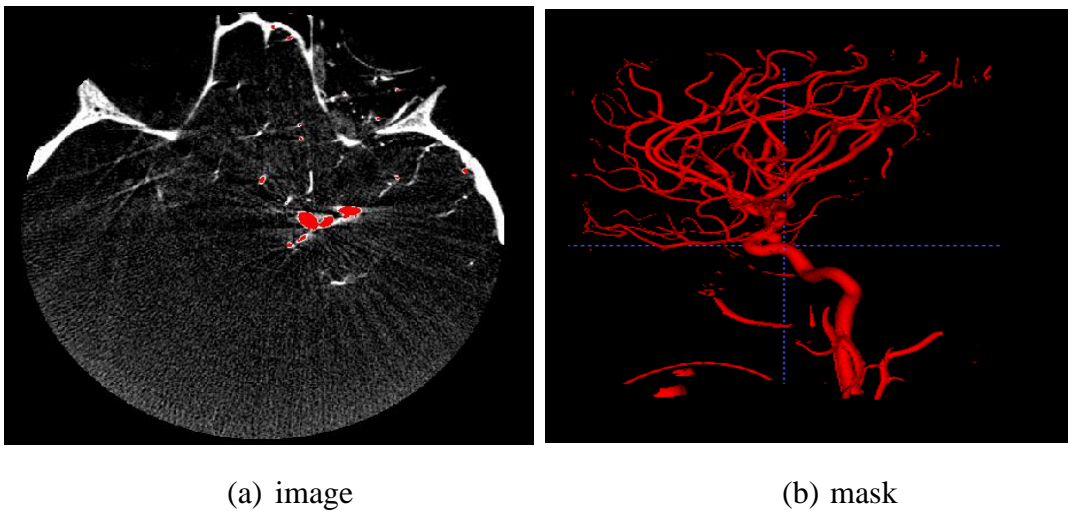
Fig. 3 A sample dataset was used in our proposed model.

3.2. Experimental Results

For evaluation purposes, the SHINY-ICARUS challenge provides two test sets (primary and secret). We have trained the proposed model using patches of input volume due to the limitation of memory. We have found $160 \times 160 \times 160$ the best patch for training. We have generated 10 random patches with a single epoch during training and optimization. In this experiment, we used Adam optimizer, batch size 2, and set the learning rate to 0.0001 and epoch to 1000. We have used a combined loss functions function between the prediction mask and the ground truth. We have used early stopping criteria and training was ended after 20 epochs on similar validation dice. We have used NVIDIA GTX 3070 GPU having a 24GB memory machine and developed the model in PyTorch library and trained from scratch. Figure 4 illustrates the visualization of the validation set. Table 1 shows the results of the proposed model using the validation dataset.

Figure 4 shows the predicted and ground truth segmentation for subject1 and subject2 validation datasets. The proposed model produced a similar segmentation mask as compared to ground truth segmentation. Our proposed model produced overestimated some branches and we can remove these extra segmentation branches using connected components.

Table 1. Results on the proposed model on the validation dataset.

Models	Dice	Re	Pr	clDice	clRe	clPr
3DUNet+MSFE	0.8674	0.9990	0.7664	0.8496	0.9988	0.7392
3DUNet+HSCA	0.8969	0.9882	0.8210	0.8573	0.9890	0.7566
3DUNet+HSCA+MSFE	0.9374	0.9432	0.9317	0.8968	0.9567	0.8440

GT input   GT volume   GT input   GT volume

Predict   Predict Volume   Predict   Predict Volume
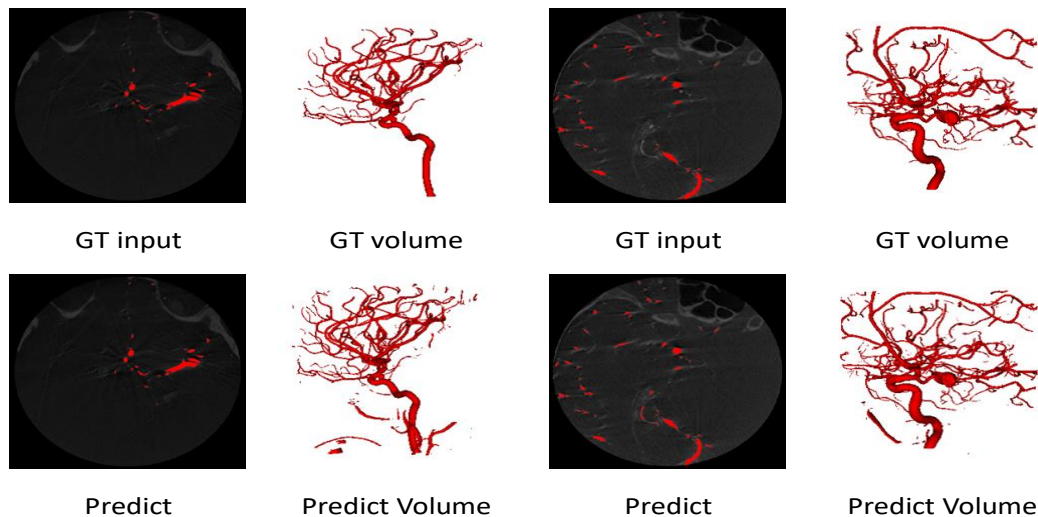
Fig. 4 The first shows the prediction and ground truth (GT) segmentation mask for validation subject 1 and the second row shows results on subject 2.

4. Conclusion

In this paper, we presented an efficient brain vascular segmentation deep framework aided with hybrid attention and multi-scale feature aggregation modules. Experiments are conducted on the SHINY-ICARUS challenge which shows that our model is ranked 2nd on the secret test and 3rd on the combo test that validates the effectiveness of the proposed framework. We have divided the dataset into five-fold cross-validation and based on the best validation score, the proposed model was uploaded in a docker container. Experiment results showed efficient (dice and clDice5) segmentation of brain vascular.

References

1. McMahan, Brendan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. "Communication-efficient learning of deep networks from decentralized data." In Artificial intelligence and statistics, pp. 1273-1282. PMLR, 2017.
2. Liu, Quande, Cheng Chen, Jing Qin, Qi Dou, and Pheng-Ann Heng. "Feddg: Federated domain generalization on medical image segmentation via episodic learning in continuous frequency space." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 1013-1023. 2021.
3. Wicaksana, Jeffry, Zengqiang Yan, Dong Zhang, Xijie Huang, Huimin Wu, Xin Yang, and Kwang-Ting Cheng. "FedMix: Mixed supervised federated learning for medical image segmentation." IEEE Transactions on Medical Imaging (2022).
4. Xu, Xuanang, Tianyi Chen, Han Deng, Tianshu Kuang, Joshua C. Barber, Daeseung Kim, Jaime Gateno, Pingkun Yan, and James J. Xia. "Federated cross learning for medical image segmentation." arXiv preprint arXiv:2204.02450 (2022).
5. Qiu, Liang, Jierong Cheng, Huxin Gao, Wei Xiong, and Hongliang Ren. "Federated Semi-Supervised Learning for Medical Image Segmentation via Pseudo-Label Denoising." IEEE Journal of Biomedical and Health Informatics (2023).
6. Moccia, S. et al., "Blood vessel segmentation algorithms—review of methods, datasets and evaluation metrics," Computer methods and programs in biomedicine 158, 71–91 (2018).
7. Payette, Kelly, Hongwei Bran Li, Priscille de Dumast, Roxane Licandro, Hui Ji, Md Mahfuzur Rahman Siddiquee, Daguang Xu et al. "Fetal brain tissue annotation and segmentation challenge results." Medical Image Analysis 88 (2023): 102833.
8. De Vente, Coen, Koenraad A. Vermeer, Nicolas Jaccard, He Wang, Hongyi Sun, Firas Khader, Daniel Truhn et al. "AIROGS: Artificial Intelligence for robust glaucoma screening challenge." IEEE Transactions on Medical Imaging (2023).
9. Chen, Zhihao, Alain Lalande, Michel Salomon, Thomas Decourselle, Thibaut Pommier, Abdul Qayyum, Jixi Shi, Gilles Perrot, and Raphaël Couturier. "Automatic deep learning-based myocardial infarction segmentation from delayed enhancement MRI." Computerized Medical Imaging and Graphics 95 (2022): 102014.
10. Qayyum, Abdul, Mona Mazhar, Imran Razzak, and Mohamed Reda Bouadjenek. "Multilevel depth-wise context attention network with atrous mechanism for segmentation of COVID19 affected regions." Neural Computing and Applications (2021): 1-13.
11. Ahmad, Iftikhar, Abdul Qayyum, Brij B. Gupta, Madini O. Alassafi, and Rayed A. AlGhamdi. "Ensemble of 2D residual neural networks integrated with atrous spatial pyramid pooling module for myocardium segmentation of left ventricle cardiac MRI." Mathematics 10, no. 4 (2022): 627.
12. Qayyum, Abdul, Aamir Malik, Naufal M Saad, and Moona Mazher. "Designing deep CNN models based on sparse coding for aerial imagery: a deep-features reduction approach." European Journal of Remote Sensing 52, no. 1 (2019): 221-239.
13. Ahmad, R. F., Malik, A. S., Kamel, N., Amin, H., Zafar, R., Qayyum, A., & Reza, F. (2014, November). Discriminating the different human brain states with EEG signals using Fractal dimension: A nonlinear approach. In 2014 IEEE International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA) (pp. 1-5). IEEE.
14. Qayyum, Abdul, Moona Mazher, Tariq Khan, and Imran Razzak. "Semi-supervised 3D-InceptionNet for segmentation and survival prediction of head and neck primary cancers." Engineering Applications of Artificial Intelligence 117 (2023): 105590.
15. Eisenmann, Matthias, Annika Reinke, Vivienn Weru, Minu Dietlinde Tizabi, Fabian Isensee, Tim J. Adler, Patrick Godau et al. "Biomedical image analysis competitions: The state of current participation practice." arXiv preprint arXiv:2212.08568 (2022)

Authors Introduction

Dr. Moona Mazher

She is a senior postdoc research fellow at Department of Computer Science, University College London. She received her Ph.D. from the University of Rovira i Virgili, Spain, 212 a specialization in Neuroscience from Universiti Teknologi PETRONAS, Malaysia in 2017.. Her areas of interest are machine learning, deep learning, medical imaging, signal processing, computer vision, and explainable AI.

Dr. Abdul Qayyum

He is currently working at National Heart and Lung Institute Imperial College London, UK. Previously, he was joined as lecturer at University of Bourgogne Franche-Comté France. He received his Ph.D in electrical & electronics engineering with specialization in deep learning and image processing in 2017 from Universiti Teknologi Petronas Malaysia. His area of interest is machine learning, deep learning and quantum machine learning for signal processing and bioemdical imaging..

Dr M. K. A. Ahamed Khan

He is currently working at UCSI University, Malaysia. He received his Ph.D in Robotics and controls from USA. His area of research is robotics, AI and controls. He has published more than100 papers. He is also an IEEE Senior member. He is also the past chair for IEEE RAS Malaysia chapter.

.

Prof. Steven Niederer

He completed his undergraduate degree in Engineering Science at the University of Auckland in 2003 and his DPhil at the University of Oxford in 2008. In 2023, he moved to Imperial College London as the Chair in Biomedical Engineering at the National Heart and Lung Institute. His current work is focused on reducing barriers to adopting digital twin technology, developing virtual patient cohorts for in-silico trials, mapping organ scale function through to cellular and molecular physiology, and using modelling and simulation to personalize and guide therapies.

Dr . Mastaneh Mokayef

He is currently working at National Heart and Lung Institute Imperial College London, UK. Previously, he was joined as lecturer at University of Bourgogne Franche-Comté France. He received his Ph.D in electrical & electronics engineering with specialization in deep learning and image processing in 2017 from Universiti Teknologi Petronas Malaysia. His area of interest is machine learning, deep learning and quantum machine learning for signal processing and bioemdical imaging.

Ts Amar Ridzuan Bin Abd Hamid

He is a lecturer of Mechanical and Mechatronic programmes from Department of Mechanical Engineering, UCSI University, Malaysia. He has completed his Master Degree from Universiti Putra Malaysia, Postgraduate Diploma from UCSI, and a Bachelor Degree with Hounours in Mechanical Engineering (Automotive) from Universiti Teknikal Malaysia Melaka (UTeM), Malaysia.

Dr Cik Suhana Hassan

She currently works at UCSI University. She received her PhD from UTP Malaysia. She has been doing research in Mechanical Engineering and Materials Engineering. She is passionate about turning environmental waste into value-added products as part of her quest to live a more environmentally friendly life.

Hybrid Classical and Quantum Deep Learning Models for Medical Image Classification

Moona Mazher
Centre for Medical Image Computing, Department of Computer Science, University College London, UK
Email: m.mazher@ucl.ac.uk

Abdul Qayyum
National Heart & Lung Institute, Imperial College, London, United Kingdom

M. K. A. Ahamed Khan
UCSI University, Faculty of Engineering, Malaysia

Steven Niederer
National Heart & Lung Institute, Imperial College, London, United Kingdom
Alan Turing Institute, London, United Kingdom

Mastaneh Mokayef
UCSI University, Faculty of Engineering, Malaysia

C. S. Hassan
UCSI University, Faculty of Engineering, Malaysia

M. Ridzuan, A.
UCSI University, Faculty of Engineering, Malaysia

Abstract

Quantum machines enhance the capabilities of classical counterparts across various domains, notably in addressing real-world challenges. The classification of brain MR images for tumor detection is a crucial diagnostic process in the analysis of brain images. Traditional approaches, such as classical machine learning techniques and conventional deep learning structures like convolutional neural networks, are frequently employed for image classification. However, as the network size increases, training these models becomes increasingly arduous. Quantum algorithms offer advantages by optimizing the performance of classical algorithms through the incorporation of the intrinsic properties of quantum bits. In this paper, we proposed a hybrid classical and quantum convolutional neural network for Alzheimer's disease (AD) classification. The proposed model was further validated on the brain tumor classification task. The fundamental concept involves encoding data into quantum states, facilitating quicker information extraction, and subsequently utilizing this information to discern the data class. The proposed model results underscore the reliability and robustness and demonstrated by optimal performance accuracies across various datasets, the proposed model substantiates its efficacy in detecting and classifying AD disease and brain tumors.

Keywords: Quantum machine learning, Deep learning, Classification, Alzheimer's disease, Brain Tumour, Hybrid Quantum technique in medical imaging.

1. Introduction

Alzheimer's disease (AD) is the most common reason of dementia worldwide [1]. Dementia belongs to assess abnormal changes in the brain and commonly interrupts the communication between the brain cells [1]. Disrupted communication between cells can lead to impairments in an individual's cognitive functions, including memory loss, emotional regulation, reasoning and decision-making, behavior, and language proficiency [1]. In recent years, substantial progress has been achieved in the advancement of cerebrospinal fluid (CSF) biomarkers [2] and cutting-edge imaging techniques like amyloid and tau positron emission tomography (PET) [2]. Although there have been significant advancements, numerous emerging diagnostic and treatment approaches are still primarily confined to research settings. As a result, the primary means of diagnosis before death continues to rely on traditional clinical evaluation, neuropsychological testing [10] and magnetic resonance imaging (MRI) [11]. Mild cognitive impairment (MCI), which serves as an early precursor to dementia, can also represent a subtle initial manifestation of Alzheimer's disease (AD). Diagnosing MCI as an early sign of AD demands a high level of clinical expertise from skilled specialists. Alzheimer's disease (AD) is a neurological disorder that can be detected through brain imaging, and numerous

Moona Mazher, Abdul Qayyum, M. K. A. Ahamed Khan, Steven Niederer, Mastaneh Mokayef, C. S. Hassan, M. Ridzuan, A.

studies have concentrated on using machine learning or deep learning methods to classify AD based on brain images [3].

Recently, various studies proposed deep learning models for AD classifications [3], [4]. They proposed traditional deep learning classification models, in this study we investigate a hybrid model based on classical deep learning and quantum machine learning. Recently Quantum machine learning research has gained success in the medical community.

Few studies proposed quantum machine learning for brain tumor classification [5]. We have proposed hybrid deep learning and quantum layers for the classification of AD and compared the performance with state-of-the-art methods using a hybrid proposed approach. We proposed first-time hybrid classical deep learning and quantum machine learning models for AD classification. We have compared performance using the brain tumor dataset using our proposed technique.

2. Methodology

Our proposed approach combines a classical network with a quantum network to harness the strengths of both worlds, creating a model that can effectively identify AD

abnormalities using MRI analysis. The architecture primarily emphasizes the utilization of classical algorithms for enhanced detection, thereby expediting the diagnosis and treatment processes. The substructure integrated into constructing a quantum layer aims to transform classical data points into quantum states, facilitating faster information extraction and improving the efficiency of feature detection and pattern recognition.

The Convolutional Neural Network (CNN) is a widely employed architecture for a range of processing tasks, such as image recognition, segmentation, and classification [6], [7], [8], [9], [10], [11], [12]. It is composed of three basic layers: Convolution, pooling, and fully connected layers. The convolution and pooling layers are involved in feature extraction from input images, while the fully connected layer extracts features to the output, facilitating classification based on the identified features. A typical CNN structure comprises multiple blocks that integrate convolution, pooling, and fully connected layers. We have proposed three blocks of conv layers. Each block consisted of 2D Conv, 2D MaxPool, and 2D Normalization layers [13], [14], [15]. We have used one fully connected (FC) layer and one hybrid quantum layer for AD and brain tumor classification. The structure of the proposed model is shown in Figure 1.
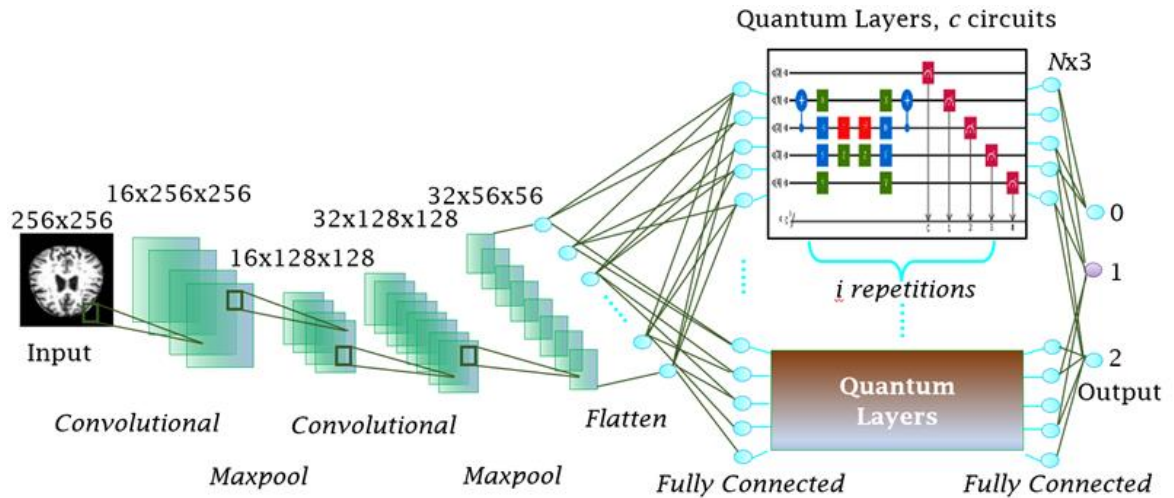
Fig. 1 The structure of the proposed classical and Quantum deep learning network.

3. Quantum Layers

Quantum supremacy manifests when solving intricate problems more effectively than traditional classical methods [5]. Quantum mechanics provides the fundamental framework for the operation of quantum machines. In a quantum computer, the fundamental unit of information processing is a quantum bit, or qubit, akin to the role played by classical bits in conventional computers. A standard quantum computer utilizes the

unique features of a quantum bit, primarily superposition and quantum entanglement, to process data and information. Unlike a classical bit, a qubit's superposition property enables it to simultaneously exist in multiple states, including zero and one. In essence, qubits can be described as state vectors within Hilbert space. The hybrid layer has been proposed with classical deep learning models for AD classification and further compared on the brain tumor classification dataset.

Figure 2 shows the complete block architecture of the quantum layers.
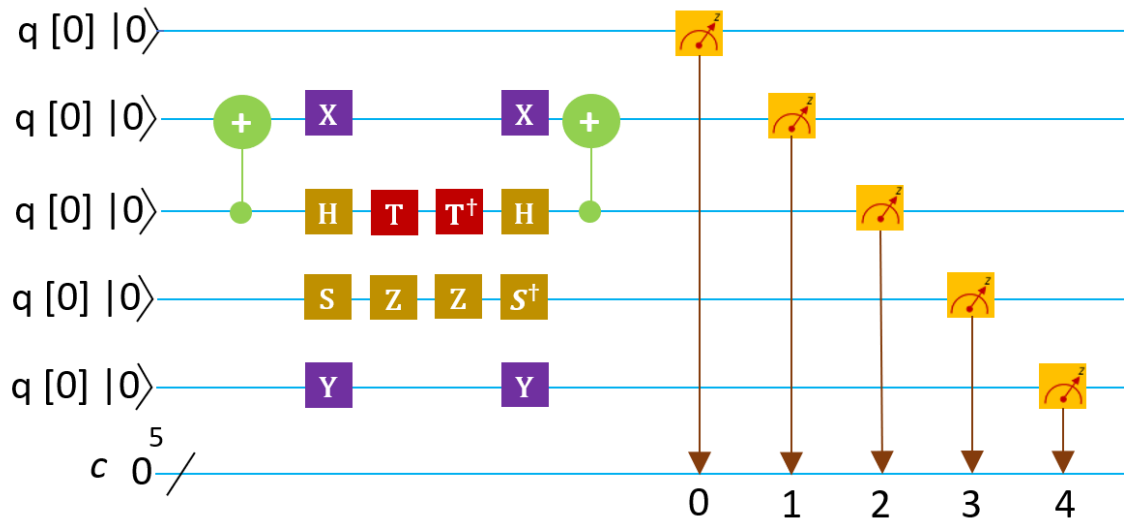
Fig. 2 Block of the quantum layer.

4. Dataset

We have used an Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset, consisting of 1,821 participants based on magnetic resonance imaging (MRI) scans. After rigorous inclusion criteria were applied, the dataset comprised a total of 8,916 participants. Subjects were labeled according to the clinical diagnoses provided by the ADNI study cohort. For any subjects with documented dementia and a primary diagnosis of Alzheimer's disease dementia, an AD label was assigned regardless of the presence of additional dementing comorbidities. We compared the performance of our proposed method using a brain tumor dataset. All algorithms undergo training and performance validation using the Kaggle Brain Tumor MRI dataset. The Kaggle Brain dataset comprises 7,023 human brain MRI scan images in DICOM format, categorized into four distinct classes, namely glioma, meningioma, no tumor, and pituitary [5].

5. Results and Discussion

In this study, a series of tests have been performed to assess the benefits of employing quantum circuits in AD disease and brain cancer classification. These tests specifically compared the performance of a traditional CNN with quantum circuits used for feature extraction, with a focus on highlighting the advantages of quantum circuits. The proposed approach with Quantum layers produced comparable performance using both datasets as shown in Table 1.

Table 1 Classification performance of brain tumor using classical and quantum deep learning models

Methods	Accuracy	precision	recall	F1-score
A. Pashaei et al. [16]	93.68	--	--	--
Classical CNN	94.33	95.07	94.33	94.98
Classical and Quantum CNN	95.33	96.34	95.88	94.23

Table 2 The Alzheimer classification using classical and quantum deep learning models.

Methods	Accuracy	precision	recall	F1-score
Di Wang et al. [17]	87.25	--	--	--
Classical CNN	88.33	88.90	89.90	90.39
Classical and Quantum CNN	89.56	90.87	91.63	91.22

Moona Mazher, Abdul Qayyum, M. K. A. Ahamed Khan, Steven Niederer, Mastaneh Mokayef, C. S. Hassan, M. Ridzuan, A.

Table 2 shows the performance of the proposed model using classical and hybrid deep learning models. The quantum-based hybrid deep learning model produced optimal performance on brain tumor and AD disease classification tasks.

Furthermore, the results found have the capacity of the hybrid quantum model to enhance the efficiency of medical image classification and diagnosis. However, additional research is necessary to evaluate its performance in classifying different types of data. This research offers valuable insights into the benefits of integrating quantum circuitry into CNN models for the analysis of medical images, opening possibilities for quantum-enhanced machine learning within this field.

6. Conclusion

The proposed study introduces an efficient Hybrid Quantum deep learning model for the classification and diagnosis of AD disease and Kaggle brain tumors medical image datasets. The proposed model attains a comparable performance accuracy as compared to the conventional CNN's performance. The validation accuracy results provide additional evidence of this superior performance. Future research efforts will be directed towards assessing the resilience of hybrid quantum-based models against adversarial attacks. The research on quantum with classical deep learning models will continue to enrich the body of knowledge in quantum computing and its potential applications in medical imaging research, ultimately driving progress in precise diagnosis and enhanced patient care.

References

1. Nichols, Emma, Jaimie D. Steinmetz, Stein Emil Vollset, Kai Fukutaki, Julian Chalek, Foad Abd-Allah, Amir Abdoli et al. "Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019." The Lancet Public Health 7, no. 2 (2022): e105-e125.
2. Palmqvist, Sebastian, Shorena Janelidze, Yakeel T. Quiroz, Henrik Zetterberg, Francisco Lopera, Erik Stomrud, Y. I. Su et al. "Discriminative accuracy of plasma phospho-tau217 for Alzheimer disease vs other neurodegenerative disorders." Jama 324, no. 8 (2020): 772-781.
3. Mahendran, Nivedhitha, PM Durai Raj Vincent, Kathiravan Srinivasan, and Chuan-Yu Chang. "Improving the classification of alzheimer's disease using hybrid gene selection pipeline and deep learning." *Frontiers in genetics* 12 (2021): 784814.
4. Qiu, Shangran, Matthew I. Miller, Prajakta S. Joshi, Joyce C. Lee, Chonghua Xue, Yunruo Ni, Yuwei Wang et al. "Multimodal deep learning for Alzheimer's disease dementia assessment." *Nature communications* 13, no. 1 (2022): 3404.
5. Ajlouni, Naim, Adem Özyavaş, Mustafa Takaoğlu, Faruk Takaoğlu, and Firas Ajlouni. "Medical Image Diagnosis Based on Adaptive Hybrid Quantum CNN." (2023).
6. Payette, Kelly, Hongwei Bran Li, Priscille de Dumast, Roxane Licandro, Hui Ji, Md Mahfuzur Rahman Siddiquee, Daguang Xu et al. "Fetal brain tissue annotation and segmentation challenge results." Medical Image Analysis 88 (2023): 102833.
7. Qayyum, Abdul, Iftikhar Ahmad, Mohsin Iftikhar, and Moona Mazher. "Object Detection and Fuzzy-Based Classification Using UAV Data." Intelligent Automation & Soft Computing 26, no. 4 (2020).
8. De Vente, Coen, Koenraad A. Vermeer, Nicolas Jaccard, He Wang, Hongyi Sun, Firas Khader, Daniel Truhn et al. "AIROGS: Artificial Intelligence for robust glaucoma screening challenge." IEEE Transactions on Medical Imaging (2023).
9. Chen, Zhihao, Alain Lalande, Michel Salomon, Thomas Decourselle, Thibaut Pommier, Abdul Qayyum, Jixi Shi, Gilles Perrot, and Raphaël Couturier. "Automatic deep learning-based myocardial infarction segmentation from delayed enhancement MRI." Computerized Medical Imaging and Graphics 95 (2022): 102014.
10. Qayyum, Abdul, Mona Mazher, Imran Razzak, and Mohamed Reda Bouadjenek. "Multilevel depth-wise context attention network with atrous mechanism for segmentation of COVID19 affected regions." Neural Computing and Applications (2021): 1-13.
11. Ahmad, Iftikhar, Abdul Qayyum, Brij B. Gupta, Madini O. Alassafi, and Rayed A. AlGhamdi. "Ensemble of 2D residual neural networks integrated with atrous spatial pyramid pooling module for myocardium segmentation of left ventricle cardiac MRI." Mathematics 10, no. 4 (2022): 627.
12. Qayyum, Abdul, Aamir Malik, Naufal M Saad, and Moona Mazher. "Designing deep CNN models based on sparse coding for aerial imagery: a deep-features reduction approach." European Journal of Remote Sensing 52, no. 1 (2019): 221-239.
13. Ahmad, R. F., Malik, A. S., Kamel, N., Amin, H., Zafar, R., Qayyum, A., & Reza, F. (2014, November). Discriminating the different human brain states with EEG signals using Fractal dimension: A nonlinear approach. In 2014 IEEE International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA) (pp. 1-5). IEEE.
14. Qayyum, Abdul, Moona Mazher, Tariq Khan, and Imran Razzak. "Semi-supervised 3D-InceptionNet for segmentation and survival prediction of head and neck primary cancers." Engineering Applications of Artificial Intelligence 117 (2023): 105590.
15. Eisenmann, Matthias, Annika Reinke, Vivienn Weru, Minu Dietlinde Tizabi, Fabian Isensee, Tim J. Adler, Patrick Godau et al. "Biomedical image analysis competitions: The state of current participation practice." arXiv preprint arXiv:2212.08568 (2022).
16. A. Pashaei, H. Sajedi and N. Jazayeri, "Brain tumor classification via convolutional neural network and

extreme learning machines", Proc. 8th Int. Conf. Comput. Knowl. Eng. (ICCKE), pp. 314-319, Oct. 2018.

17. Wang, Di, Nicolas Honnorat, Peter T. Fox, Kerstin Ritter, Simon B. Eickhoff, Sudha Seshadri, Mohamad Habes, and Alzheimer's Disease Neuroimaging Initiative. "Deep neural network heatmaps capture Alzheimer's disease patterns reported in a large meta-analysis of neuroimaging studies." Neuroimage 269 (2023): 119929.

Authors Introduction

Dr. Moona Mazher

She is a senior postdoc research fellow at Department of Computer Science, University College London. She received her Ph.D. from the University of Rovira i Virgili, Spain, 212 a specialization in Neuroscience from Universiti Teknologi PETRONAS, Malaysia in 2017.. Her areas of interest are machine learning, deep learning, medical imaging, signal processing, computer vision, and explainable AI.

Dr . Abdul Qayyum

He is currently working at National Heart and Lung Institute Imperial College London, UK. Previously, he was joined as lecturer at University of Bourgogne Franche-Comté France. He received his Ph.D in electrical & electronics engineering with specialization in deep learning and image processing in 2017 from Universiti Teknologi Petronas Malaysia. His area of interest is machine learning, deep learning and quantum machine learning for signal processing and bioemdical imaging.

Dr M. K. A. Ahamed Khan

He is currently working at UCSI University, Malaysia. He received his Ph.D in Robotics and controls from USA. His area of research is robotics, AI and controls. He has published more than100 papers. He is also an IEEE Senior member. He is also the past chair for IEEE RAS Malaysia chapter.

Prof. Steven Niederer

He completed his undergraduate degree in Engineering Science at the University of Auckland in 2003 and his DPhil at the University of Oxford in 2008. In 2023, he moved to Imperial College London as the Chair in Biomedical Engineering at the National Heart and Lung Institute. His current work is focused on reducing barriers to adopting digital twin technology, developing virtual patient cohorts for in-silico trials, mapping organ scale function through to cellular and molecular physiology, and using modelling and simulation to personalize and guide therapies.

Dr . Mastaneh Mokayef

He is currently working at National Heart and Lung Institute Imperial College London, UK. Previously, he was joined as lecturer at University of Bourgogne Franche-Comté France. He received his Ph.D in electrical & electronics engineering with specialization in deep learning and image processing in 2017 from Universiti Teknologi Petronas Malaysia. His area of interest is machine learning, deep learning and quantum machine learning for signal processing and bioemdical imaging.

Dr Cik Suhana Hassan

She currently works at the Faculty of Engineering, Technology and Built Environment, UCSI University. She received her PhD from UTP Malaysia. Suhana does research in Mechanical Engineering and Materials Engineering. Dr. Suhana is passionate about turning environmental waste into value-added products as part of her quest to live a more environmentally friendly life.

Ts Amar Ridzuan Bin Abd Hamid

.He is a lecturer of Mechanical and Mechatronic programmes from Department of Mechanical Engineering, UCSI University, Malaysia. He has completed his Master Degree from Universiti Putra Malaysia, Postgraduate Diploma from UCSI, and a Bachelor Degree with Hounours in Mechanical Engineering (Automotive) from Universiti Teknikal Malaysia Melaka (UTeM), Malaysia.

Global Stabilization of A Class of Nonholonomic Integrators via Discontinuous Control

Lixia Yan

The Seventh Research Division and the Center for Information and Control, School of Automation Science and Electrical Engineering, Beihang University (BUAA), Beijing 100191, China

Yingmin Jia

The Seventh Research Division and the Center for Information and Control, School of Automation Science and Electrical Engineering, Beihang University (BUAA), Beijing 100191, China
Email: yanlixia@buaa.edu.cn, ymjia@buaa.edu.cn

Abstract

This paper investigates the discontinuous state feedback control for stabilizing a class of nonholonomic integrators with drift terms. The control design relies on constraining state trajectory in an invariant set. To this end, we apply constant controls to drive the states moving into the invariant set and then switch to a continuous control law with suitable gain selections. It is proven in the Lyapunov sense that the proposed control scheme achieves global exponential stabilization of the states, and the control switch would only occur at most once. Numerical simulations are carried out to validate the proposed control law.

Keywords: Nonholonomic integrators, Discontinuous feedback, Nonlinear control

1. Introduction

Nonholonomic integrators refer to one kind of control systems that obstruct Brockett's necessary condition for the existence of static time-invariant stabilizers [1]. They have attracted much attention in the control community because various robotic systems can be converted into nonholonomic integrators [2].

Due to non-integrable properties, only discontinuous, time-varying, or hybrid control laws are applicable to stabilize the nonholonomic integrators [3], [4], [5], [6]. Via forcing the state trajectories to move on a sliding mode surface, a stabilization and a tracking control law were developed in [3] for nonholonomic integrators, achieving global asymptotical convergence of the states. In [4], a novel logic-based hybrid control law with the switching mechanism that achieved global exponential stabilization was reported. Using virtual control and variable structure design, the control scheme in [5] stabilized the states of nonholonomic integrators to zero exponentially from any initial states. An alternative with the control Lyapunov function approach for stabilizing the nonholonomic integrators can be found in [6]. The leading results in [3], [4], [5], [6] are helpful in understanding the structural properties and solution trajectory of nonholonomic integrators. Yet, the drift terms that play affect on the dynamic performance of the

system are not considered in the literature mentioned above.

It is necessary to consider drift terms for utilizing the results developed for pure nonholonomic integrators on practical robotic systems. In [2], an adaptive leader-following formation control scheme of multiple wheeled mobile robots was developed by applying techniques associated with nonholonomic integrators, regulating formation errors globally convergent to the neighborhood of the origin. After converting the kinematic model of unicycles into the form of nonholonomic integrators with drift terms, a control law capable of rendezvous and tracking of networked unicycles was proposed in [7], which, however, only consider the case that the error states are initialized inside an invariant set.

Motivated by the discussions above, this paper makes further endeavors on the control design of nonholonomic integrators. The concerned nonholonomic integrators can be viewed as an augmented version of that in [1] by adding a drift term. The control design involves a state feedback control law and a constant control law, which achieves asymptotical convergence of the states in an invariant set and forces the state trajectory to move into the invariant set, respectively. A simple switch would occur if the initial states are outside the invariant set. Lyapunov stability theory is utilized to prove the obtained theoretical results.

The rest is organized as follows. Section 2 formulates the control problem and presents the control design.

Numerical simulations are carried out in Section 3. Section 4 concludes the work briefly.

2. Main results
2.1. Problem formulation

The concerned nonholonomic integrator with a drift term in this work holds the form below,

$$
\begin{cases}
\dot{z}_1 = u_1 \\
\dot{z}_2 = u_2 \\
\dot{z}_3 = z_2 u_1 - z_1 u_2 + f(z_1, z_2)
\end{cases} \tag{1}
$$

where $z_1, z_2, z_3, u_1, u_2 \in \Box, f : \Box \times \Box \mapsto \Box$, and there is a positive number κ_1 such that $|f(z_1, z_2)| \leq \kappa_1 \|[z_1, z_2]\|$. The nonholonomic dynamic (1) obstructs the famous Brockett's necessary condition for the existence of full-state time-invariant static stabilizer. To this end, we plan to find a discontinuous control law for (1) so that

$$
\lim_{t \to \infty} z_1 = 0, \lim_{t \to \infty} z_2 = 0, \lim_{t \to \infty} z_3 = 0 \tag{2}
$$

from any initial states.

Remark 1. Either kinematic or dynamic models of various nonholonomic systems, including nonholonomic unicycles and underactuated hovercrafts [7], can be converted into the form of (1). Thus, the addressed nonholonomic integrator (1) has general property though its form is simple.

2.2. Control design

The discontinuous control design includes two steps. First, we design a state feedback control law so that the state trajectory is convergent to zero in an invariant set. Second, we propose a constant control to force the states to move into the invariant set. The design process above is depicted in Fig. 1.

$$
C \triangleq \left\{ Z = [z_1, z_2, z_3]^T \in \mathbb{R}^3 \middle| z_1^2 + z_2^2 > |z_3| \right\}
$$
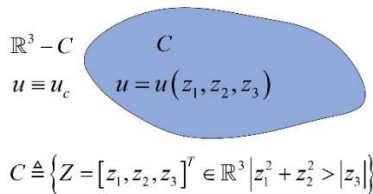
Fig. 1. The invariant set and switching mechanism.

Design the following control law,

$$
\begin{aligned}
u_1 &= -k_1 \tanh z_1 - \frac{z_2}{z_1^2 + z_2^2}(f + g + k_2 z_3) \\
u_2 &= -k_1 \tanh z_2 + \frac{z_1}{z_1^2 + z_2^2}(f + g + k_2 z_3)
\end{aligned} \tag{3}
$$

where $g \Box -k_1 z_2 \tanh z_1 + k_1 z_1 \tanh z_2, k_2 > 2k_1 > 0$. The effectiveness of the control law (3) is summarized in the lemma below.

LEMMA 1. *The application of (3) on (1) achieves that $z_1 \to 0, z_2 \to 0, z_3 \to 0$ as $t \to +\infty$ and $u_1, u_2 \in L_\infty$, if $k_2 > 2k_1 > 0$ and $Z(0) \in C \Box \left\{ Z \in \Box^3 \middle| z_1^2 + z_2^2 > |z_3| \right\}$.*

Proof. Substituting (3) into (1) results in,

$$
\begin{aligned}
\dot{z}_1 &= -k_1 \tanh z_1 - \frac{z_2}{z_1^2 + z_2^2}(f + g + k_2 z_3) \\
\dot{z}_2 &= -k_1 \tanh z_2 + \frac{z_2}{z_1^2 + z_2^2}(f + g + k_2 z_3) \\
\dot{z}_3 &= -k_2 z_3
\end{aligned} \tag{4}
$$

Obviously, one has $z_3(t) = z_3(0)e^{-k_2 t}$ and obtains that z_3 would converge to zero exponentially. Choose a positive function,

$$
V = 0.5(z_1^2 + z_2^2) \tag{5}
$$

The time derivative of V along with the solution trajectory of (4) can be calculated as,

$$
\dot{V} = -k_1 z_1 \tanh z_1 - k_1 z_2 \tanh z_2 \leq 0 \tag{6}
$$

which implies that the states z_1 and z_2 converge to zero asymptotically. Next, we prove the boundedness of the control inputs and invariant properties of the set C. Due to the forms of f and g, it is direct to obtain

$$
\begin{aligned}
|f| &\leq \kappa_1 \sqrt{z_2^2 + z_2^2} \\
|g| &\leq k_1 |z_1| + k_1 |z_2| \leq 2k_1 \sqrt{z_2^2 + z_2^2}
\end{aligned} \tag{7}
$$

which, together with (3), demonstrates that

$$
\begin{aligned}
|u_1| &\leq k_1 + \frac{|f| + |g|}{\sqrt{z_1^2 + z_2^2}} + \frac{k_2 |z_3|}{\sqrt{z_1^2 + z_2^2}} \\
&\leq 3k_1 + \kappa_1 + \frac{k_2 |z_3|}{\sqrt{z_1^2 + z_2^2}}
\end{aligned} \tag{8}
$$

According to $|x| \geq |\tanh x|, \forall x \in \Box$, we derive that

$$
\dot{V} \geq -k_1(z_1^2 + z_2^2) = -2k_1 V \tag{9}
$$

Using the comparison principle then yields [8],

$$
V(t) \geq e^{-2k_1 t} V(0) \tag{10}
$$

Therefore, one has

$$
\|[z_1(t), z_2(t)]\| \geq \|[z_1(0), z_2(0)]\| e^{-k_1 t} \tag{11}
$$

and

$$
\frac{|z_3|}{\sqrt{z_1^2 + z_2^2}} \leq \frac{|z_3(0)|}{\sqrt{z_1^2(0) + z_2^2(0)}} \tag{12}
$$

Combining (8) and (12), we find out that

$$
|u_1| \leq U \Box 3k_1 + \kappa_1 + k_2 \frac{|z_3(0)|}{\sqrt{z_1^2(0) + z_2^2(0)}} \tag{13}
$$

and u_1 is bounded. Following the same routine above, we can also prove the boundedness $|u_2| \leq U$. Additionally, the inequality (11) and the fact $z_3(t) = z_3(0)e^{-k_2 t}$ imply that $z_1^2 + z_2^2 > |z_3|, \forall t \geq 0$ as $k_2 > 2k_1$ and $Z(0) \in C$.

Therefore, the set C is invariant. This completes the proof. □

The Lemma 1 shows that the state trajectory would converge to zero asymptotically in the invariant set C. In what follows, we would like to use constant controls to drive the state trajectory moving into the set C if the states are not initialized therein.

Without losing generality, the initial states are assumed to be outside the invariant set C, i.e., $Z(0) \in \Box^3 - C$. Consider the constant control inputs,

$$u_1 \equiv u_{1c}, u_2 \equiv u_{2c} \qquad (14)$$

Using (14) and integrating (1) with respect to time then lead to,

$$z_1(t) = z_1(0) + u_{1c}t$$
$$z_2(t) = z_2(0) + u_{2c}t$$
$$z_3(t) = z_3(0) + [z_2(0)u_{1c} - z_1(0)u_{2c}]t \qquad (15)$$
$$+ \int_0^t f(z_1(\tau), z_2(\tau))d\tau$$

The term $\int_0^t f(z_1(\tau), z_2(\tau))d\tau$ satisfies,

$$\int_0^t f(z_1(\tau), z_2(\tau))d\tau \le \kappa_1 \int_0^t (|z_1(0)| + |z_2(0)|)d\tau$$
$$+ \kappa_1 \int_0^t (|u_{1c}| + |u_{2c}|)\tau d\tau \qquad (16)$$
$$= c_1 t + c_2 t^2$$

where $c_1 = \kappa_1(|z_1(0)| + |z_2(0)|), c_2 = 0.5\kappa_1(|u_{1c}| + |u_{2c}|)$.

Let $c_0 = |z_3(0)|, c_3 = \kappa_1(|z_1(0)| + |z_2(0)|) + [z_2(0)u_{1c} - z_1(0)u_{2c}]$, the estimation for $z_3(t)$ can be given by,

$$z_3(t) \le c_0 + c_3 t + c_2 t^2 \qquad (17)$$

It hence follows that,

$$z_1^2(t) + z_2^2(t) - |z_3(t)| \ge (z_1(0) + u_{1c}t)^2 + (z_2(0) + u_{2c}t)^2$$
$$- c_0 - c_3 t - c_2 t^2 \qquad (18)$$
$$= b_2 t^2 + b_1 t + b_0$$

with

$$b_2 \Box u_{1c}^2 + u_{2c}^2 - c_2,$$
$$b_1 \Box 2z_1(0)u_{1c} + 2z_2(0)u_{2c} - c_3 \qquad (19)$$
$$b_0 \Box z_1^2(0) + z_2^2(0) - c_0 < 0$$

Observing (19), if u_{1c}, u_{2c} are chosen satisfying $u_{1c}^2 + u_{2c}^2 > c_2$, then we conclude from $b_2 > 0$ and $b_0 < 0$ that there exists a finite time instant $t_1 > 0$ so that,

$$z_1^2(t_1) + z_2^2(t_1) - |z_3(t_1)| = 0$$
$$z_1^2(t) + z_2^2(t) - |z_3(t)| > 0, \forall t > t_1 \qquad (20)$$

Such a t_1 can be estimated as,

$$0 < t_1 \le t_2 \Box \frac{-b_1 + \sqrt{b_1^2 - 4b_2 b_0}}{2b_2} \qquad (21)$$

The derivations above are gathered together in the following lemma.

LEMMA 2. *Given the constant control inputs $[u_{1c}, u_{2c}]$ in (14) satisfying $u_{1c}^2 + u_{2c}^2 > c_2$ and $Z(0) \in \Box^3 - C$, there is a finite time instant t_1 given by (21) so that $Z(t_1) \in C$.*
Proof. The proof is direct via following the derivations (14)-(21), and hence omitted. □

According to Lemmas 1-2, the discontinuous control law can be constructed as

$$u_1 = \begin{cases} -k_1 \tanh z_1 - \frac{z_2}{z_1^2 + z_2^2}(f + g + k_2 z_3), & \text{if } Z \in C \\ u_{1c}, & \text{if } Z \in \Box^3 - C \end{cases}$$
$$u_2 = \begin{cases} -k_1 \tanh z_2 + \frac{z_1}{z_1^2 + z_2^2}(f + g + k_2 z_3), & \text{if } Z \in C \\ u_{2c}, & \text{if } Z \in \Box^3 - C \end{cases} \qquad (22)$$

THEOREM 1. *Given the discontinuous control law (22), the system (1) is globally asymptotically stable if $k_2 > 2k_1 > 0$ and $u_{1c}^2 + u_{2c}^2 > c_2$.*
Proof. According to Lemma 1, the states would converge to zero asymptotically in the invariant set C by the control law (22). Moreover, the states would be driven moving into the set C in finite time if they are initialized outside the invariant set. Thus, the discontinuous scheme (22) would always ensure the asymptotic convergence to zero of the states. The claims in the theorem are established immediately. □

Remark 2. *Note that the time instant t_1 is not necessarily known for control switch in (22). We introduce the estimation on t_1 to prove that the state trajectory would move into the invariant set C in finite time.*

Remark 3. *Due to the invariant property of C, the switch of the control law (22) would only occur once at most.*

2.3. Application to the tracking control of a unicycle robot

This subsection illustrates the application of (22) on the trajectory tracking of a unicycle robot. The kinematic model of a unicycle robot can be given by

$$\dot{x} = v\cos\theta, \dot{y} = v\sin\theta, \dot{\theta} = \omega \qquad (23)$$

where $[x, y]^T$ denotes the Cartesian position, θ is the orientation angle, u represents linear velocity and ω stands for the angular velocity. Let $[x_r, y_r, \theta_r]^T$ denote the reference trajectory generated by

$$\dot{x}_r = v_r \cos\theta_r, \dot{y}_r = v_r \sin\theta_r, \dot{\theta}_r = \omega_r \qquad (24)$$

where the reference velocities v_r and ω_r are bounded. Define the following tracking errors,

$$z_1 = (x - x_r)\cos\theta + (y - y_r)\sin\theta$$
$$z_2 = \theta - \theta_r \qquad (25)$$
$$z_3 = -2(x - x_r)\sin\theta + 2(y - y_r)\cos\theta + z_1 z_2$$

The time derivative of (25) can be calculated as

$$\dot{z}_1 = v + 0.5\omega(z_3 - z_1 z_2) - v_r \cos z_2$$
$$\dot{z}_2 = \omega - \omega_r \qquad (26)$$
$$\dot{z}_3 = z_2 \dot{z}_1 - z_1 \dot{z}_2 + 2(v_r \sin z_2 - \omega_r z_1)$$

Obviously, the error dynamics (26) features the same structure as (1) if we define

$$u_1 = v + 0.5\omega(z_3 - z_1 z_2) - v_r \cos z_2$$
$$u_2 = \omega - \omega_r$$
$$f = 2(v_r \sin z_2 - \omega_r z_1) \qquad (27)$$
$$\kappa_1 = 2\sqrt{v_r^2 + \omega_r^2}$$

Therefore, applying the control law (22) on (26) would steer the error (25) to zero globally asymptotically. As the state transformation (25) is globally invertible, the original tracking errors $x - x_r$, $y - y_r$ and $\theta - \theta_r$ would converge to zero asymptotically from any initial states.

3. Numerical Simulations

This section validates the proposed control law (22) by numerical simulations. To this end, we initialize the nonholonomic integrator (1) by $Z(0) = [-1, 1, -4]^T$ and set the control coefficients as $k_1 = 0.25, k_2 = 0.55$. The drift term is $f = 0.2(z_1 + z_2)$. The constant control inputs are $u_{1c} = 0.2, u_{2c} = 0.2$. We depict the results in Figs 2-3.
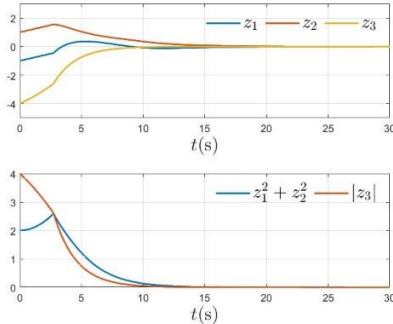
Fig. 2. The state trajectories.

As can be seen from Fig. 2. the initial states are outside the invariant set, and the control inputs are constant in this phase. The control is automatically switched into state feedback control law as soon as the switching condition $z_1^2(t) + z_2^2(t) > |z_3(t)|$ is satisfied, which can be drawn from the turnings in Fig. 2. Then, the states are constrained in the invariant set and converge to zero asymptotically. In addition, the control inputs are always bounded, which can be concluded from Fig. 3.
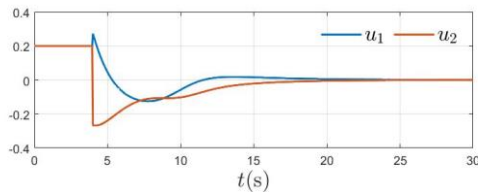
Fig. 3. The control inputs.

To validate the control law on solving the trajectory tracking control problem of a unicycle robot, we set the

reference signal by $v_r = 0.2, \omega_r = 0.1$ with initial condition $x_r(0) = 0, y_r(0) = -2, \theta_r(0) = 0$. The initial pose of the unicycle is chosen as $x(0) = -2, y(0) = -2.5, \theta(0) = \pi$. Meanwhile, the control coefficients k_1, k_2, u_{1c} and u_{2c} are set the same as previous case. We depict the position trajectory and errors in Fig. 4 and Fig. 5, respectively.

As can be seen, the Fig. 4-5 illustrate the success of applying discontinuous control law (22) on the trajectory tracking control of a typical unicycle robot. The pose tracking errors are convergent to zero asymptotically.
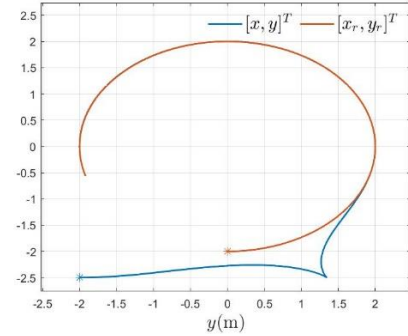
Fig. 4. The position trajectory(*:starting point).
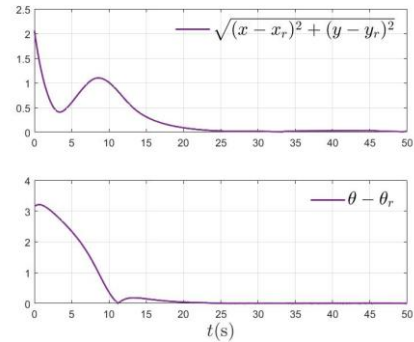
Fig. 5. The tracking errors.

4. Conclusion

This brief introduces a discontinuous global stabilizer for a class of nonholonomic integrators with drift terms. The proposed control scheme includes a state feedback controller and a constant control law, fusing with a simple criterion for control switching. It is proven in the Lyapunov sense that the system states converge to zero globally asymptotically, with the control switching occurring at most once. In the future, the authors will generalize the current control scheme to solve global stabilization problems of other nonholonomic systems.

Acknowledgements

References

1. R.W. Brockett, Asymptotic stability and feedback stabilization, Differential Geometric Control Theory. Birkhauser, 1983.
2. L. Yan and B. Ma, Adaptive practical leader-following formation control of multiple nonholonomic wheeled mobile robots, International Journal of Robust Nonlinear Control, Vol.30(17), 2020, pp.7216-7237.
3. A. Bloch and S. Drakunov, Stabilization and tracking in the nonholonomic integrator via sliding modes, Systems & Control Letters, Vol. 29(2), 1996, pp.91-99.
4. J.P. Hespanha and A.M. Stephen, Stabilization of nonholonomic integrators via logic-based switching, Automatica, Vol.35(3), 1999, pp.385-393.
5. W. Xu and W. Huo, Variable structure exponential stabilization of chained systems based on the extended nonholonomic integrator, Systems & Control Letters, Vol.41(4), 2000, pp.225-235.
6. P. Braun, L. Grüne and C.M. Kellett, Feedback design using nonsmooth control Lyapunov functions: A numerical case study for the nonholonomic integrator. 2017 IEEE 56th Annual Conference on Decision and Control (CDC); 2017: IEEE.
7. L. Yan, B. Ma, and Y. Jia. Universal Control for Both Rendezvous and Tracking of Multiple Nonholonomic Unicycles. IEEE Transactions on Control of Network Systems, 2023(Early Access).
8. H.K., Khalil, Nonlinear Systems (3rd Edition), Patience Hall, 2002.

Authors Introduction

Dr. Lixia Yan

He received B.S. degree from Beijing Jiaotong University, Beijing, China in 2013, and the M.S. and Ph.D degrees from Beihang University, Beijing, China in 2016 and 2021, respectively. He is currently a postdoctor with the Seventh Research Division and the Center for Information and Control at Beihang University. His research interests include nonlinear control theory, multiagent systems and embedded system applications. He is a member of IEEE.

Prof. Yingmin Jia

He received B.S. degree from Shandong University, Jinan, China, in 1982, and the M.S. and Ph.D. degrees from Beihang University, Beijing, China, in 1990 and 1993, respectively.
He is currently a professor with the Seventh Research Division and the director of the Center for Information and Control at Beihang University. His research interests include robust control, robotic systems, spacecraft coordination and on-orbit servicing. He is a senior member of IEEE.

Frequency Dependence Performance Limit of Vibration Absorbers

Jiqiang Wang

*Zhejiang Provincial Engineering Research Centre for Special Aircrafts, Ningbo Institute of Materials Technology &
Engineering, Chinese Academy of Sciences, Ningbo, China
Email: wangjiqiang@nimte.ac.cn*

Jingjing Fang

*Zhejiang Development & Planning Institute, Zhangzhou, China
Email: fjjzdpri@foxmail.com*

Weicun Zhang

*School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing, China
Email: weicunzhang@ustb.edu.cn*

Xinmin Chen

*Zhejiang Provincial Engineering Research Centre for Special Aircrafts, Ningbo Institute of Materials Technology &
Engineering, Chinese Academy of Sciences, Ningbo, China
Email: chenxinmin@nimte.ac.cn*

Abstract

Optimal design of vibration absorbers has been extensively investigated. Most of the design methods are approached by optimizing certain performance indices, resulting in a set of optimal parameters that are independent of exogenous forcing frequencies. In practical designs, however, it is often desirable to know the performance limits over a frequency band of interest. This problem is tackled in the present paper where both lower and upper bounds are obtained. Numerical examples are given to validate the corresponding designs.

Keywords: Vibration absorbers, optimal design, performance limit

1. Introduction

Tuned mass dampers (TMD) are widely used for vibration attenuation either at a frequency or over a frequency range. Many configurations have been proposed [1], [2], [3] while one typical TMD consists of a secondary mass-damper-spring system, attached to a vibrating primary system. The secondary system is also called an absorber whose parameters are to be designed. Besides the classical "equal height" methods, many approaches have been proposed for optimal design of TMD parameters. Most of the developments are preceded with formulating optimization problems by optimizing a properly chosen performance index. For example, H2/H∞ forms of performance indices can be optimized where even analytical solution can be found for optimal parameter selections [4], [5]. A comparison between different optimization criteria is given in [6], while more elaborated examples can be found in [7]. As a result of the optimization, optimal parameters can be obtained which is usually represented as optimal damping and frequency ratios as a function of the mass ratio.

In review of the available results in the literature, it is seen that the optimization of performance indices leads to the results that only optimize the pre-designated indices, yet performance limits, particularly the frequency-dependence performance bounding information cannot be obtained. This frequency-dependence performance limit of vibration absorbers is considered in the paper. It is organized as follows: section 2 formulates the problem to be considered. Section 3 proceeds to develop the lower and upper bounds. Numerical examples are also provided before a conclusion in section 4.

2. Problem Formulation

The model can be represented as:

$$m_1 \ddot{x}_1 + (c_1 + c_2)\dot{x}_1 - c_2 \dot{x}_2 + (k_1 + k_2)x_1 - k_2 x_2 = d$$
$$m_2 \ddot{x}_2 - c_2 \dot{x}_1 + c_2 \dot{x}_2 - k_2 x_1 + k_2 x_2 = 0 \qquad (1)$$
$$F = c_1 \dot{x}_1 + k_1 x_1$$

where the dependence on time has been omitted for easy reference. The frequency response property for the transmission force can then be written down:

$$\frac{F(j\omega)}{D(j\omega)} = \frac{(k_1 + jc_1\omega)(-m_2\omega^2 + k_2 + jc_2\omega)}{[-m_1\omega^2 + k_1 + k_2 + j(c_1+c_2)\omega](-m_2\omega^2 + k_2 + jc_2\omega) - (k_2 + jc_2\omega)^2}$$
(2)

where $F(j\omega)$ and $D(j\omega)$ are the Fourier transforms of $F(t)$ and $d(t)$, respectively.

The aim of the TMD design is to design the absorber parameters m_2, c_2, and k_2 in such a way, so that certain appropriate performance indices such as energy/magnitude defined by H2/H∞ norms are optimized. Although the optimization "routines" exist for obtaining feasible solutions to the corresponding optimization problems, it is still of great significance and interest to seek the boundary or limit of performance. As the limits will dictate the achievable performance, on the one hand, *they are not to be conquered by any form of performance index to be optimized; on the other hand, they will actually provide guidance to the designers* if a choice of performance index is suitable by knowing its "distance" to the limits. Henceforth, performance limits should be treated as a benchmarking that any design through any optimization method with any performance index must compare with. These issues are treated in the following sections.

3. Performance Limit: Lower & Upper Bounds

From equation (1), it is known that the following frequency responses relationship holds:

$$F(j\omega) = (k_1 + j\omega c_1)X_1(j\omega)$$
(3)

For any particular configuration, the primary system parameters c_1 and k_1 are known, hence optimizing $F(j\omega)/D(j\omega)$ is equivalent to optimizing $X_1(j\omega)/D(j\omega)$. From equation (2), it is known:

$$\frac{X_1(j\omega)}{D(j\omega)} = \frac{(-m_2\omega^2 + k_2 + jc_2\omega)}{[-m_1\omega^2 + k_1 + k_2 + j(c_1+c_2)\omega](-m_2\omega^2 + k_2 + jc_2\omega) - (k_2 + jc_2\omega)^2}$$
(4)

Thus the objective of TMD design can be re-stated to reduce the transmission magnitude $|X_1(j\omega)/D(j\omega)|$ through the optimal selection of the absorber parameters m_2, c_2, and k_2. In the following, the performance bounds for attenuation of the magnitude of $X_1(j\omega)/D(j\omega)$ through the to-be-designed parameters of m_2, c_2, and k_2 will be sought. This is preceded by boldly stating the results.

Theorem 1 (Lower Bound): The performance of $|X_1(j\omega)/D(j\omega)|$ is bounded from below by:

$$\left|\frac{X_1(j\omega)}{D(j\omega)}\right| > \frac{1}{\sqrt{(k_1 - m_1\omega^2)^2 + c_1^2\omega^2} + \omega\sqrt{\frac{m_2(k_2^2 + c_2^2\omega^2)}{2}}}$$
, $\forall \omega$
(5)

The relationship is strictly "greater than" implying that the lower bound is absolute and never to be attained.

Theorem 2: The performance of $|X_1(j\omega)/D(j\omega)|$ is bounded from above by:

$$\left|\frac{X_1(j\omega)}{D(j\omega)}\right| < \frac{\sqrt{K^2 + M^2\omega^4 + C^2\omega^2}}{h(\omega)}, \quad \forall \omega$$
(6)

where: $h(\omega)$ is a positive function dependent on frequency.

From the above result, a series of observations follow:

(1) While the minimum lower bound is attained at natural frequency of the primary system $\omega_1 = \sqrt{k_1/m_1}$, the minimum upper bound is achieved at the natural frequency of the absorber $\omega_2 = \sqrt{k_2/m_2}$.

(2) As the lower bound, the upper bound is also inversely proportional to the natural frequency ω_2 —to increase the bound, ω_2 needs to be decreased!

(3) Yet one of the most important applications for upper bound is the assertion that the performance $|X_1(j\omega)/D(j\omega)|$ will always be attenuated over the frequency bands where the upper bound is less than unity. This can be developed into a very useful and powerful design methodology.

For example, assume c_1 and m_2 are unities, then the bound becomes:

$$\frac{1}{\omega + \dfrac{c_2\omega^5}{(k_2 - \omega^2)^2 + c_2^2\omega^2}}$$
(7)

Then a calculation for cubic unities of k_2, c_2, and ω with a grid of $10 \times 10 \times 10$ shows that a set of solutions exist for satisfying (7). This is shown in Figure1.
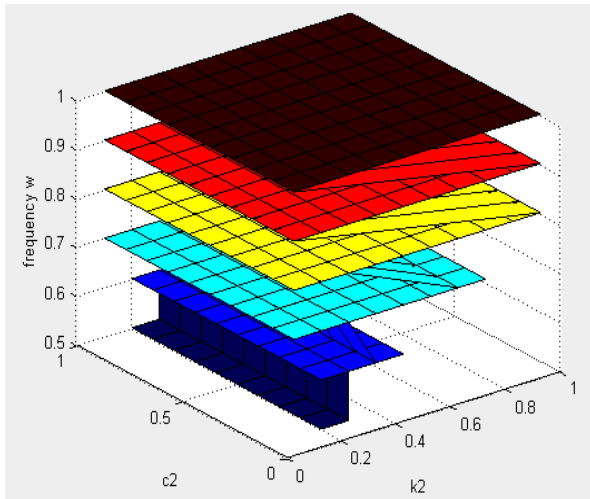
Figure 1: feasible choices of parameters k_2, c_2, and ω within a $10\times10\times10$ cubic unity grid

4. Conclusion

Frequency-dependent performance limits of tuned mass dampers have been considered. It has been demonstrated that the performance is not only bounded from below, but also bounded from above. The existence of the lower bound is very useful providing guidance upon best performance to be expected. Combining with the concept of attenuation by bounding from above, it has been shown that this can lead to design methods with such features as guaranteed performance.

Acknowledgements

References

1. S.G. Kelly, Fundamentals of Mechanical Vibrations, 2nd edition, McGraw-Hill, 2000.
2. Y. Ishida, Recent development of the passive vibration control method, Mechanical Systems & Signal Processing 2012, 29: 2-18.
3. S. Bakre, R. Jangid, Optimal multiple tuned mass dampers for base excited damped main system, International Journal of Structural Stability & Dynamics 2004, 4(4): 527-542.
4. O. Nishihara, T. Asami, Closed-form solutions to the exact optimizations of dynamic vibration absorbers, Journal of Vibration & Acoustics 2002, 124: 576-582.
5. T. Asami, O. Nishihara, A. Baz, Analytical solutions to H∞ and H2 optimization of dynamic vibration absorbers attached to damped linear systems, Journal of Vibration & Acoustics 2002, 124: 284-295.
6. G. Marano, R. Greco, B. Chiaia, A comparison between different optimization criteria for tuned mass dampers design, Journal of Sound & Vibration 2010, 329: 4880-4890.
7. G. Bekdaş, S. Nigdeli, Mass ratio factor for optimum tuned mass damper strategies, International Journal of Mechanical Sciences 2013, 71: 68-84.

Authors Introduction

Prof. Jiqiang Wang

He received his PhD degree in Control Systems in 2009 from the Unversity of Sheffield, UK. He is currently a Professor at Ningbo Institute of Materials Technology & Engineering, Chinese Academy of Sciences. He is also the deputy director of Zhejiang Provincial Engineering Research Centre for Special Aircrafts.

Ms. Jingjing Fang

She graduated with a bachelor's degree from Nanjing University of Aeronautics & Astronautics in 2013 and a master's degree from Shanghai Academy of Spaceflight Technology in 2016. She has been working on instrumentation and testing of rocket engines for many years before moving to Zhejiang Development & Planning Institute for her current career. She now focuses upon the economic and industrial analysis of cutting-edge technologies in aerospace and high-end equipment sectors.

Prof. Weicun Zhang

He received his M.S. degree in Automatic Control from Beijing Institute of Technology in 1989, and the Ph.D. degree in Control Theory and Applications from Tsinghua University in 1993. From March 1997 to May 1998, he was a visiting research fellow in Industrial and Operations Engineering Department, University of Michigan at Ann Arbor, From September 2006 to August 2007; he was a visiting professor in Department of Electrical and Computer Engineering, Seoul National University, South Korea.

Prof. Xinmin Chen

He graduated with a bachelor's degree from School of Astronautics of Beijing University of Aeronautics & Astronautics (BUAA) in 1995, and graduated from China Academy of Launch Vehicle Technology (CALT) with a master degree in 1998, and a PhD degree in 2009. In CALT he successively served as assistant engineer, engineer, senior engineer (2005) and chief engineer (2010). He has been under the supervision of Academician YU Menglun of Chinese Academy of Sciences for many years and has expertise in flight mechanics and aircraft designs.

Adaptive Concurrent Learning Algorithm Based on Pontryagin's Maximum Principle for Nonlinear System Optimal Tracking Control with State Inequality Constraints

Yuqi Zhang

Beijing University of Posts and Telecommunications, School of Artificial Intelligence, Beijing 100876, China

Bin Zhang

Beijing University of Posts and Telecommunications, School of Artificial Intelligence, Beijing 100876, China
Email: zhangbinzdh@bupt.edu.cn

Abstract

In this paper, an adaptive iterative learning algorithm is proposed to solve the optimal tracking control problem (OTCP). Unlike the existing method, we select the finite-time horizon cost function to measure tracking performance. Our method doesn't require an accurate system dynamic and the concurrent learning (CL) technique is utilized to learn the system identification model. Based on identification model, we present our concurrent iterative learning algorithm under the Pontryagin's framework to learn the solution for OTCP. The proposed algorithm overcomes the limitation of Adaptive Dynamic Programming (ADP) methods when dealing with time-varying systems or situations involving state inequality constraints. The algorithm's effectiveness is demonstrated through a numerical simulation.

Keywords: Optimal tracking control, Concurrent learning, State inequality constraints, Adaptive iterative algorithm

1. Introduction

Optimal tracking control [1], [2] is a comprehensive problem to find an optimal control input to minimize the tracking cost function to ensure that the system follows the prescribed trajectory. Due to the complexity of the nonlinear systems, it is difficult to obtain analytical solutions for OTCP. This paper aims to design an iterative algorithm to realize the tracking control without the accurate system information.

Many researchers apply Reinforcement learning (RL) technique [3], [4], [5], [6] to solve the above problems. The advantage of reinforcement learning is that it does not require a precise prior knowledge of system dynamics. Most of the methods use the ADP technique to solve the tracking Hamilton–Jacobi–Bellman (HJB) equation. Approximation methods like neural networks (NNs) are widely used in ADP literature for value function approximation. Although ADP is effective, it still has inherent technical obstacles. The HJB equation reduced to the form of an ODE under an infinite-horizon cost function and an affine nonlinear system. It is easier to learn solutions compared to PDE.

Another important problem for the OTCP is the is the existence of constraints. For control input constraints, a nonquadratic performance function [7], [8] is used in the optimal regulation problem. As for the state constraints problem, the explicit expression between optimal control and value function no longer holds in affine nonlinear systems. Therefore, within the ADP framework, the problem of state constraints for OTCP remains unresolved.

This paper aims to develop a new adaptive iterative algorithm to solve the OTCP with finite-horizon cost function and state constraints. CL technique [9] is used to learn the system identification model. Based on Pontryagin's framework and the identification model, we design a new adaptive iterative method to learn the optimal control input which minimize the tacking cost function in prescribed time interval without the exact system dynamic. Moreover, the system trajectory not only tracks the predetermined trajectory, but also satisfies the state constraints.

2. Problem statement

We consider the following nonlinear dynamic system:

$$\dot{x} = f(x, u, t), \quad x(t_0) = x_0 \quad (1)$$

where $x \in \mathbb{R}^n$ is the system state with initial state x_0, and $u \in \mathbb{R}^m$ is the control input. The unknown map $f: \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}^n$ is Lipschitz continuous. It is assumed that the state must be limited to satisfy the following inequality constraints:

$$s(x, t) \leq 0 \quad (2)$$

where $s = (s_1, s_2, \ldots, s_z)^T \in \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^z$ is a z-dimensional column vector function. Let $x_d(t) \in \mathbb{R}^n$ is a prescribed trajectory with initial state x_{d0}, the dynamic of tracking error and the constraints can be expressed as:

$$\begin{cases} \dot{e} = F(e, u, t) & (3) \\ S(e, t) \leq 0 & (4) \end{cases}$$

where $e = x - x_d$ is the tracking error, $F(e, u, t) = f(e + x_d, u, t) - \dot{x}_d$ and $S(e, t) = s(e + x_d, t)$. We define the following finite-horizon cost function for tracking performance:

$$J(e, u) = \int_{t_0}^{t_f} L(e, u) dt \qquad (5)$$

where $L(e, u) = e^T Q e + u^T R u$, $Q \in \mathbb{R}^{n \times n}$ and $R \in \mathbb{R}^{m \times m}$ are positive definite symmetric matrices. The optimal tracking control problem in this paper aims to find an optimal control input $u^*(t), t \in [t_0, t_f]$ to minimize the cost function (5) subject to system (3) and inequality constraints (4).

It is proved [10] that the above state constraints optimal control problem can be reduced to solve a sequence of unconstrained problem by minimizing the following Kelley-Bryson penalty cost function as $k \to \infty$:

$$J(e, u, r^k) = \int_{t_0}^{t_f} L(e, u) + r^k \sum_{j=1}^{z} h(S_j) S_j^2(e, t) \, dt \quad (6)$$

where $\lim_{k \to \infty} r^k = \infty$ and $h(\sigma) = \begin{cases} 1, \sigma > 0 \\ 0, \sigma \leq 0. \end{cases}$

The necessary condition for each optimal control problem associated with $J(e, u, r^k)$ is concluded by:

$$\begin{cases} \dot{\lambda} = -H_e - 2r^k \sum_{j=1}^{z} h(S_j) S_j S_{je} \\ u = \underset{u}{\operatorname{argmin}} H(e, u, \lambda, t) \end{cases} \qquad (7)$$

where $H(e, u, \lambda, t) = L(e, u) + \lambda^T F(e, u, t)$.

3. Concurrent learning iterative algorithm

3.1. *System identification*

First, we suppose that the unknown error dynamic (3) can be represented through a finite set of basis functions:

$$F(e, u) = \Theta^T(t) \Phi(e, u) + \varepsilon \qquad (8)$$

where $\Phi(e, u): \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^N$ is the vector of basis functions. $\Theta: \mathbb{R} \to \mathbb{R}^{N \times n}$ is the unknown time-varying weight matrix. ε is the vector of approximation errors. According to the Weierstrass approximation theorem, ε can be reduced to zero with arbitrarily high precision as $N \to \infty$. We define the estimate of weight matrix and tracking error as $\widehat{\Theta}$ and $\hat{e}$, respectively.

For error dynamic system (3), we define the historical data stack as $\left\{ \mathbb{E}_j(t), \mathbb{U}_j(t), \dot{\mathbb{E}}_j(t), \ t \in [t_0, t_f] \right\}_{j=1}^{W}$, which stores the tracking error, the control input, and the derivative of the tracking error. We have the following assumption for history trajectory stack:

Assumption 1. Let $\Pi^t = \sum_{j=1}^{W} \Phi^T(\mathbb{E}_j^t, \mathbb{U}_j^t) \Phi(\mathbb{E}_j^t, \mathbb{U}_j^t)$. It is assumed that there exist positive constants $\overline{\mu} > \underline{\mu} > 0$, for $\forall t[t_0, t_f]$, we have $\underline{\mu}I < \Pi^t < \overline{\mu}I$.

We design the following adaptive updating law of the i-th estimates for tracking error $\hat{e}^i$ and weight matrix $\widehat{\Theta}^i$:

$$\begin{cases} \dot{\hat{e}}^i = (\widehat{\Theta}^i(t))^T \Phi(e, u) + \omega \tilde{e}^i, \ \hat{e}^i(t_0) = x_0 - x_{d0} \ (9) \\ \widehat{\Theta}^i(t) = \widehat{\Theta}^{i-1}(t) + (I - \mu \Pi^t)^2 \Phi(e, u)(\tilde{e}^i)^T \\ \quad + \mu \sum_{j=1}^{W} \Phi(\mathbb{E}_j^t, \mathbb{U}_j^t) \left(\dot{\mathbb{E}}_j^t - (\widehat{\Theta}^{i-1}(t))^T (\Phi(\mathbb{E}_j^t, \mathbb{U}_j^t)) \right)^T \\ = \widehat{\Theta}^{i-1}(t) + (I - \mu \Pi^t)^2 \Phi(e, u)(\tilde{e}^i)^T + \mu \Pi^t \widetilde{\Theta}^{i-1}(t) \\ \hspace{8cm} (10) \end{cases}$$

where $0 < \mu < \frac{1}{\overline{\mu}}$ is a positive constant. $\tilde{e}^i = e - \hat{e}^i$ and $\widetilde{\Theta}^i(t) = \Theta(t) - \widehat{\Theta}^i(t)$ are i-th estimation errors of the tracking error and the weight matrix, respectively. Under Assumption 1, the adaptive updating law (10) guarantees that the weight matrix $\widehat{\Theta}^i$ converges to its true values.

3.2. *Iterative learning algorithm*

In this subsection, we will provide our iterative learning algorithm to solve the optimal tracking problem. To start with, we define:

$$\begin{cases} P(e, u, r, t) = L(e, u) + r^k \sum_{j=1}^{z} h(S_j) S_j^2(e, t) \\ Q(e, u, r, t, \lambda, \Theta) = P(e, u, r, t) + \lambda^T \Theta^T(\text{t}) \Phi(e, u) \end{cases} \quad (11)$$

Algorithm 1. Iterative learning algorithm for optimal tracking problem

Step 1: Initialize parameters $r^0 > 0$, $a^0 > 0$, $b^0 > 0$, $\widehat{\Theta}^0(t)$, $t \in [t_0, t_f]$, convergence error $\epsilon_1 > 0$, $\epsilon_2 > 0$, $\epsilon_3 > 0$. Select the initial control input $\text{u}^0(t)$, $t \in [t_0, t_f]$. Update the initial estimation of tracking error by solving:

$$\dot{e}^0 = (\widehat{\Theta}^0(t))^T \Phi(e, u) + \omega \tilde{e}^0, \quad \hat{e}^0(t_0) = x_0 - x_{d0} \quad (12)$$

Calculate the initial tracking cost function:

$$J(e^0, u^0, r^0) = \int_{t_0}^{t_f} P(e^0, u^0, r^0, t)dt \quad (13)$$

Let $i = 0$ and $k = 0$.

Step 2: Calculate u^{i+1} and λ^i by following equation:

$$\begin{cases} \dot{\lambda}^i = -Q_e(e^i, u^{i+1}, r^k, t, \lambda^i, \widehat{\Theta}^i), \quad \lambda^i(t_f) = 0 \\ u^{i+1} = \underset{u}{argmin} \, Q(e^i, u, r^k, t, \lambda^i, \widehat{\Theta}^i) \\ \qquad + a^i \|u - u^i\|_1 + \frac{1}{2} b^i \|u - u^i\|_2^2 \end{cases} \quad (14)$$

Step 3: Calculate $\hat{e}^i$ and $\widehat{\Theta}^i$ by following equation:

$$\begin{cases} \dot{\hat{e}}^i = (\widehat{\Theta}^i(t))^T \Phi(e, u) + \omega \tilde{e}^i, \quad \hat{e}^i(t_0) = x_0 - x_{d0} \\ \widehat{\Theta}^i(t) = \widehat{\Theta}^{i-1}(t) + (I - \mu \Pi^t)^2 \Phi(e, u)(\tilde{e}^i)^T \\ \qquad + \mu \Pi^t \widetilde{\Theta}^{i-1}(t) \end{cases} \quad (15)$$

Step 4: Calculate the tracking cost function:

$$J(e^{i+1}, u^{i+1}, r^k) = \int_{t_0}^{t_f} P(e^{i+1}, u^{i+1}, r^k, t)dt \quad (16)$$

If $J(e^{i+1}, u^{i+1}, r^k) > J(e^i, u^i, r^k)$, let $a^i \leftarrow a^i + d_a$ and $b^i \leftarrow b^i + d_b$, where d_a and d_b are positive step size, then go to Step 2. Else go to Step 5.

Step 5: If $\|u^{i+1} - u^i\| \leq \epsilon_1$, go to Step 6. Else let $i \leftarrow i + 1$, $a^i \leftarrow a^0$, $b^i \leftarrow b^0$, go to Step 2.

Step 6: If $\|\widehat{\Theta}^{i+1} - \widehat{\Theta}^i\| \leq \epsilon_2$, go to Step 7. Else let $i \leftarrow i + 1$, $a^i \leftarrow a^0$, $b^i \leftarrow b^0$, go to Step 2.

Step 7: If $S(e^{i+1}, t) \leq \epsilon_3$, go to Step 8. Else let $i \leftarrow i + 1$, $a^i \leftarrow a^0$, $b^i \leftarrow b^0$, $r^k \leftarrow r^k + d_r$ where d_r is the positive step size, then go to Step 2.

4. Simulation

In this section, we provide a nonlinear dual motor servo system with backlash to validate the efficiency of our algorithm. The system can be modeled as:

$$\begin{cases} J_i \ddot{\theta}_i + b_i \dot{\theta}_i = u_i - \tau_i \\ J_m \dot{\theta}_m + b_m \dot{\theta}_m = \sum_{i=1}^{2} \tau_i \end{cases}$$

where θ_i and θ_m represent the angles of motors and the load, J_i and J_m are the moment of inertia for the motor and the load, b_i and b_m are the resistance coefficients of the motor and load, respectively. u_i is the input torque of the motor, τ_i is the torque transmitted when the driving motor and load come into contact, which can be expressed as the following dead zone function:

$$\tau_i = \begin{cases} k(z_i + \alpha), & z_i \leq \alpha \\ 0, & |z_i| < \alpha \\ k(z_i - \alpha), & z_i \geq \alpha \end{cases}$$

where k is the torque torsion coefficient, $z_i = \theta_i - \gamma \theta_m$ is the angular error between the motor and the load, γ is transmission ratio, 2α is the backlash width. Due to the non differentiability of the dead zone function, it is generally approximated by the following smooth continuous differentiable functions:

$$\tau_i = k\left(z_i - \alpha\left(\frac{2}{1 + e^{-rz_i}} - 1\right)\right) = k(z_i - \alpha g_i)$$

Assuming two motors have the same parameters, the transmission ratio γ is selected as 1, the parameter r in the approximation function is selected as 10, and the parameters of the system are selected as Table 1.

Table 1. The system parameters.

Parameter	Value	Units
J_m	0.185	$kg \cdot m^2$
b_m	1.2	$N \cdot m \cdot s/rad$
J_i	0.028	$kg \cdot m^2$
b_i	1.3	$N \cdot m \cdot s/rad$
k	56	$N \cdot m/rad$
α	0.2	rad

Let $x = (\theta_m, \dot{\theta}_m, \theta_1, \dot{\theta}_1, \theta_2, \dot{\theta}_2)^T$. The target is selected as $x_d(t) = [1\ 0\ 1\ 0\ 1\ 0]^T$, we choose cost function:

$$J(e, u) = \int_0^{20} \frac{1}{2}\left(\sum_{i=1}^{6} e_i^2 + \sum_{j=1}^{2} u_j^2\right)dt$$

the state constraints are assumed to be $|x_2| \leq 0.4 rad/s$, $|x_4| \leq 0.4 rad/s$, $|x_6| \leq 0.4 rad/s$, which can be rewritten as: $s_1 = e_2 - 0.4$, $s_2 = -e_2 - 0.4$, $s_3 = e_4 - 0.4$, $s_4 = -e_4 - 0.4$, $s_5 = e_6 - 0.4$, $s_6 = -e_6 - 0.4$. The basis function is selected as $\Phi(e, u) = (e_1, e_2, e_3, e_4, e_5, e_6, g_1(e_1, e_3), g_2(e_1, e_5), u_1, u_2)^T$. The initial state is $x_0 = (1.5\ 0\ 1.6\ -0.1\ 1.5\ 0.1)^T$. The initial parameters are set to be $r_0 = c_0 = d_0 = 1$, $\widehat{\Theta}_0^0 = rand(6,10)^T$, $\epsilon_1 = \epsilon_2 = \epsilon_3 = 0.001$.

Yuqi Zhang, Bin Zhang

Fig. 1 and Fig. 2 are visual representations of the results. We can see that compared with the initial trajectory, the angles of load and motor have stabilized around the predetermined values.
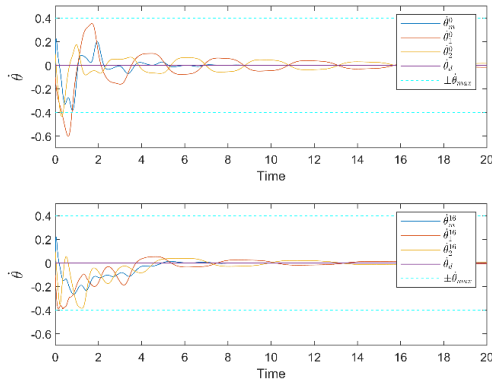
Fig. 1. Comparisons of the tracking angles.
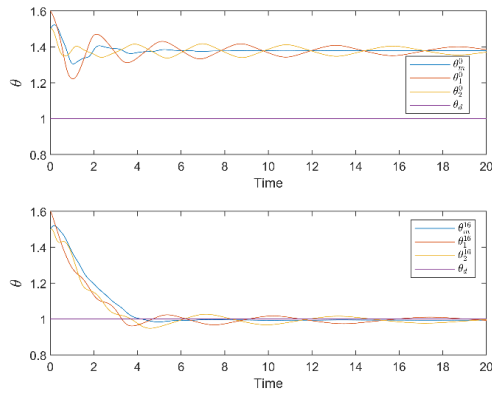
Fig. 2. Comparisons of the tracking angular velocities.

5. Conclusion

In this paper, we present a novel concurrent iterative learning algorithm. The tracking cost function is selected as a finite-horizon form. Our method does not require precise dynamic information and can handle with the time-varying systems. Moreover, state constraints are considered in the process of solving OTCP.

Acknowledgements

References

1. Firdaus E Udwadia. Optimal tracking control of nonlinear dynamical systems. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 464(2097):2341–2363, 2008.
2. Mohamed Boukattaya, Mohamed Jallouli, and Tarak Damak. On trajectory tracking control for nonholonomic mobile manipulators with dynamic uncertainties and external torque disturbances. Robotics and autonomous systems, 60(12):1640–1647, 2012.
3. Hamidreza Modares and Frank L Lewis. Optimal tracking control of nonlinear partially-unknown constrained-input systems using integral reinforcement learning. Automatica, 50(7):1780–1792, 2014.
4. Ruizhuo Song, Frank L Lewis, Qinglai Wei, and Huaguang Zhang. Off policy actor-critic structure for optimal control of unknown systems with disturbances. IEEE transactions on cybernetics, 46(5):1041–1050, 2015.
5. Yuanheng Zhu, Dongbin Zhao, and Xiangjun Li. Using reinforcement learning techniques to solve continuous-time non-linear optimal tracking problem without system dynamics. IET Control Theory & Applications, 10(12):1339–1347, 2016.
6. Fayez El-Sousy, Mahmoud M Amin, and Ahmed Al-Durra. Adaptive optimal tracking control via actor-critic-identifier based adaptive dynamic programming for permanent-magnet synchronous motor drive system. IEEE Transactions on Industry Applications, 57(6):6577–6591, 2021.
7. Murad Abu-Khalaf and Frank L Lewis. Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network hjb approach. Automatica, 41(5):779–791, 2005.
8. Bahare Kiumarsi, Frank L Lewis, Hamidreza Modares, Ali Karimpour, and Mohammad-Bagher Naghibi-Sistani. Reinforcement q-learning for optimal tracking control of linear discrete-time systems with unknown dynamics. Automatica, 50(4):1167–1175, 2014.
9. Girish V Chowdhary and Eric N Johnson. Theory and flight-test validation of a concurrent-learning adaptive controller. Journal of Guidance, Control, and Dynamics, 34(2):592–607, 2011.
10. Milind M Lele and David H Jacobson. A proof of the convergence of the kelley-bryson penalty function technique for state-constrained control problems. Journal of Mathematical Analysis and Applications, 26(1):163–169, 1969.

Authors Introduction

Mr. Yuqi Zhang

He received the B.S. degree in Automation from Beijing University of Posts and Telecommunications, Beijing, China, in 2020. He is currently working toward the M.D. degree in Control Science and Engineering. His research interests include differential games and reinforcement learning.

Dr. Bin Zhang

He received his B.S. and Ph.D. degrees both in control theory and applications from Beihang University, Beijing, China, in 2010 and 2016, respectively. He is currently an Associate Professor with the School of Artificial Intelligence at Beijing University of Posts and Telecommunications. His research interests include reinforcement learning, multi-agent systems, and intelligent control.

Privacy preserving Mean-square consensus for discrete-time heterogeneous multi-agent systems with Communication Noises

Tongqing Yang

School of Mathematics and Statistics, Beijing Technology and Business University, 11 Fucheng Road, Haidian District, Beijing, 100048, China

Lipo Mo

School of Computer and Artificial Intelligence, Beijing Technology and Business University, 11 Fucheng Road, Haidian District, Beijing, 100048, China

Yingmin Jia

*The seventh Division, Beihang university, 37 Xueyuan Road, Haidian District, Beijing, 100191, China
Email: yangtongqing08@126.com, molipo@th.btbu.edu.cn, ymjia@buaa.edu.cn.
www.btbu.edu.cn*

Abstract

This paper investigates privacy-preserving mean-square consensus in distributed heterogeneous multi-agent systems with communication noise on fixed undirected graphs. To mitigate the impact of communication noise, we introduce a stochastic approximation step rule in the control protocol. Utilizing graph theory, stochastic analysis, and Lyapunov theory, consensus conditions are derived. Subsequently, a cryptographic cryptosystem encrypts the designed protocol, safeguarding against eavesdropping and information privacy loss between agents during consensus. Numerical simulations confirm the efficacy of the proposed consensus protocol and privacy protection algorithm.

Keywords: Distributed, Heterogeneous multi-agent systems (HMAS), Consensus, Privacy preserving

1. Introduction

In recent years, research on multi-agent systems has profoundly impacted the development of engineering technology [1]. Consensus control stands out as a key issue in multi-agent system studies, aiming for the convergence of state and output values of each agent through a control protocol [2]. Extensive research results exist for consensus control in multi-agent systems comprising first or second-order agents [3]. However, the focus is expanding to include Heterogeneous Multi-Agent Systems (HMAS) consisting of both first-order and second-order agents. HMAS demonstrates superior load and task configuration capabilities, making it suitable for more complex environments [4].

In the above research, the consensus control protocol relies on the interaction of state information between adjacent nodes to achieve consensus. For instance, a group of agents aiming to converge at a specific location may wish to keep their initial locations confidential. In recent years, various privacy protection schemes have been explored. A commonly used method is differential privacy, where random noise is introduced to the interaction state to obscure the true state [5]. However, this approach is susceptible to eavesdropping, as the added noise compromises accuracy [6], [7]. Another approach involves a state decomposition mechanism introduced in literature. To address privacy concerns without sacrificing accuracy, the study delves into the homomorphic encryption scheme and extends the average consensus of first-order agent systems to second-order agent systems [7], [8]. Nevertheless, practical networks often operate in uncertain communication environments with random perturbations. Consequently, instead of assuming a deterministic form, literature considers the mean square consensus of communication noise [9], [10]. However, it overlooks privacy protection in the presence of communication noise, leaving the challenge of achieving privacy-preserving consensus for Heterogeneous Multi-Agent Systems (HMAS) with communication noises unresolved. This paper presents a distributed consensus algorithm for HMAS based on cryptography. The main contributions of this paper are as follows:

Tongqing Yang, Lipo Mo, Yingmin Jia.

1) We propose a privacy consensus algorithm for HMAS with communication noises based on partial homomorphic cryptography;

2) The privacy protection for the initial state of the agent is analyzed in detail. In particular, compared with the existing approach in [7], [8], our algorithm can achieve higher privacy when there exists an agent connected with only one neighbor;

3) The effectiveness of our method is verified by simulation.

2. Preliminaries and problem description

The network of HAMS consisting of n agents is represented by an undirected connected graph G. For HMAS with first-order and second-order agents, the Laplacian matrix of the graph G can be described as

$$L = \begin{bmatrix} L_s + D_{sf} & -A_{sf} \\ -A_{fs} & L_f + D_{fs} \end{bmatrix},$$

where L_s and L_f represent the Laplacian matrices of the graph comprising the second-order agents and first-order agents, respectively. A_{sf} and A_{fs} denote the adjacency matrix between the second-order agents (first-order agents) and first-order agents (second-order agents) respectively, D_{sf} and D_{fs} denote their row sums.

Paillier Cryptosystem: The Pailler crypto-system has key generation, encryption and decryption functions. The encryption operator E and the decryption operator D are usually used in the encryption and decryption process. Now, we introduce the additive homomorphism property of the Paillier cryptosystem:

$$E(z)^k = \prod_{i=1}^{k} E(z) = E(\sum_{i=1}^{k} z) = E(kz).$$

We assume that there are n agents in the HMAS, and the first group has m second-order agents, labeled $V_m = \{1, 2, \cdots, m\}$, while the rest are first-order agents, labeled $V_{n-m} = \{m+1, m+2, \cdots, n\}$. The dynamic equation of the composed HMAS is as follows:

$$\begin{cases} x_i(k+1) = x_i(k) + v_i(t)T \\ v_i(k+1) = v_i(k) + u_i(t)T, & i \in V_m, \\ x_i(k+1) = x_i(k) + u_i(t)T, & i \in V_{n-m}. \end{cases} \quad (1)$$

where $x_i(k) \in \Box$, $v_i(k) \in \Box$ and $u_i(k) \in \Box$ are the position, velocity and control input of agent i, respectively. The sample time is $0 < T < 1$. All results can obtained exist in $\Box^n$ by using the Kronecker product.

Definition 1. The HMAS consisted of (1) is said to achieve mean-square consensus if for any initial condition, we have

$$\lim_{k \to \infty} E\Box x_j(k) - x_i(k)\Box^2 = 0 \quad i, j = 1, 2, \cdots, n$$

$$\lim_{k \to \infty} E\Box v_i(k)\Box^2 = 0 \quad i, j = 1, 2, \cdots, m$$

Lemma 1. Let $\{\tau(k)\}, \{\beta(k)\}$ and $\{q(k)\}, k = 0, 1, \cdots$, be real sequences, satisfying $0 < q(k) \leq 1, \tau(k) \geq 0$, $\beta(k) \geq 0$, $\sum_{k=0}^{\infty} q(k) = \infty, \frac{\beta(k)}{q(k)} \to 0, k \to \infty$, and $\tau(k+1) \leq (1 - q(k))\tau(k) + \beta(k)$, then $\tau(k) \to 0, k \to \infty$.

3. Main contents

3.1. *Mean-square Consensus in Heterogeneous Multi-agent Systems*

We propose the following consensus protocol:

$$\begin{cases} u_i(k) = \alpha(k)\sum_{j=1}^{n} a_{ij}(x_j(k) + w_{ji}(k) - x_i(k)) - k_1 v_i(k), & i \in V_m, \\ u_i(k) = \alpha(k)\sum_{j=1}^{n} a_{ij}(x_j(k) + w_{ji}(k) - x_i(k)), & i \in V_{n-m}. \end{cases} \quad (2)$$

where $\alpha(k)$ and k_1 are the control gains, $\{w_{ji}(k), k \geq 0\}$ are communication noise sequences.

Let $y_i(k) = x_i(k) + v_i(k)$, $i \in V_m$, then derived from (1)
$$x_i(k+1) = (1-T)x_i(k) + Ty_i(k).$$

Define $\xi(k) = [x_1(k), \cdots, x_m(k), y_1(k), \cdots, y_m(k), x_{m+1}(k), \cdots, x_n(k)]$, $W(k) = [W_1^T(k), \cdots, W_m^T(k), W_1^T(k), \cdots, W_n^T(k)]^T$, $\Lambda = diag\{0, \cdots, 0, \Lambda_1, \cdots, \Lambda_n\}, \Lambda_i = [a_{i1}, a_{i2}, \cdots, a_{in}]$, $W_i(k) = [w_{1i}(k), \cdots, w_{ni}(k)]^T$. Then the HMAS (1) can be rewritten as

$$\xi(k+1) = [I_{m+n} - \alpha(k)T\Psi_1 - (1 - \alpha(k))T\Psi_2]\xi(k) + \alpha(k)T\Lambda W(k) \quad (3)$$

where

$$\Psi_1 = \begin{bmatrix} I_m & -I_m & 0 \\ (1-k_1)I_m + (L_s + D_{sf}) & (k_1-1)I_m & -A_{sf}T \\ -A_{fs}T & 0 & (L_f + D_{fs})T \end{bmatrix}$$

$$\Psi_2 = \begin{bmatrix} I_m & -I_m & 0 \\ (1-k_1)I_m & (k_1-1)I_m & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

It is clear that $\theta = [(k_1 - 1)\mathbf{1}_m^{\sharp}, \mathbf{1}_m, \mathbf{1}_{n-m}]$ and $\mathbf{1}_{n+m}$ are the left eigenvector and right eigenvector associated with zero eigenvalue of $\alpha(k)T\Psi_1 + (1 - \alpha(k))T\Psi_2$, respectively.

Assumption 1. The additive independent noise $\{w_{ji}(k), k \geq 0\}$, $E(w_{ji}(k)) = 0$, $Var(\{w_{ji}(k)\}) = \sigma_W$.

Assumption 2. Suppose that the control gains satisfy the following conditions

$$1)\sum_{k=0}^{\infty}a(k)=\infty, 0<a(k)<1, a(k+1)\le a(k), 2)\sum_{k=0}^{\infty}a^2(k)<\infty.$$

Lemma 2. If Assumptions 1-2 hold, $0<T\le\min\{\max_{\lambda_i\neq 0}$

$\frac{1}{\sqrt{\text{Re}^2(\lambda_i)+\text{Im}^2(\lambda_i)}}, \max_{\eta_i\neq 0}\frac{1}{\sqrt{\text{Re}^2(\eta_i)+\text{Im}^2(\eta_i)}}\}$ and $k_1\ge 1+\max l_{ii}$, l_{ii},

$1,2,\cdots,n$, then $0<\alpha(k)T\rho_1<1$, $0<(1-\alpha(k))T\rho_2<1$,

where $\lambda_i, i=1,2,\cdots,n+m$ and $\eta_i, i=1,2,\cdots,n+m$ are the eigenvalues of Ψ_1 and Ψ_2 , ρ_1 and ρ_2 are the spectral radius of Ψ_2 and Ψ_2 , respectively. l_{ii} is the diagonal element of the Laplacian matrix.

Theorem 1. If Assumptions 1-2 hold, $0<T\le\min\{\max_{\lambda_i\neq 0}$

$\frac{1}{\sqrt{\text{Re}^2(\lambda_i)+\text{Im}^2(\lambda_i)}}, \max_{\eta_i\neq 0}\frac{1}{\sqrt{\text{Re}^2(\eta_i)+\text{Im}^2(\eta_i)}}\}$ and $k_1\ge 1+\max l_{ii}$, l_{ii}, i

$=1,2,\cdots,n$, then the HMAS (3) reach mean-square consensus.

3.2. *Privacy protection*

In this paper, honest-but-curious and eavesdropper were considered.

Algorithm 1 Encrypted information exchange

Initialization: Each agent i initializes its state $x_i(0)$, the public key pk_i and private key sk_i are generated, and broadcasts pk_i to its neighbor(s) while keeps sk_i private. Each time a message is sent, a independent random variable $w_{j\to i}(k), \forall i,j\in\mathbb{V}_m, k=\ 0,1,\cdots$ is added.

Input: $x_i(k), k_1$; **Output:** $u_i(k)$.

1: Agent i encrypts $x_i(k)$ with its public key pk_i :
$$-x_i(k)\to\text{E}_i(-x_i(k))$$
then sends $\text{E}_i(-x_i(k))$ to each neighbor $j\in N_i$, the equivalent of sending $\text{E}_i(-x_i(k)+w_{i\to j}(k))$.

2: Each agent j encrypts $x_j(k)$, and $v_j(k)$ with its public key pk_i :
$$x_j(k)\to\text{E}_i(x_j(k))$$

3: Each agent i generates the number $a_i(k)$ and agent $j\in N_i$ generates the number $a_j(k)$.

4: Each agent $j\in N_i$ computes the encrypted differences based on property of the Paillier cryptosystem:

$\text{E}_i(x_j(k))\text{E}_i(-x_i(k)+w_{i\to j}(k))=\text{E}_i(x_j(k)-x_i(k)$

$+w_{i\to j}(k))\to(\text{E}_i(x_j(k)-x_i(k)+w_{i\to j}(k)))^{a_j(k)}$

$=\text{E}_i(a_j(k)(x_j(k)-x_i(k)+w_{i\to j}(k)))$

then sends encrypted differences back to agent i .

5: Agent i decrypts the received ciphertexts with private key sk_i and computes weighted differences:

$\text{E}_i(a_j(k)(x_j(k)-x_i(k)+w_{i\to j}(k)+w_{j\to i}(k)))$

$\to a_j(k)(x_j(k)-x_i(k)+w_{ji}(k))$

$\xrightarrow{\times a_i(k)}a_{ij}(x_j(k)-x_i(k)+w_{ji}(k))$

6: The values calculated in step 5 are taken into protocol (2) to update $x_i(k)$ and $v_i(k)$.

7: Set $k\leftarrow k+1$.

Definition 2 [7]: For a connected network of n nodes, the privacy of the initial value $x_i(0)$ of node i is preserved if an honest-but-curious adversary cannot estimate $x_i(0)$ with any accuracy.

Similar to [7], the fixed coupling weight can be decomposed, $a_{ij}=a_{ij}=a_i(k)a_j(k)>0$, where $a_i(k)$ is generated when agent i transmits information and only known by itself. Similarly, we split the noise delivered by a communication link, i.e., $w_{ji}(k)=w_{i\to j}(k)+w_{j\to i}(k)$ in (2).

Theorem 2: In a connected network of nodes with system dynamics given by (1), each agent adheres to the consensus protocol (2) and Algorithm 1 for state updates. An honest-but-curious node, Eve, capable of receiving messages from a neighboring node, Alice, cannot ascertain Alice's initial state if Alice is also connected to another legitimate node, Bob. Furthermore, Alice's privacy remains undisclosed to external eavesdroppers.

Theorem 3: In a connected network of nodes governed by system dynamics as described in (1), each agent adheres to the consensus protocol (2) and Algorithm 1 for state updates. If a node, Alice, is solely connected to the rest of the network through an honest-but-curious node (or a colluding group of such nodes) Eve, and the noise introduced by the communication link is symmetrical, Eve can asymptotically deduce Alice's initial state. Conversely, in the case of asymmetric noise, while Eve can receive messages from its neighbor Alice, it lacks the means to ascertain Alice's initial state.

4. Numerical simulation

Consider the HMAS with eight agents, where nodes 1-4 represent first-order agents with input saturation and nodes 5-8 represent second-order agents. The network of the eight agents between communication topology shown

Tongqing Yang, Lipo Mo, Yingmin Jia.

in Fig. 1 (a). We assume $k_1 = 5$, given the initial value $x(0) = [1,2,3,4,5,6,7,8]^{\sharp}, v(0) = [1,2,3,4]$. Fig. 1 (b) represents the case where the state of the agent converges to 0. Fig. 1 (c) is the comparison between the real value of the input of agent 1 and the encrypted ciphertext.

a. Network   b. Convergence   c. Control input
Fig.1. Numerical simulation picture
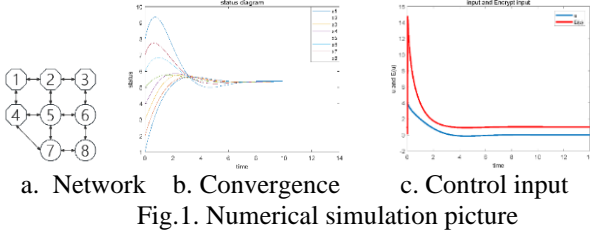
5. Conclusion

This paper explores privacy-preserving consensus in distributed heterogeneous multi-agent systems with first-order and second-order agents, addressing communication noise using graph theory, stochastic analysis, Lyapunov theory, and cryptography. It proves that the closed-loop system attains consensus. The designed protocol is encrypted to prevent eavesdropping and information privacy loss. However, there are limitations when a secure agent connects with an adversary alone. Numerical simulations validate the consensus protocol's effectiveness and the privacy protection algorithm's ability to safeguard agent privacy.

Acknowledgements

References

1. W. Yue, Y. Yang and W. Sun, "Resilient Consensus Control for Heterogeneous Multiagent Systems via Multiround Attack Detection and Isolation Algorithm," in IEEE Transactions on Industrial Informatics, doi: 10.1109/TII.2023.3327175.
2. A. Nedic, A. Ozdaglar and P. A. Parrilo, "Constrained Consensus and Optimization in Multi-Agent Networks," in IEEE Transactions on Automatic Control, vol. 55, no. 4, pp. 922-938, April 2010.
3. H. Du, S. Li, and P. Shi, "Robust consensus algorithm for second-order multi-agent systems with external disturbances," International Journal of Control, vol. 85, no. 12, pp. 1913–1928, 2012.
4. Y. Liu, H.B. Min, S.C. Wang, Z.G. Liu, S.Y. Liao, Distributed consensus of a class of networked heterogeneous multi-agent systems, J. Frankl. Inst. 351 (2014) 1700–1716.
5. X. Li, H. Yan, Z. Cheng, W. Sun and H. Li, "Protecting Regression Models With Personalized Local Differential Privacy," in IEEE Transactions on Dependable and Secure Computing, vol. 20, no. 2, pp. 960-974, 1 March-April 2023.
6. Zhang K, Li Z, and Wang Y, "Privacy-preserving dynamic average consensus via state decomposition: Case study on multi-robot formation control," Automatica, vol. 139, May. 2022.
7. M. Ruan, H. Gao and Y. Wang, "Secure and Privacy-Preserving Consensus," in IEEE Transactions on Automatic Control, vol. 64, no. 10, pp. 4035-4049, Oct. 2019.
8. W. Fang, M. Zamani, and Z. Chen, "Secure and privacy preserving consensus for second-order systems based on paillier encryption," Syst.Control Lett., vol. 148, p. 104869, 2021.
9. S. Guo, L. Mo, Y. Yu, Mean-square consensus of heterogeneous multi-agent systems with communication noises, Journal of the Franklin Institute 355 (8) (2018) 3717-3736.
10. T. Li, J.-F. Zhang, Consensus conditions of multi-agent systems with time-varying topologies and stochastic communication noises, IEEE Transactions on Automatic Control 55 (9) (2010) 2043-2057.

Authors Introduction

Mr. Tongqing Yang

He received his Bachelor degree in Engineering in 2022 from the School of Mechanical and Electrical Engineering, Guizhou Minzu University in China. He is currently a master student in Beijing Technology and Business University, China.

Dr. Lipo Mo

He received his B.S. degree in Department of Mathematics, Shihezi University, China, in 2003, and the Ph.D. degree in School of Mathematics and Systems Science from the Beihang University, Beijing, in 2010.

Yingmin Jia (Member, IEEE)

He received the B.S. degree in control theory from Shandong University, Jinan, China, in 1982, and the M.S. and Ph.D. degrees in control theory and applications from Beihang University, Beijing, China, in 1990 and 1993, respectively.

Event-Triggered Consensus Control for Nonlinear Singular Multi-Agent Systems under Directed Topology

Lin Li*, Tong Yuan

Automation Division, University of Shanghai for Science and Technology, Shanghai, China, 200093
Email: lilin0211@163.com, ytong77@163.com

Abstract

The event-triggered consensus problem for singular multi-agent systems with Lipschitz nonlinearity under directed topology is investigated in this paper. A sampled-data-based event-triggered mechanism is constructed to decide when the current data-packet should be broadcast. The objective is to design an event-triggered protocol such that the considered multi-agent system can achieve admissible consensus. Based on graph theory and singular system theory, sufficient conditions of consensus of nonlinear singular multi-agent systems are derived. Finally, a numerical example shows the effectiveness of our proposed approach.

Keywords: Singular multi-agent systems, Event-triggered scheme, Consensus control, Lipschitz nonlinearity

1. Introduction

In recent decades, the distributed coordinated control of multi-agent systems has attracted tremendous attention due to its wide adoption in various fields, such as robot formation [1], [2], [3], sensor network [4], [5], [6], neural network [7], [8], target tracking control [9] and so on. Consensus is the basis problem of the coordinated control of multi-agent systems. In practical, each agent may be equipped by the embedded microprocessor, which is with limited computing and communication capabilities. How to utilize these limited resources to operate system efficiently becomes a concerned issue. Therefore, event-triggered consensus has attracted much attention and achieved some results [10], [11], [12]. Singular systems, can provide convenient and accurate description in economic systems, power systems and aircraft modeling from practical considerations. Some literatures have studied the event-triggered for singular systems [13], [14], [15]. However, to the best of our knowledge, there are few works regarding to the event-triggered consensus for singular multi-agent systems (SMAS), which motivates the current study.

In the present study, we focus on the event-triggered consensus control for a class of SMAS with Lipschitz nonlinearity under directed topology. An even-triggered mechanism excluding Zeno behavior is proposed, and an artificial transition time-delay is introduced to design the event-trigged consensus control protocol. Sufficient conditions that can guarantee the consensus of the considered SMAS are obtained. Throughout this paper, $P > 0$ $(P \geq 0)$ means that P is positive definite (positive semi-define). I_N is the $N \times N$ identity matrix. $\otimes$ and $\|\cdot\|$ are used to represent Kronecker product and Euclidean norm of a vector or a matrix, respectively. In a symmetric matrix, * denotes the matrix entries implied by symmetry. **diag**$\{\cdots\}$ stands for a block-diagonal matrix.

2. Problem formulation and preliminaries
2.1. Graph theory

A directed graph $\mathcal{G}(\mathcal{V}, \mathcal{E}, \mathcal{A})$ is considered in this paper, in which $\mathcal{V} = \{v_1, v_2, \cdots v_N\}$ is the node set, $\mathcal{E}$ is the edge set, and $\mathcal{A} = [a_{ij}]_{N \times N}$ is the adjacency matrix, $a_{ij} > 0$ if v_i can receive the information from $v_j (i, j = 1,2,\cdots N)$, otherwise $a_{ij} = 0$. The set of neighbors of the v_i is denotes as $\mathcal{N}_i$. Let $d_i = \sum_{j=1}^{j=N} a_{ij}$ be the in-degree of node v_i and $D = \text{diag}(d_1, \cdots, d_N)$. The Laplacian matrix of $\mathcal{G}$ can be defined as $\mathcal{L} = D - \mathcal{A}$. Suppose the graph $\mathcal{G}$ has a directed spanning tree.

2.2. Problem formulation

Consider a group of N agents, and the dynamic of ith agent is described by

$$E\dot{x}_i(t) = Ax_i(t) + Bu_i(t) + f(x_i(t)), i = 1,2\cdots, N \quad (1)$$

where $x_i(t) \in \mathbb{R}^n$, $u_i(t) \in \mathbb{R}^m$ denote the state and the input, respectively. E, A and B are constant with $rank(E) = r < n$. The nonlinear function $f(x_i(t))$ satisfies $\|f(\varsigma_1) - f(\varsigma_2)\| \leq \mu \|\varsigma_1 - \varsigma_2\|$, $\forall \varsigma_1, \varsigma_1 \in \mathbb{R}^n$ with $\mu > 0$ is the Lipschitz constant.

Definition 1: The SMAS (1) is said to achieve admissible consensus if it is regular, impulse free while satisfying

$$\lim_{t \to \infty} \| x_i(t) - x_j(t) \| = 0, \forall i, j = 1, 2 \cdots N.$$

In this paper, we assume that the state of the multi-agent system (1) is periodically sampled at a constant sampling period $h > 0$. The sampling sequence is described by the set $\mathbb{S}_1 = \{0, h, 2h, \cdots, kh, \cdots\}$. The transmission event is determined by an event-triggered scheme. The transmission sequence of agent i is described by the set $\mathbb{S}_2 = \{0, t_1^i h, t_2^i h, \cdots, t_k^i h, \cdots\}$.

The event detector is described as follows

$$\begin{cases} t_{k+1}^i h = \inf\{kh|kh > t_k^i h, \psi_i(kh) \geq 0\} \\ \psi_i(kh) = \chi_i^T(t_k^i h + l_i h)\Phi\chi_i(t_k^i h + l_i h) - \\ \qquad \delta_i\lambda_i^T(t_k^i h + l_i h)\Phi\lambda_i(t_k^i h + l_i h)\} \end{cases} \quad (2)$$

where $kh = t_k^i h + l_i h$, $l_i \in \mathbb{N}$, $\delta_i > 0$ is the threshold parameter, Φ is a positive definite matrix, and

$$\begin{cases} \chi_i(t_k^i h + l_i h) = x_i(t_k^i h + l_i h) - x_i(t_k^i h) \\ \lambda_i(t_k^i h + l_i h) = \sum_{j \in \mathcal{N}_i} a_{ij}(x_i(t_k^i h) - x_j(t_{k_j'}^j h)) \end{cases}$$

with $k_j' = argmin_p\{t_k^i + l_i - t_p^j | t_k^i + l_i > t_p^j, p \in \mathbb{N}\}$.

Under the event-triggered scheme (2), $t_{k+1}^i h - t_k^i h$ denotes the transmission period of the Event generators. Let τ_k^i denote the signal transmission delay of ith agent. Suppose τ_k^i is bound, that is, $\tau_k^i \in (0, \tau_M]$, where τ_M is a positive integer. In view of the effect of signal transmission delay, the released states $x_i(t_k^i h)$ will reach the controller at the time instants $t_k^i h + \tau_k^i$. Obviously, $t_k^i h + \tau_k^i < t_{k+1}^i h + \tau_{k+1}^i (k = 1, 2 \cdots)$. Figure 1 shows the above-mentioned event-triggered transmission scheme.
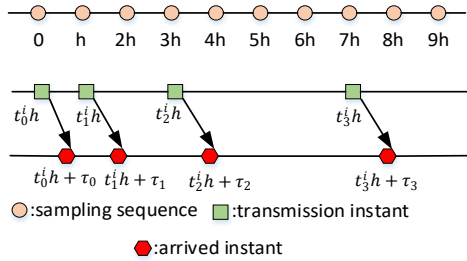
Figure 1. Example of time evolution of the sampling and transmission series of agent i

In system (1), based on the event-triggered mechanism (2), consider the following control protocol

$$u_i(t) = -K \sum_{j \in \mathcal{N}_i} a_{ij}\left(x_i(t_k^i h) - x_j\left(t_{k_j'}^j h\right)\right) \quad (3)$$

where $t \in \left[t_k^i h + \tau_k^i, t_{k+1}^i h + \tau_{k+1}^i\right)$, $i = 1, 2 \cdots N$, and K is the control gain matrix to be determined.

Define the function $\tau_i(t)$ as

$$\tau_i(t) = \begin{cases} t - t_k^i h, & t \in [t_k^i h + \tau_k^i, t_k^i h + h + \tau_M) \\ t - t_k^i h - \Delta h, & t \in [t_k^i h + \Delta h + \tau_M, t_k^i h + \Delta h + h + \tau_M) \\ t - t_k^i h - dh, & t \in [t_k^i h + dh + \tau_M, t_{k+1}^i h + \tau_{k+1}^i) \end{cases}$$

where $\tau_M = \max\{\tau_k^i\}$, Δ is a positive integer satisfying $\Delta \geq 1$. It is easy to see that

$$e_i(t_k^i h + \Delta h) = 0, \ t \in [t_k^i h + \tau_k^i, t_k^i h + h + \tau_M),$$
$$e_i(t_k^i h + \Delta h) = x_i(t_k^i h + \Delta h) - x_i(t_k^i h),$$
$$t \in [t_k^i h + \Delta h + \tau_M, t_k^i h + \Delta h + h + \tau_M)$$
$$e_i(t_k^i h + \Delta h) = x_i(t_k^i h + dh) - x_i(t_k^i h),$$
$$t \in [t_k^i h + dh + \tau_M, t_{k+1}^i h + \tau_{k+1}^i)$$

Then, we get
$$x_i(t_k^i h) = x_i(t - \tau_i(t) - e_i(t_k^i h + \Delta h))$$

Then, the event-triggered scheme (2) can be rewritten as

$$\chi_i^T(t_k^i h + l_i h)\Phi\chi_i(t_k^i h + l_i h) < \delta_i\left[\sum_{j \in \mathcal{N}_i} a_{ij}(x_i(t_k^i h) - x_j(t_{k_j'}^j h))\right]^T \Phi\left[\sum_{j \in \mathcal{N}_i} a_{ij}(x_i(t_k^i h) - x_j(t_{k_j'}^j h))\right]$$

Substituting (3) into the system (1) yields
$$(I_N \otimes E)\dot{x}(t) = (I_N \otimes A)x(t) - (\mathcal{L} \otimes BK)x(t - \tau(t)) + (\mathcal{L} \otimes BK)e(kh) + F(x(t))$$

Let $z_i(t) = x_1(t) - x_{i+1}(t)$, $\varepsilon_i(kh) = e_1(kh) - e_{i+1}(kh)$, $z(t) = [z_1^T(t), z_2^T(t), \cdots, z_{N-1}^T(t)]^T = (T_1 \otimes I_n)x(t)$, $\varepsilon(kh) = [\varepsilon_1^T(kh), \varepsilon_1^T(kh), \cdots, \varepsilon_{N-1}^T(kh)] = (T_2 \otimes I_n)e(kh)$, $T_1 = [1, -I_{N-1}] \in \mathbb{R}^{(N-1) \times N}$, $T_2 = [0, -I_{N-1}]^T \in \mathbb{R}^{N \times (N-1)}$. The following system is immediate

$$(I_{N-1} \otimes E)\dot{z}(t) = (I_{N-1} \otimes A)z(t) - (T_1 \mathcal{L} T_2 \otimes BK)z(t - \tau(t)) + (T_1 \mathcal{L} T_2 \otimes BK)\varepsilon(kh) + (T_1 \otimes I_n)F(x(t))$$

$$(4)$$

Now, the consensus problem of SMAS (1) is transformed to the admissible of system (4).

3. Main results

Theorem 1: Given h, τ_m, $\bar{\tau}_M$, and μ, the SMAS (1) achieves admissible consensus if there exist real matrices P, positive definite matrices Q, R_1, R_2, Z_1, Z_2 such that the following matrix inequalities hold

$$\bar{E}^T P \bar{E} \geq 0 \quad (5)$$

$$\Omega_1 = \begin{bmatrix} \Psi_0 + \Upsilon + \Upsilon^T & \sqrt{\tau_m}\Psi_1 & \sqrt{\tau_{Mm}}\Psi_2 & \sqrt{\tau_m}L & \sqrt{\tau_{Mm}}M \\ * & Z_1 & 0 & 0 & 0 \\ * & * & Z_2 & 0 & 0 \\ * & * & * & Z_1 & 0 \\ * & * & * & * & Z_2 \end{bmatrix} < 0 \quad (6)$$

$$\Omega_2 = \begin{bmatrix} \Psi_0 + \Upsilon + \Upsilon^T & \sqrt{\tau_m}\Psi_1 & \sqrt{\tau_{Mm}}\Psi_2 & \sqrt{\tau_m}L & \sqrt{\tau_{Mm}}N \\ * & Z_1 & 0 & 0 & 0 \\ * & * & Z_2 & 0 & 0 \\ * & * & * & Z_1 & 0 \\ * & * & * & * & Z_2 \end{bmatrix} < 0 \quad (7)$$

where

$$\varphi_0 = \bar{A}^T \mathbb{P}^T + \mathbb{P}\bar{A} + R_1 + R_2 + \mu^2 I,$$

$$\varphi_1 = -Q + \lambda(T_2^T L^T \Lambda L T_2) \otimes \Phi,$$

$$\varphi_2 = -\lambda(T_2^T L^T \Lambda L T_2) \otimes \Phi,$$

$$\varphi_3 = -(I_{N-1} \otimes \Phi) + \lambda(T_2^T L^T \Lambda L T_2) \otimes \Phi,$$

$\Upsilon = [L\bar{E} \quad 0 \quad (N-M)\bar{E} \quad (M-N)\bar{E} \quad -N\bar{E} \quad 0],$

$\Psi_1 = [Z_1\bar{A} \quad 0 \quad Z_1\bar{L} \quad 0 \quad 0 \quad -Z_1\bar{L}]^T,$

$\Psi_2 = [Z_2\bar{A} \quad 0 \quad Z_2\bar{L} \quad 0 \quad 0 \quad -Z_2\bar{L}]^T,$

$\mathbb{P} = I_{N-1} \otimes (E^T P + TS^T).$

$$\Psi_0 = \begin{bmatrix} \varphi_0 & \mathbb{P} & -\mathbb{P}\bar{L} & 0 & 0 & \mathbb{P}\bar{L} \\ * & -I & 0 & 0 & 0 & 0 \\ * & * & \varphi_1 & 0 & 0 & \varphi_2 \\ * & * & * & -R_1 & 0 & 0 \\ * & * & * & * & -R_2 & 0 \\ * & * & * & * & * & \varphi_3 \end{bmatrix}$$

Proof: Firstly, similar to the analysis in [14], it can be proved that LMIs (6) and LMIs (7) can guarantee that system (4) is regular and impulse-free.

Next, we will prove that the system (4) is asymptotically stable. Construct the Lyapunov functional

$V(t) = V_1(t) + V_2(t) + V_3(t)$

$V_1(t) = z^T(t)\bar{E}^T P\bar{E}z(t) + (\bar{\tau}_M - \tau(t))z^T(t-\tau(t))Qz(t-\tau(t))$

$V_2(t) = \int_{t-\tau_m}^t z^T(\alpha)ER_1z(\alpha)d\alpha + \int_{t-\bar{\tau}_M}^t z^T(\alpha)R_2z(\alpha)d\alpha$

$V_3(t) = \int_{t-\tau_m}^t \int_{t+\beta}^t \dot{z}^T(\alpha)\bar{E}^T Z_1\bar{E}\dot{z}(\alpha)d\alpha d\beta$
$$+ \int_{t-\bar{\tau}_M}^{t-\tau_m} \int_{t+\beta}^t \dot{z}^T(\alpha)\bar{E}^T Z_2\bar{E}\dot{z}(\alpha)d\alpha d\beta$$

From $\tau_m \leq \tau(t) \leq \bar{\tau}_M$. And taking the derivative of $V(t)$ with respect to t along the solution of (4) yields

$\dot{V}(t) = \dot{z}^T(t)\bar{E}^T(I_{N-1}\otimes P)\bar{E}z(t) + z^T(t)\bar{E}^T(I_{N-1}\otimes P)\bar{E}\dot{z}(t) - z^T(t-\tau(t))Qz(t-\tau(t)) + z^T(R_1+R_2)z(t) - z^T(t-\tau_m)R_1z(t-\tau_m) - z^T(t-\bar{\tau}_M)R_2z(t-\bar{\tau}_M) + \tau_m\dot{z}^T(t)\bar{E}^T Z_1\bar{E}\dot{z}(t) + (\bar{\tau}_M - \tau_m)z^T(t)\bar{E}^T Z_2\bar{E}\dot{z}(t) - \int_{t-\tau_m}^t \dot{z}^T(\alpha)\bar{E}^T Z_1\bar{E}\dot{z}(\alpha)d\alpha - \int_{t-\bar{\tau}_M}^{t-\tau_m} \dot{z}^T(\alpha)\bar{E}^T Z_2\bar{E}\dot{z}(\alpha)d\alpha.$

For any free-weighting matrices *L*, *M* and *N*, we have

$2\eta^T(t)L\left[\bar{E}z(t) - \bar{E}z(t-\tau_m) - \int_{t-\tau_m}^t \bar{E}\dot{z}(\alpha)d\alpha\right] = 0$

$2\eta^T(t)M\left[\bar{E}z(t-\tau_m) - \bar{E}z(t-\tau(t)) - \int_{t-\tau(t)}^{t-\tau_m}\bar{E}\dot{z}(\alpha)d\alpha\right] = 0$

$2\eta^T(t)N\left[\bar{E}z(t-\tau(t)) - \bar{E}z(t-\bar{\tau}_M) - \int_{t-\bar{\tau}_M}^{t-\tau(t)}\bar{E}\dot{z}(\alpha)\right] = 0$

where $\eta(t) = [z^T(t), F^T(z(t)), z^T(t-\tau(t)), z^T(t-\tau_m), z(t-\bar{\tau}_M), \varepsilon^T(kh)]^T$. Then, we have

$-2\eta^T(t)L\int_{t-\tau_m}^t \bar{E}\dot{z}(\alpha)d\alpha \leq \tau_m\eta^T(t)LZ^{-1}L^T\eta(t)$
$$+ \int_{t-\tau_m}^t \dot{z}^T\bar{E}^T Z_1\bar{E}\dot{z}(\alpha)d\alpha$$

$-2\eta^T(t)M\int_{t-\tau(t)}^{t-\tau_m}\bar{E}\dot{z}(\alpha)d\alpha \leq (\tau(t)-\tau_m)\eta^T(t)$

$MZ_1^{-1}M^T\eta(t) + \int_{t-\tau(t)}^{t-\tau_m}\dot{z}^T\bar{E}^T Z_2\bar{E}z(\alpha)d\alpha$

$-2\eta^T(t)N\int_{t-\bar{\tau}_M}^{t-\tau(t)}\bar{E}\dot{z}(\alpha)d\alpha \leq (\tau_M - \tau(t))\eta^T(t)$

$NZ^{-1}N^T\eta(t) + \int_{t-\bar{\tau}_M}^{t-\tau(t)}\dot{z}^T(\alpha)\bar{E}^T Z_2\bar{E}z(\alpha)d\alpha$

By using the Kronecker product, the event-triggered scheme can be written in the compact form as follows:

$\varepsilon^T(kh)(I_{N-1} \otimes \Phi_1)\varepsilon(kh)$
$$= e^T(kh)(T_1^T I_{N-1}T_1 \otimes \Phi_1)e(kh).$$

It is obviously that

$\varepsilon^T(kh)(I_{N-1} \otimes \Phi_1)\varepsilon(kh) \leq \lambda e^T(kh)(I_N \otimes \Phi_1)e(kh)$

$\leq \lambda[z(kh) - \varepsilon(kh)]^T[(T_2^T L^T \Lambda LT_2) \otimes \Phi_2][z(kh) - \varepsilon(kh)]$

where λ is the largest eigenvalue of $T_1^T T_1$. $\Lambda = \text{diag}\{\delta_1, \delta_2, \cdots \delta_N\}$.

Since $\|f(x_i(t)) - f(x_j(t))\| \leq \mu\|x_i(t) - x_j(t)\|$, Then, $F^T(z(t))F(z(t)) \leq \mu^2 z^T(t)(I_{N-1} \otimes I_n)z(t)$. To sum up, we can get

$\dot{V}(t) \leq \eta^T(t)\Omega\eta(t)$

where
$\Omega = \Psi + \Upsilon + \Upsilon^T + \tau_m\Psi_1 Z_1^{-1}\Psi_1^T + (\bar{\tau}_M - \tau_m)\Psi_2 Z_1^{-1}\Psi_2^T + \tau_m L Z_1^{-1}L^T + (\tau(t) - \tau_m)MZ_1^{-1}M^T + ((\bar{\tau}_M - \tau(t))NZ_1^{-1}N^T.$

The inequalities (6) and (7) are equivalent to $\Omega < 0$, that is $\dot{V}(t) < 0$, which implies system (4) is asymptotically stable. Hence, system (4) is admissible. From Definition 1, the multi-agent system (1) achieves admissible consensus. This completes the proof.

4. Simulation example

Consider a SMAS with six agents shown in Figure 2. There is a directed spanning tree in this graph.
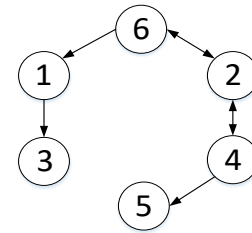
Figure 2. The interaction topology of agents

The dynamics of each agent described by (1) with $f(x_i(t)) = [0 \quad 0 \quad -\mu\sin(x_{i1})]^T$, and

$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, A = \begin{bmatrix} -2 & 1 & 0.5 \\ 0 & 0 & -1 \\ 2 & 1 & -1 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

Let $h = 0.05, \tau_m = 0.02, \bar{\tau}_M = 0.07, \mu = 0.5$ and $x_i = [x_{i1}, x_{i2}, x_{i3}, x_{i4}, x_{i5}, x_{i6}]^T$. The initial state is selected as

$$[x_1(0) \quad x_2(0) \quad x_3(0) \quad x_4(0) \quad x_5(0) \quad x_6(0)]$$
$$= \begin{bmatrix} 1.5 & -1 & 1 & 1.5 & -1.5 & 0.5 \\ 1 & -1.5 & 1.5 & 1 & -1 & -1.5 \\ -1.5 & 0.5 & -1 & 1.5 & 1 & -0.5 \end{bmatrix}$$

Solve the matrix inequalities (5), (6), (7), we can obtain
$K = [0.0445 \quad -0.2153 \quad 0.1489]$, and

$$\Phi = \begin{bmatrix} 29.4174 & 9.0284 & 9.4232 \\ 9.0284 & 43.4806 & 32.8987 \\ 9.4232 & 32.8987 & 48.9758 \end{bmatrix}$$

The state trajectories of system (1) are shown in Figure 3, 4, 5. One can clearly see that all agent's states can indeed reach consensus.
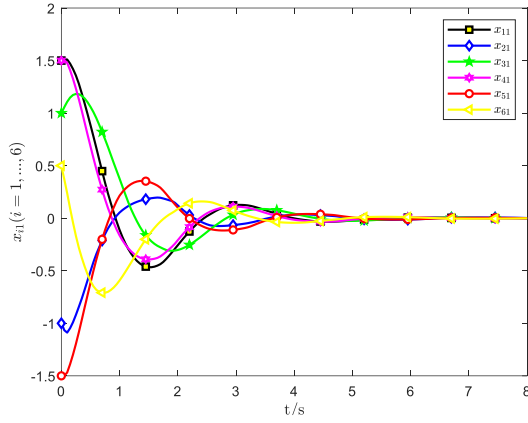
Figure 3. State trajectories $x_{i1,}(i = 1, 2 \cdots 6)$
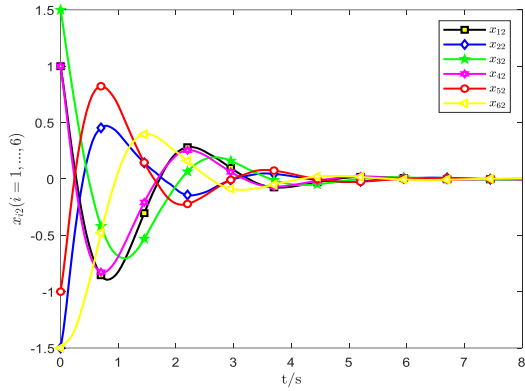
Figure 4. State trajectories $x_{i2,}(i = 1, 2 \cdots 6)$
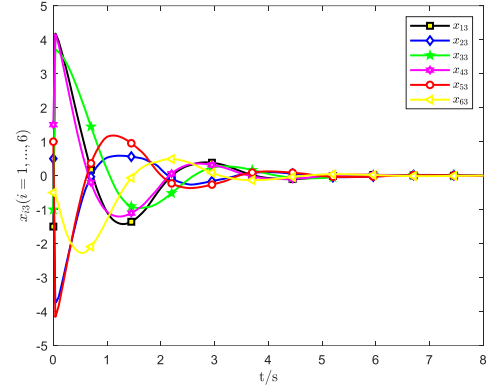
Figure 5. State trajectories $x_{i3,}(i = 1, 2 \cdots 6)$

The transmission instants and release intervals are illustrated in Figure 6, which shows that the number of the sampled-data transmission is significantly decreased.
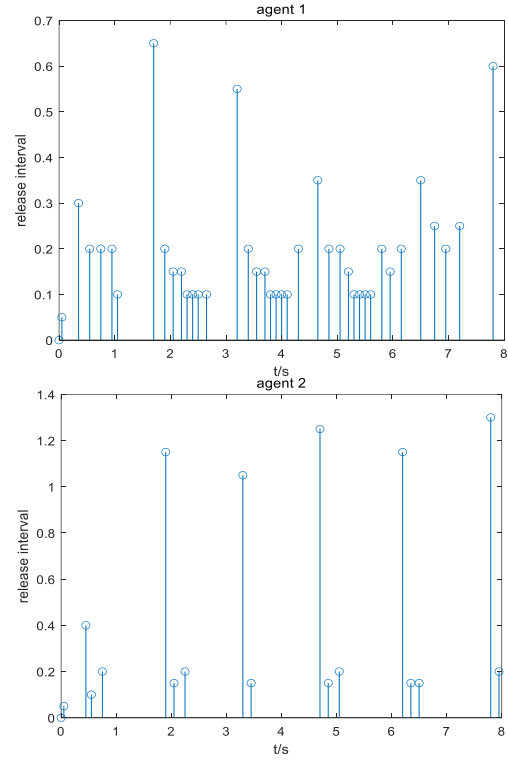
Figure 6. Transmission instants and release interval of agent $i(i = 1, 2)$

5. Conclusion

Event-triggered consensus problem of nonlinear SMAS with directed topologies has been investigated in this paper. The conditions of achieving event-triggered consensus were obtained, while consensus control gain matrix is designed. The effectiveness of the proposed method was verified by a numerical example. It would be important and more practical to study consensus for SMAS with leaders or under switching topologies, which will be our further analysis.

Acknowledgements

References

1. T. Liu and Z. Jiang, Distributed formation control of nonholonomic mobile robots without global position measurements, Automatica, 49(2), 2013, pp. 592-600.
2. X. Ge, Q. Han and X. Zhang, Achieving cluster formation of multi-agent systems under aperiodic sampling and communication delays, IEEE Transactions on Industrial Electronics, 65(4), 2018, pp. 3417-3426.
3. X. Dong, B. Yu and Shi Z. Time-varying formation control for unmanned aerial vehicles: theories and applications, IEEE Transactions on Control Systems Technology, 23(1), 2014, pp. 340-348.
4. J. He, H. Li and J. Chen, Study of consensus-based time synchronization in wireless sensor networks, ISA Transactions, 53(2), 2014, pp. 347-357.
5. H. Ji, L. Lewis and Z. Hou, Distributed information-weighted Kalman consensus filter for sensor networks, Automatica, 77, 2017, pp. 18-30.
6. Y. Zhang and Y. Tian, A fully distributed weight design approach to consensus Kalman filtering for sensor networks, Automatica, 104, 2019, pp. 34-40.
7. C. Chen, G. Wen and Y. Liu, Adaptive consensus control for a class of nonlinear multiagent time-delay systems using neural networks, IEEE transactions on neural networks and learning systems, 25(6), 2014, pp. 1217-1226.
8. G. Dong, H. Li and H. Ma, Finite-Time consensus tracking neural network FTC of multi-agent systems, IEEE Transactions on Neural Networks and Learning Systems, 32(2), 2020, pp. 653-662.
9. Y. Wang, H. Li and X. Qiu, Consensus tracking for nonlinear multi-agent systems with unknown disturbance by using model free adaptive iterative learning control, Applied Mathematics and Computation, 365, 2020, pp. 124701.
10. D. V. Dimarogonas, E. Frazzoli and K. H. Johansson, Distributed Event-triggered control for multi-agent systems, IEEE Transactions on Automatic Control, 57(5), 2012, pp. 1291-1297.
11. W. Zhu, Z. Jiang and G. Feng, Event-based consensus of multi-agent systems with general linear models, Automatica, 50(2), 2014, pp. 552-558.
12. E. Garcia, Y. Cao and D. W. Casbeer, Periodic event-triggered synchronization of linear multi-agent systems with communication delays, IEEE Transactions on Automatic Control, 62(1), 2015, pp. 366-371.
13. X. Fan and Wang Z. Event-triggered sliding mode control for singular systems with disturbance, Nonlinear Analysis Hybrid Systems, 40(11), 2021, pp. 101011.
14. H. Wang, R. Lu and A. Xue, Network-based H∞ control for singular systems with event-triggered sampling scheme, Information Sciences, 329, 2016, pp. 540-551.
15. M. Li, J. Xia, J. Zhao. Event-triggered extended dissipative control for networked singular systems. International Journal of Control Automation and Systems, 19(1), 2021, pp. 382–391.

Authors Introduction

Dr. Lin Li

She received her Ph.D. degree in Control Theory and Control Engineering from Beihang University in 2010, Beijing, China. She is currently an Associate Professor at University of Shanghai for Science and Technology. Her current research interests include robust control, filtering, the multi-agent systems control, and machine learning.

Mr. Tong Yuan

He received his Master degree in Applied Mathematics from Shandong University of Science and Technology in 2022. He is currently a PhD student at the University of Shanghai for Science and Technology, China.

An Improved Conversion Technique from EPNAT Models to VDM++ Specifications for Simulation of Abstract Software Behavior

Sho Matsumoto
Graduate School of Science for Creative Emergence, Kagawa University
2217-20 Hayashi-cho, Takamatsu-shi, Kagawa 761-0396, Japan

Ryoichi Ishigami
Graduate School of Science for Creative Emergence, Kagawa University
2217-20 Hayashi-cho, Takamatsu-shi, Kagawa 761-0396, Japan

Tetsuro Katayama
Department of Computer Science and Systems Engineering, Faculty of Engineering, University of Miyazaki
1-1 Gakuen-kibanadai nishi, Miyazaki-shi, Miyazaki 889-2192, Japan

Tomohiko Takagi
Department of Engineering and Design, Faculty of Engineering and Design, Kagawa University
2217-20 Hayashi-cho, Takamatsu-shi, Kagawa 761-0396, Japan
Email: s23g364@kagawa-u.ac.jp, takagi@eng.kagawa-u.ac.jp

Abstract

Formal software models based on EPNAT (Extended Place/transition Net with Attributed Tokens) can be converted to VDM++ specifications that enable simulation of abstract software behavior before implementation processes. However, the conversion technique has two problems, that is, (1) extracting all properties to be checked from the VDM++ specifications requires time and effort, and (2) the structure of the VDM++ specifications has less readability and maintainability. In this study, we improve the conversion technique by (1) adding features to extract an abstract current state of software, and (2) dividing into classes that correspond to subnets of EPNAT models. This paper shows a new conversion rule, a new structure of VDM++ specifications, a simple example, and the discussion about their effectiveness.

Keywords: Software model, Place/transition net, Vienna development method, Conversion technique

1. Introduction

Formal software models are useful to define specifications and designs clearly, and to find failures early in software development. EPNAT (Extended Place/transition Net with Attributed Tokens), which is based on place/transition nets and VDM++, has been proposed as a formal software modeling language especially for concurrent and distributed software that is composed of multiple subsystems (and/or systems) [1]. In an EPNAT model (that is, a formal software model that is defined by using EPNAT), the behavior of each subsystem is defined as a subnet, and the interactions between the subsystems are defined as glue transitions. EPNAT models can be converted to VDM++ specifications, and enable simulation of abstract software behavior before implementation processes [2].

However, the conversion technique has two problems. One is that extracting all properties (that is, data to be checked in the simulation) from the VDM++ specifications requires time and effort. In the simulation, a common VDM++ interpreter [3] is used to execute the VDM++ specifications, and moreover, a particular feature of our support tool to manage its original EPNAT model needs to be used to process and obtain some properties. Another is that the structure of the VDM++ specifications has less readability and maintainability. The structure is not appropriately divided into modules that correspond to the elements of its original EPNAT model.

To tackle the above problems, we propose an improved conversion technique from EPNAT models to VDM++ specifications. In this technique, features to extract an abstract current state of software are added to the VDM++ specifications in order to tackle the first problem. Additionally, the structure of the VDM++ specifications is divided into classes that correspond to subnets of EPNAT models in order to tackle the second problems. In general, the important things in the conversion of

Sho Matsumoto, Ryoichi Ishigami, Tetsuro Katayama, Tomohiko Takagi

software models are correctness and automation [4], and thus we try to keep or improve them in this study. This paper shows a new conversion rule, a new structure of VDM++ specifications, and a simple example in section 2. The discussion about their effectiveness is given in section 3, and then conclusion and future work are shown in section 4.

2. Conversion Technique

The improvement of the previous technique to convert from EPNAT models to VDM++ specifications consists of introducing the features to obtain marking data (discussed in section 2.1) and the basic class structure of the VDM++ specifications (discussed in section 2.2). The former is proposed for tackling the first problem, and the latter is proposed for tackling the second problem.

For the discussion in this study, we have constructed an experimental EPNAT model of BSS (imaginary Business Support Software) that is composed of a purchase subsystem, a production subsystem, and a stock management subsystem. Its overview is shown in Fig. 1.
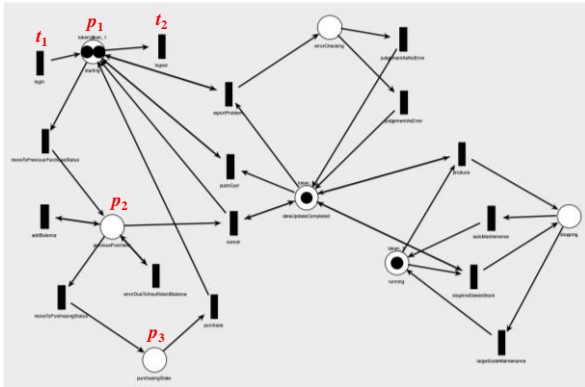
Fig. 1. Overview of an experimental EPNAT model.

2.1. *Features to obtain marking data*

The features to extract an abstract current state of software should be added to the VDM++ specifications. In EPNAT models, the abstract states are expressed as the distributions of attributed tokens, that is, markings. Therefore, in this technique, the features to obtain original markings and partitioned markings are added to the VDM++ specifications. The idea of the partitioned markings has been derived from equivalence partitioning testing and coverability trees [5].

An original marking is the list of the numbers of attributed tokens in every place. For example, the original marking in the purchase subsystem of BSS (Fig. 1) is expressed as $[p_1, p_2, p_3] = [2, 0, 0]$. In general, there is a huge number of original markings in an EPNAT model, and it is not so easy to cover all of them in the simulation. In partitioned markings, the ranges of the numbers of attributed tokens in which software seems to have the

same behavior are expressed as symbols (the ranges are defined by engineers). A symbol represents all the numbers of attributed tokens in a corresponding range, and therefore the number of markings to be covered can be reasonably reduced. For example, in the purchase subsystem of BSS (Fig. 1), an execution path by using original markings $[1, 0, 0]$->t_1->$[2, 0, 0]$->t_1->$[3, 0, 0]$->t_2->$[2, 0, 0]$->t_2->$[1, 0, 0]$ can be converted into an execution path by using partitioned markings $[1, 0, 0]$->t_1->$[w, 0, 0]$->t_1->$[w, 0, 0]$->t_2->$[w, 0, 0]$->t_2->$[1, 0, 0]$, when the number of attributed tokens in p_1 in the range of two or more is represented as a symbol w. In this example, the original markings $[2, 0, 0]$ and $[3, 0, 0]$ are replaced with the partitioned marking $[w, 0, 0]$.

In the simulation, if possible, a search algorithm should avoid to revisit the same partitioned markings to improve its efficiency. The features to obtain marking data will be frequently used in the simulation, since the marking data is important properties to capture the abstract software behavior. The way of adding the features to VDM++ specifications is discussed in the next section.

2.2. *Basic class structure of VDM++ specifications*

The structure of VDM++ specifications should be divided into modules that correspond to the elements of its original EPNAT model. Therefore, the basic class structure of VDM++ specifications is constructed in this study.

The basic class structure is shown in Fig. 2. It is composed of five kinds of classes, that is, "Software", "Subsystem", "AttributedToken", "Places", and "Marking". "Subsystem" and "AttributedToken" are the sets of classes for specifying each subsystem. In each of the sets, there are N classes (N expresses the number of subsystems), and one class peculiar to each subsystem needs to be defined in the VDM++ specifications. The others are classes for managing the data of all of the subsystems. In the VDM++ specifications, they need to be defined as one class, respectively.

"Subsystem" is a part of "Software", and corresponds to the subnets of the original EPNAT model. There are relations between "Subsystem$_i$" and "AttributedToken$_i$", and also there are relations between "AttributedToken$_i$" and "Places" ($1 \leq i \leq N$; *ClassName$_i$* expresses the *i*-th class in the set *ClassName*). They mean that each subsystem keeps corresponding attributed tokens in its own places. There are relations between "Marking" and "Subsystem$_i$", which means that original markings and partitioned markings are derived from the places of subsystems. The operations that correspond to glue transitions are defined in "Software", since they have a role in the interaction between different subsystems. The other transitions (hereinafter, referred to as normal transitions) are peculiar to each subsystem, and thus the operations that correspond to normal transitions are

defined in corresponding "Subsystem*ᵢ*". Executing the operations of the transitions changes the current marking, which corresponds to arcs. Therefore, the VDM++ specifications have no explicit elements that correspond to arcs.

Fig. 3 shows the overview of the VDM++ specifications of the experimental EPNAT model that are based on the basic class structure. The classes "PurchaseSubsystem", "StockManagementSubsystem", and "ProductionSubsystem" correspond to "Subsystem" of the basic class structure. The classes "Customer", "Goods", and "ProductionMachine" correspond to "AttributedToken" of the basic class structure.
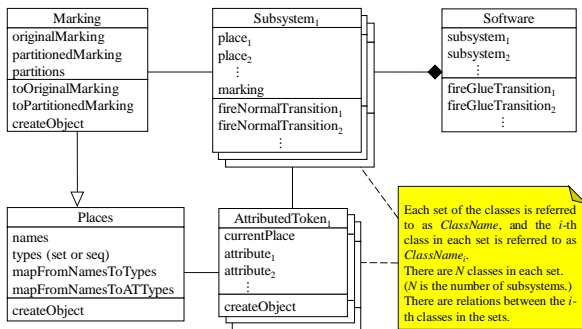
Fig. 2. Overview of the basic class structure.
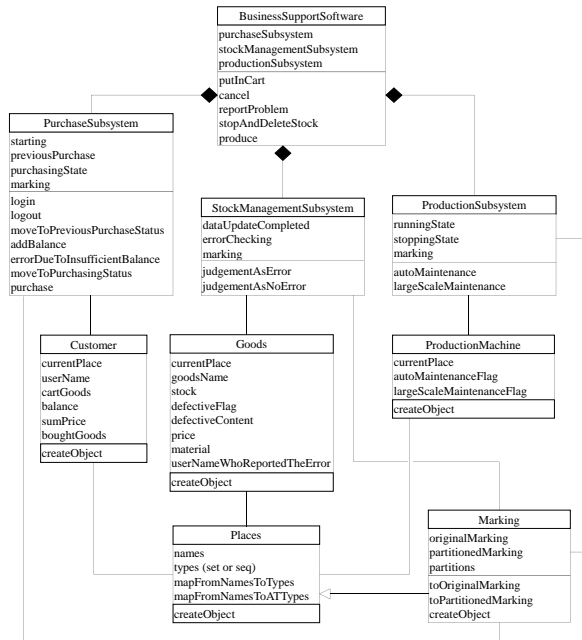
Fig. 3. Overview of VDM++ specifications
of the experimental EPNAT model
based on the basic class structure.

3. Discussion

This section gives the discussion about the effectiveness of this technique. We have manually converted the experimental EPNAT model to VDM++

specifications by using this technique. Also, we have converted the experimental EPNAT model to VDM++ specifications by using the previous technique. Hereinafter, the former is referred to as improved specifications, and the latter is referred to as previous specifications. We have compared the improved specifications and the previous specifications, and have found the followings.

- The improved specifications can additionally provide marking data (original markings and partitioned markings). Also, the improved specifications can provide the information about individual attributed tokens by subsystems, and variables by subsystems. These will be useful to search the VDM++ specifications for the failures of software in the simulation.
- The elements of EPNAT models correspond to the elements of the improved specifications. There is traceability between EPNAT models and the improved specifications to some extent. Additionally, the improved technique can avoid the errors that the actions of normal transitions of a subsystem try to access the elements of other subsystems, since subsystems are separately defined as classes in the improved specifications.
- The improved specifications and the previous specifications consist of about 440 and 280 lines of code in VDM++, respectively. The previous specifications do not have the features to obtain marking data discussed in section 2.1, and the improved specifications excluding the features consist of about 340 lines of code. In any case, the improved specifications are larger than the previous specifications.

4. Conclusion and Future Work

In this paper, we have proposed an improved conversion technique from EPNAT models to VDM++ specifications for the simulation of abstract software behavior. The improvement consists of (1) introducing the features to obtain marking data and (2) introducing the basic class structure of the VDM++ specifications. Regarding (1), the features will be frequently used in the simulation, since the marking data is important properties to capture the abstract software behavior. We have concentrated on obtaining marking data, but other properties also need to be collected to be efficiently checked in the simulation. The features to collect other properties will be considered in future work. Regarding (2), the basic class structure establishes traceability between EPNAT models and VDM++ specifications to some extent, and avoids a certain kind of errors. However, VDM++ specifications by the improved conversion technique are larger than VDM++ specifications by the previous conversion technique. The basic class structure

Sho Matsumoto, Ryoichi Ishigami, Tetsuro Katayama, Tomohiko Takagi

may need to be refined further to improve the readability and maintainability in future work.

Acknowledgements

References

1. T. Takagi and R. Kurozumi, Software Modeling Technique and its Prototype Tool for Behavior of Multiple Objects Using Extended Place/Transition Nets with Attributed Tokens, *Journal of Robotics, Networking and Artificial Life*, Vol.7, No.3, pp.194-198, Dec. 2020.
2. S. Matsumoto, T. Katayama and T. Takagi, Automated Random Simulation for Checking a Behavioral Model of Systems Based on Extended Place/Transition Net with Attributed Tokens, *Proceedings of International Conference on Artificial Life and Robotics*, online, Japan, pp.337-340, Feb. 2023.
3. J. Fitzgerald, P.G. Larsen, P. Mukherjee, N. Plat and M. Verhoef, *Validated Designs for Object-Oriented Systems*, Springer-Verlag London, 2005.
4. K. Lano and S. Kolahdouz-Rahimi, Model Transformation Specification and Design, *Advances in Computers*, Vol.85, pp.123-163, Elsevier, 2012.
5. T. Takagi and Z. Furukawa, Test Case Generation Technique Based on Extended Coverability Trees, *Proceedings of 13th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing*, IEEE, Kyoto, Japan, pp. 301-306, Aug. 2012.

Authors Introduction

Mr. Sho Matsumoto

He received the B.S. degree from Kagawa University in 2023. He is a master's student in the Graduate School of Science for Creative Emergence at Kagawa University. His research interests are in software engineering, particularly software quality control.

Mr. Ryoichi Ishigami

He received the B.S. degree from Kagawa University in 2023. He is a master's student in the Graduate School of Science for Creative Emergence at Kagawa University. His research interests are in software engineering, particularly software quality control.

Dr. Tetsuro Katayama

He received a Ph.D. degree in engineering from Kyushu University, Fukuoka, Japan, in 1996. From 1996 to 2000, he has been a Research Associate at the Graduate School of Information Science, Nara Institute of Science and Technology, Japan. Since 2000 he has been an Associate Professor at the Faculty of Engineering, Miyazaki University, Japan. He is currently a Professor with the Faculty of Engineering, University of Miyazaki, Japan. His research interests include software testing and quality. He is a member of the IPSJ, IEICE, and JSSST.

Dr. Tomohiko Takagi

He received the B.S., M.S. and Ph.D. degrees from Kagawa University in 2002, 2004 and 2007, respectively. He became an assistant professor in 2008, and a lecturer in 2013 in the Faculty of Engineering at Kagawa University. Since 2018 he has been an associate professor in the Faculty of Engineering and Design at Kagawa University. His research interests are in software engineering.

Prototype of RAGESS Which Is a Tool for Automatically Generating SwiftDiagram to Support iOS App Development

Haruki Onaga*, Tetsuro Katayama*, Yoshihiro Kita†,
Hisaaki Yamaba*, Kentaro Aburada*, Naonobu Okazaki*

*Department of Computer Science and Systems Engineering, Faculty of Engineering, University of Miyazaki,
1-1 Gakuen-kibanadai nishi, Miyazaki, 889-2192 Japan
†Department of Information Security, Faculty of Information Systems, Siebold Campus, University of Nagasaki
1-1-1 Manabino, Nagayo-cho, Nishi-Sonogi-gun, Nagasaki, 851-2195 Japan
E-mail: onaga@earth.cs.miyazaki-u.ac.jp, kat@cs.miyazaki-u.ac.jp, kita@sun.ac.jp,
yamaba@cs.miyazaki-u.ac.jp, aburada@cs.miyazaki-u.ac.jp, oka@cs.miyazaki-u.ac.jp

Abstract

It is difficult to understand the structure of large and complex mobile applications. To support iOS app development, we proposed SwiftDiagram, a visualization of the static structure of Swift source code and confirmed its high usefulness. However, manually drawing SwiftDiagram is labor-intensive. This paper implemented a prototype of RAGESS (Real-time Automatic Generation of SwiftDiagram System), a tool that automatically generates SwiftDiagram by performing static analysis on Swift source code every time an iOS app build is successful.

Keywords: Software visualization, Mobile application, Swift

1. Introduction

The smartphone and tablet application market is expanding every year [1]. As a result, mobile applications are becoming larger and more complex. The following problems exist for developers of increasingly complex mobile applications.

- Difficult to understand the overall structure of the application
- Difficult to keep track of where changes to the source code may have an impact

To solve the above problems through software visualization, we have proposed SwiftDiagram. SwiftDiagram is a diagram that visualizes the static structure and impact scope of source code written in Swift, a programming language used to develop iOS applications. However, manually drawing a SwiftDiagram takes time and effort. Therefore, this paper implements RAGESS (Real-time Automatic Generation of SwiftDiagram System) to support the development of iOS applications in Swift. RAGESS is a tool that performs a static analysis of the Swift source code and automatically draws the corresponding SwiftDiagram when it detects that the target application builds successfully.

2. SwiftDiagram

SwiftDiagram is a diagram that targets source code written in the Swift programming language and supports

the design and maintenance of iOS applications by visualizing the following:

- Static structure of type
- Affected scope when making changes to the type

In addition, SwiftDiagram is composed of the following four types of parts and one type of arrow. The appearance of each part and arrow is shown in Fig. 1.

- **Header Part**
  Shows a type identifier, such as type categories and name
- **Details Part**
  Shows type components such as properties, methods, protocol conformance, and so on
- **Extension Part**
  Shows a type extension
- **Nest Part**
  Shows a type that is declared nested inside another type
- **Affected scope (arrow)**
  Shows that changes made to the type at the root of the arrow may affect the component of the type pointed to by the arrow tip

Haruki Onaga, Tetsuro Katayama, Yoshihiro Kita, Hisaaki Yamaba, Kentaro Aburada, Naonobu Okazaki
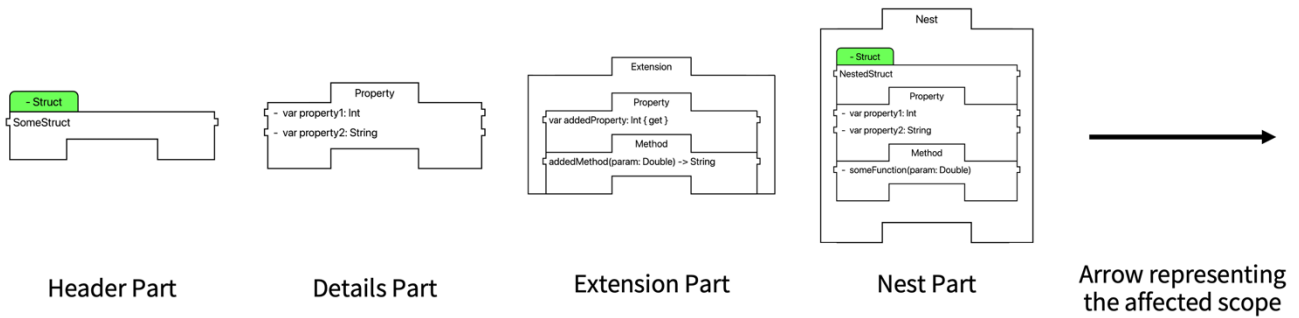
Fig. 1 Component parts of SwiftDiagram
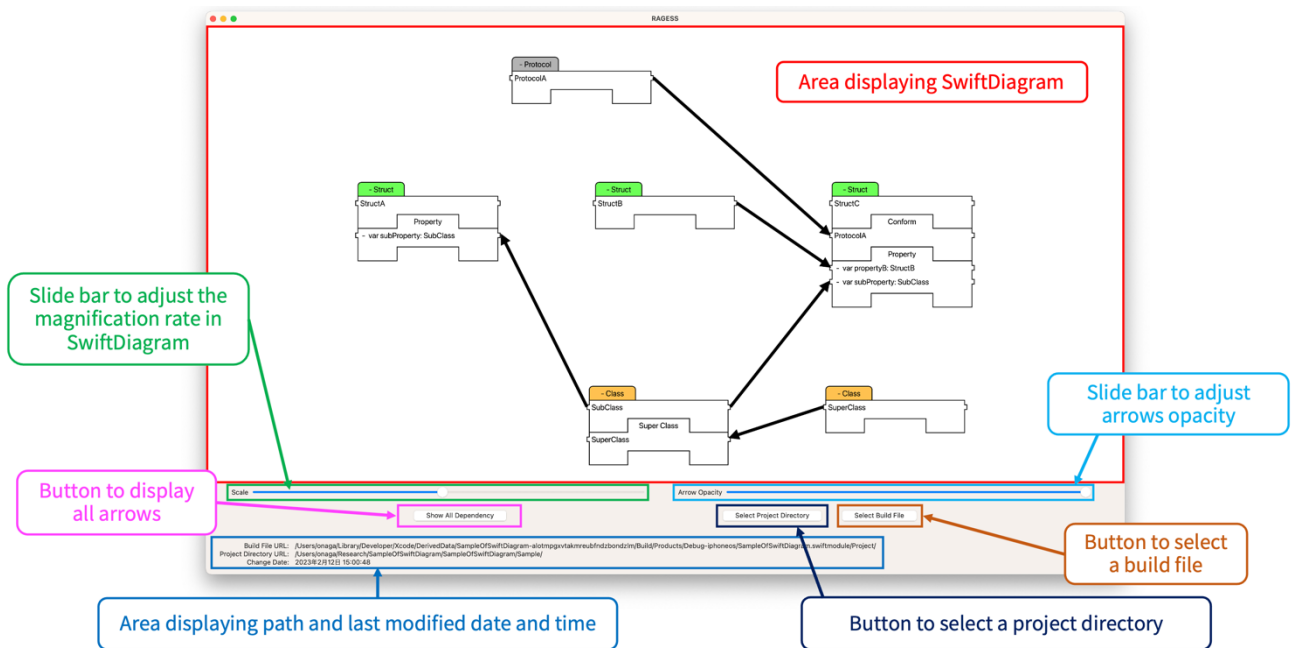
Fig. 2 Overview of RAGESS

3. RAGESS

We implement RAGESS as a macOS application in Swift. After launching RAGESS, the user can use RAGESS functions by selecting the path to the project and its build file for which the user wants to draw SwiftDiagram. This chapter describes the overview and functions of RAGESS.

3.1. *Overview of RAGESS*

Fig. 2 shows the overview of RAGESS implemented in this paper. RAGESS consists of two areas, two slide bars, and three buttons as shown below.

- Area displaying SwiftDiagram
- Area displaying path and last modified date and time
- Slide bar to adjust the magnification rate in SwiftDiagram
- Slide bar to adjust arrows opacity
- Button to display all arrows
- Button to select a project directory
- Button to select a build file

3.2. *Functions of RAGESS*

RAGESS has the following two functions.

- Automatic real-time drawing of SwiftDiagram corresponding the latest source code
- Narrowing down the affected scope

RAGESS monitors the build file of the target project. RAGESS statically analyzes the project's Swift source code every time it detects a change in the build file and automatically draws the corresponding SwiftDiagram. This saves the user from the trouble of updating the SwiftDiagram and means that the source code represented by the SwiftDiagram doesn't contain any syntax errors. Static analysis of Swift source code is
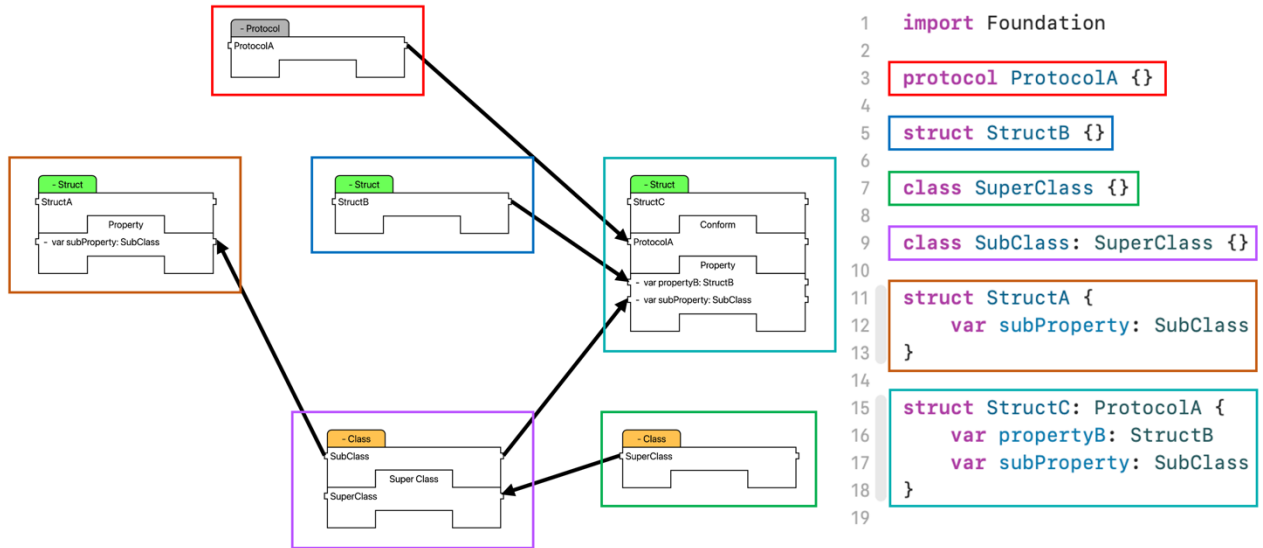
Fig. 3 Correspondence between SwiftDiagram and Swift source code in the sample project

performed by using SwiftSyntax [2], which is open source by Apple. The user can scale the drawn SwiftDiagram by operating the slide bar.

Representing the overall structure of a project in SwiftDiagram could complicate the arrows representing the affected scope and make it difficult for the user to understand the structure. When the user selects a header part, RAGESS narrows down the display to only those arrows that represent the affected scope of the changes made to the type. The user can also operate the slide bar to adjust the opacity of the arrows representing the affected scope. The button to display all arrows, returns the state of the narrowed down affected scope and the opacity of the arrows to their initial state, and displays all the arrows that represent the affected scope for the entire project.

4. Application Example

By using an application example, we verify that the RAGESS functionality works correctly for Swift source code that defines multiple types. Fig. 3 shows the correspondence between the SwiftDiagram drawn by RAGESS and the Swift source code in the example. Here, Fig. 2 shows a screen shot of RAGESS monitoring the build file of the project containing the Swift source code in the example, and when the build is successful.

By looking at the StructA of the SwiftDiagram and the Swift source code in Fig. 3, the structure name and properties of the structure in the SwiftDiagram correspond to the structure name and properties of the structure in the Swift source code, respectively. Similarly, looking at StructC, the structure name, properties, and dependencies by protocol conformance in the

SwiftDiagram correspond to the structure name, properties, and dependencies by protocol conformance in the Swift source code, respectively.

Also, ProtocolA, SuperClass, SubClass, and StructB in the SwiftDiagram extend arrows toward components of the type that may be affected by changes made to the type.

Hence, it is clear that RAGESS can statically analyze Swift source code and draw the corresponding SwiftDiagram.

5. Related work

Emerge [3] is a tool for visualizing Swift source code. It is a browser-based tool that visualizes codebase and dependencies for multiple programming languages like Swift, Kotlin, and TypeScript. Because it visualizes the source code in a graph structure composed of nodes and edges, the user can explore and analyze it visually.

Emerge analyzes source files under the source directory that are written in a language set by the user and generates files for displaying graphs in a Web browser. The user can load the file with a Web browser to analyze the dependencies. However, Emerge cannot reflect changes made to the target source code after analysis in the graph displayed in the Web browser. Also, because it takes types and files as nodes, the user cannot analyze the dependencies of the type components. On the other hand, RAGESS automatically draws SwiftDiagram every time the target application builds successfully. In addition, because arrows representing the dependencies are connected to type components such as properties and methods, the user can understand the dependencies in more detail.

Haruki Onaga, Tetsuro Katayama, Yoshihiro Kita, Hisaaki Yamaba, Kentaro Aburada, Naonobu Okazaki

And another tool for visualizing Swift source code is Swiftcity [4], [5]. Swiftcity adapts the city metaphor, which has been studied for Java, C++, and other languages, to unique type extensions, structures, and more of Swift to map and visualize source code into cities and buildings. The height of the building represents the number of lines of code (LOC) of the type, while the width and depth of the building represent the number of methods of the type. The user can use Swiftcity by uploading a project file to the user's Web browser.

Swiftcity is useful for getting an overview of the application. For example, yellow buildings represent type extension, so if a city has many yellow buildings, it means that type extension are frequently used in its source code. The variety of building colors also means the use of all the standard features provided by Swift. However, Swiftcity visualizes type categories and the number of methods, but doesn't visualize elements such as properties and enumerated type cases, nor the dependencies such as function calls and protocol conformance. Therefore, use of Swiftcity is limited to understanding and analyzing an overview of the application in the maintenance process. On the other hand, RAGESS can visualize the size of a type by the height of the SwiftDiagram that combines each part and also components and dependencies of the type.

Therefore, RAGESS provides a more detailed visualization of the structure of applications than the other two tools that visualize Swift source code. In addition, it updates the SwiftDiagram in real-time every time the application builds successfully, thus reducing the time the user needs to modify the deliverables.

6. Conclusion

In this paper, we have implemented RAGESS, a tool statically analyzes the project's Swift source code every time it detects a change in the build file and automatically draws the corresponding SwiftDiagram, in order to support the development of iOS applications in the Swift programming language.

We have applied RAGESS to Swift source code and confirmed that it can draw the corresponding SwiftDiagram. We also have confirmed that RAGESS can visualize the structure of the application in more detail and in real-time compared to other tools that visualize Swift source code.

Consequently, RAGESS is expected to support the development of iOS applications in the Swift programming language.

Future works are as follows:

- **Evaluation of usefulness by subject experimentation**

RAGESS cannot coexist with the current Swift and SwiftSyntax versions and is not buildable. Therefore, in this paper, we couldn't evaluate the usefulness of RAGESS through experimentation in which subjects actually use RAGESS. We need to implement RAGESS to adapt it to current Swift and SwiftSyntax versions and conduct subject experimentation to confirm its usefulness.

- **Implementation of SwiftDiagram editing function**

We enable SwiftDiagram to edit in the area displaying SwiftDiagrams of RAGESS and reflect the edits in the source code. This function could assist the user in designing new developments or adding new function. We need to implement it because we believe it feature will further increase the usefulness of SwiftDiagram and RAGESS.

- **Implementation of search function**

In the current RAGESS, the user must visually search for the target type in the area displaying SwiftDiagram to understand the structure of the application. We believe that implementing the function to search by type name or component will further reduce the user's time and effort.

- **Extension of supported syntax**

RAGESS does not support type inference and can only extract properties whose types are explicitly indicated by type annotations. Since type inference is used in much Swift source code, we believe that supporting type inference would further increase the usefulness of RAGESS. And even if there are more nested types within nested declared types, RAGESS cannot extract that information. It is also not yet compatible with Swift Macros released with Swift 5.9. We plan to make RAGESS compatible with them.

References

1. Ministry of Internal Affairs and Communications, Japan, Information and Communications in Japan WHITE PAPER 2022, https://www.soumu.go.jp/johotsusintokei/whitepaper/eng/WP2022/2022-index.html
2. GitHub, swift-syntax, https://github.com/apple/swift-syntax (Accessed 2023-12-14)
3. GitHub, emerge, https://github.com/glato/emerge (Accessed 2023-12-14)
4. Rafael Nunes, Marcel Rebouças, Francisco Soares-Neto, Fernando Castor, Poster: Visualizing Swift Projects as Cities, IEEE/ACM 39th International Conference on Software Engineering Companion, pp. 368-370, 2017.
5. Swiftcity, https://swiftcity.github.io/swiftcity-app/ (Accessed 2023-12-14)

Authors Introduction

Mr. Haruki Onaga

Haruki Onaga received the Bachelor's degree in engineering (computer science and systems engineering) from the University of Miyazaki, Japan in 2023. He is currently a Master's student in Graduate School of Engineering at the University of Miyazaki, Japan. His research interests include software development support through software visualization.

Dr. Tetsuro Katayama

He received a Ph.D. degree in engineering from Kyushu University, Fukuoka, Japan, in 1996. From 1996 to 2000, he has been a Research Associate at the Graduate School of Information Science, Nara Institute of Science and Technology, Japan. Since 2000 he has been an Associate Professor at the Faculty of Engineering, Miyazaki University, Japan. He is currently a Professor with the Faculty of Engineering, University of Miyazaki, Japan. His research interests include software testing and quality. He is a member of the IPSJ, IEICE, and JSSST.

Dr. Yoshihiro Kita

Yoshihiro Kita received a Ph.D. degree in systems engineering from the University of Miyazaki, Japan, in 2011. He is currently an Associate Professor with the Faculty of Information Systems, University of Nagasaki, Japan. His research interests include software testing and biometrics authentication.

Dr. Hisaaki Yamaba

He received the B.S. and M.S. degrees in chemical engineering from the Tokyo Institute of Technology, Japan, in 1988 and 1990, respectively, and the Ph D. degree in systems engineering from the University of Miyazaki, Japan in 2011. He is currently an Assistant Professor with the Faculty of Engineering, University of Miyazaki, Japan. His research interests include network security and user authentication. He is a member of SICE and SCEJ.

Dr. Kentaro Aburada

He received the B.S., M.S, and Ph.D. degrees in computer science and system engineering from the University of Miyazaki, Japan, in 2003, 2005, and 2009, respectively. He is currently an Associate Professor with the Faculty of Engineering, University of Miyazaki, Japan. His research interests include computer networks and security. He is a member of IPSJ and IEICE.

Naonobu Okazaki

He received his B.S, M.S., and Ph.D. degrees in electrical and communication engineering from Tohoku University, Japan, in 1986, 1988 and 1992, respectively. He joined the Information Technology Research and Development Center, Mitsubishi Electric Corporation in 1991. He is currently a Professor with the Faculty of Engineering, University of Miyazaki since 2002. His research interests include mobile network and network security. He is a member of IPSJ, IEICE and IEEE.

Extension to Support Types and Operation/Function Definitions in BWDM to Generate Test Case Tool from the VDM++ Specification

Shota Takakura*, Tetsuro Katayama*, Yoshihiro Kita†,
Hisaaki Yamaba*, Kentaro Aburada*, and Naonobu Okazaki*
*Department of Computer Science and Systems Engineering, Faculty of Engineering, University of Miyazaki,
1-1 Gakuen-kibanadai nishi, Miyazaki, 889-2192 Japan
†Department of Information Security, Faculty of Information Systems, Siebold Campus, University of Nagasaki
1-1-1 Manabino, Nagayo-cho, Nishi-Sonogi-gun, Nagasaki, 851-2195 Japan
E-mail: takakura@earth.cs.miyazaki-u.ac.jp, kat@cs.miyazaki-u.ac.jp, kita@sun.ac.jp,
yamaba@cs.miyazaki-u.ac.jp, aburada@cs.miyazaki-u.ac.jp, oka@cs.miyazaki-u.ac.jp

Abstract

Generating test cases from the formal specification description VDM++, which is a method for disambiguating specifications, is time-consuming and labor intensive. Therefore, our laboratory has developed BWDM, a tool that automatically generates test cases from VDM++ specifications. However, existing BWDM has problems that it only supports integer types and cannot generate test cases for operation and function definitions including recursive structure. Therefore, in order to improve the usefulness of BWDM, this paper extends BWDM to solve the above problems. Consequently, it has confirmed that the use of extended BWDM can reduce the test case generation time compared to manual test case generation.

Keywords: software testing, formal methods, VDM++, automatic generation, test cases

1. Introduction

Today, our lives are supported by a lot of software. As a result, software has become larger and more complex, and software bugs have a greater impact on society. One of the causes of bugs in software is the use of natural language in the upstream stage of software development. This is because natural language contains ambiguity. One method to solve this problem is to design software using formal methods. VDM is one of the most common formal methods used for this purpose, and VDM++ is a formal specification description language used in VDM for object-oriented modeling [1].

On the other hand, software design using VDM++ requires software testing. However, it is time-consuming and labor-intensive to manually generate test cases from specifications written in VDM++. In addition, manual test case generation may cause leakage in testing. Therefore, our laboratory has developed BWDM [2], [3], an automatic test case generation tool for VDM++ specifications. However, existing BWDM has the problem that it only supports integer types and cannot generate test cases for operation and function definitions with recursive structure.

Therefore, to improve the usefulness of BWDM, it is extended as follows.

- Addition of a function to generate test cases for enumerated type
- Addition of a function to generate test cases for operations/functions with recursive structure

In this paper, Chapter 2 describes the existing BWDM, Chapter 3 describes the extended BWDM, Chapter 4 shows application examples, and Chapter 5 discusses and confirms the usefulness of BWDM. Chapter 6 is the conclusion.

2. BWDM

BWDM is a tool that automatically generates test cases using the VDM++ specification as input. As examples, Fig. 1 shows the VDM++ specification and Fig.2 shows the test cases generated from it by BWDM, respectively.

3. The Extended BWDM

This chapter describes the details of the extension of BWDM. Fig. 3 shows the structure of the extended BWDM.

Shota Takakura, Tetsuro Katayama, Yoshihiro Kita, Hisaaki Yamaba, Kentaro Aburada, Naonobu Okazaki

```
Class calendar
functions
judgeLeapYear: nat -> seq of char
judgeLeapYear(year) ==
  if(year mod 4 = 0 and ((year mod 100 = 0) =>
    (year mod 400 = 0))) then
  "Leap Year"
  else "Normal Year "
end calendar
```

Fig. 1 Example of VDM++ specification

3.1. *Addition of a Function to Generate Test Cases for Enumerated Type*

To solve the problem of the existing BWDM only supporting specific types, we extend Syntax Analyzer and Symbolic Executor. Concretely, we extend BWDM to generate test cases for enumerated type consisting of a set of unique identifiers.

In the extended Syntax Analyzer, if a type definition of an enumerated type exists in the abstract syntax tree analyzing process, the corresponding type declaration is replaced with a list consisting of the declared name and values. If the substituted list exists in the syntax tree in the Symbolic Executor, JavaAPI [4] which generates enumerated type is used to generate constraints to be registered with the SMT solver [5] used in the propositional analyzer section. This additional process can generate input data for enumerated type by symbolic execution.

3.2. *Addition of a Function to Generate Test Cases for Operations/Functions with Recursive Structure*

To solve the problem of existing BWDM's inability to generate test cases for operation and function definitions with recursive structure, we extend Syntax Analyzer and Test Suite Generator.

In existing BWDM, when there is an operation/function call in test suite generation, test cases cannot be generated unless the parsing of the called operation/function is completed. This is the cause that the parsing of the operation/function including recursive calls is not completed and the abstract syntax tree cannot be generated.

Therefore, when there is an operation or function that has self-recursive calls, an upper limit is set on the number of recursive calls, and the abstract parse tree analysis process is modified so that it is executed after the parse tree analysis for all definitions is completed. This allows the test data generation to be performed after the parsing of the VDM++ specification given as input is completed. Additionally, a function is added that terminates the process if self-recursive calls are performed more than a limited number of times when

```
Function Name : judgeLeapYear
Argument Type : current_year : nat
Return Type : seq of (char)
Number of Test Cases : 17 cases(BVA:13 /SE:4)

Boundary Values for Each Argument
current_year : 4294967295 4294967294 0 -1 3 4
5 99 100 101 399 400 401

Test Cases for Boundary Value Analysis
No.1 : 4294967295 -> Undefined Action
No.2 : 4294967294 -> "Normal year"
No.3 : 0 -> "Leap year"
No.4 : -1 -> Undefined Action
No.5 : 3 -> "Normal year"
No.6 : 4 -> "Leap year"
No.7 : 5 -> "Normal year"
No.8 : 99 -> "Normal year"
No.9 : 100 -> "Normal year"
No.10 : 101 -> "Normal year"
No.11 : 399 -> "Normal year"
No.12 : 400 -> "Leap year"
No.13 : 401 -> "Normal year"
```

Fig. 2 Output when Fig.1 is applied to the BWDM

recursively generating output data in the output data generation section in Test Suite Generator.

4. Application Examples

This chapter shows application examples to confirm that the extended BWDM works correctly.

4.1. *Confirmation that Test Cases can be Generated for Enumerated Type*

Fig. 4, shows the VDM++ specification using enumerated type used for verification, and Fig.5 shows the results of applying it to the extended BWDM, respectively.

From Fig. 4, it can see that in the "judgeLightColor" function, "public TraficLight = <BLUE>|<YELLOW>|<RED>;" is defined as the type. As output, Fig.5 shows that test cases can be generated for the definition using the enumerated type.

4.2. *Confirmation that Test Cases can be Generated for Definitions with Recursive Structure*

Fig. 6 shows the VDM++ specification using a definition with recursive structure used for verification, and Fig 7
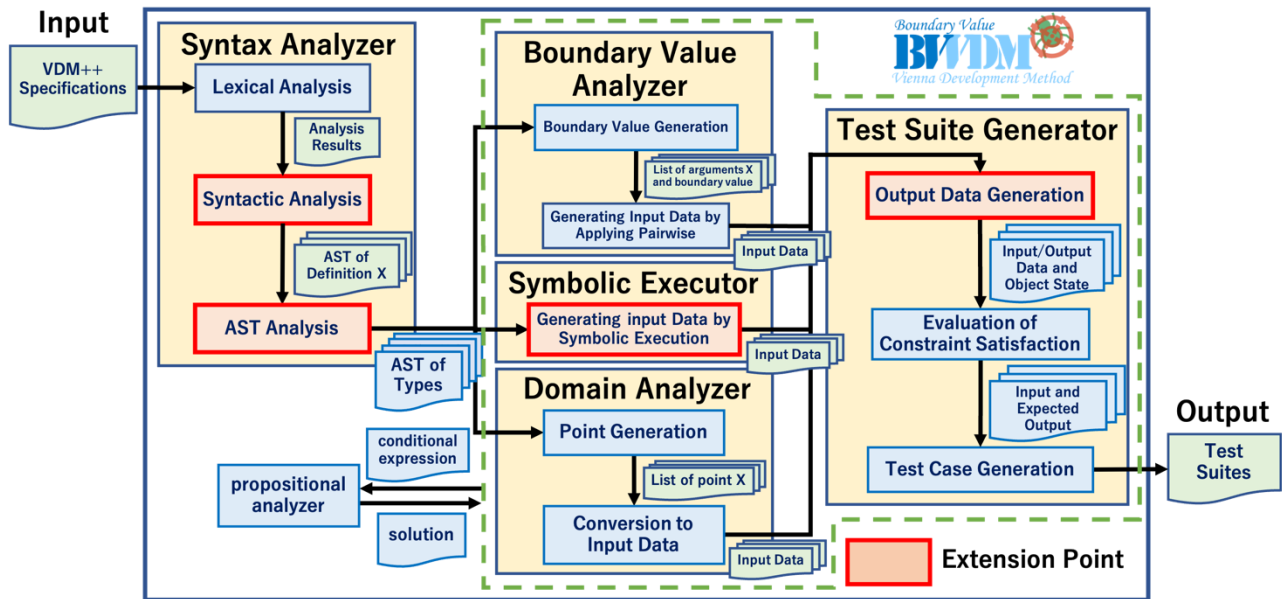
Fig. 3 The structure of the extended BWDM

```
Class judgeLightColor

Types
public TrafficLight = <BLUE> | <YELLOW> | <RED>;
functions
  judgeLightColor: TrafficLight ==> seq of char
    judgeLightCoulor (color) ==
      if color = <BLUE >
        "The color is blue. "
      if color = <YELLOW>
        "The color is yellow"
      if color = <RED>
        "The color is red."

end judgeLightColor
```

Fig. 4 VDM++ specification using enumerated type

```
Function Name : judgeLightColor
Argument Type : color:
Return Type : seq of (char)

Test Cases by Symbolic Execution
No.1 : <BLUE> -> "The color is blue"
No.2 : <YELLOW> -> "The color is yellow"
No.3 : <RED> -> "The clolor is red"
```

Fig. 5 Output when Fig.4 is applied to
the extended BWDM

shows the results of applying it to the extended BWDM, respectively.

From Fig. 6, it can see that in the "calcSum" function is defined with recursive structure. As output, Fig.7 shows that test cases can be generated for definition with recursive structure.

5. Discussion

5.1. *Evaluation of Test Cases Generation for Definitions using Enum Type*

As shown in Section 4.1, we have confirmed that the extended BWDM can generate test cases from definitions using enumerated type in the VDM++ specification of Fig. 4. Therefore, it can be said that this extension enables BWDM to generate test cases from definitions using

enumerated type. Thus, the usefulness of BWDM has been improved.

5.2. *Evaluation of Test Case Generation for Definitions with Recursive Structure*

As shown in Section 4.2, we have confirmed that the extended BWDM can generate test cases from definition with recursive structure in the VDM++ specification of Fig. 6. Therefore, it can be said that this extension enables BWDM to generate test cases from definitions with recursive structure. Thus, the usefulness of BWDM has been improved.

5.3. *Comparison with Manual Test Case Generation Time*

Using the extended BWDM, we evaluate the test case generation time for the VDM++ specifications which include definitions with enumerated type and recursive structure, by comparing it to the human case. Two VDM++ specifications, Fig. 4 and Fig. 6, are used in the

Shota Takakura, Tetsuro Katayama, Yoshihiro Kita, Hisaaki Yamaba, Kentaro Aburada, Naonobu Okazaki

```
Class calcSum

  factorial: nat ==> nat
   factorial (value) ==
     if value = 0 then
       0
     else value + factorial (value − 1)

end calcSum
```

Fig. 6 VDM++ specification with recursive structure

```
Function Name : calcSum
Argument Type : value:nat
Return Type : nat


Boundary Values for Each Argument
value: 4294967295 4294967294 0 −1


Test Cases of Boundary Values
No.1 : 4294967295 −> Undefined Action
No.2 : 4294967294 −> Undefined Action
No.3 : 0 −> 0
No.4 : −1 −> Undefined Action
```

Fig. 7 Output when Fig.6 is applied to
the extended BWDM

experiment. The manual verification is performed by a total of five people, two graduate students and three fourth-year undergraduates in our laboratory. We measure the time required to generate sufficient test cases from the VDM++ specification, which are no leakage. The time measurement is terminated when the subjects have generated the correct test cases, and when the manually test cases includes any incorrect test cases, the mistakes are pointed out. The comparison results are shown in Table.1.

Table 1 shows that about 14 minutes are saved when using the extended BWDM to generate test cases compared to manual test case creation. In addition, human error has been observed in the manual test case creation.

In this research, we have confirmed that an extended BWDM that adds a test case generation function for definitions using enumerated type and recursive structure can reduce the time and human errors required for test case generation. Therefore, it can be said that the usefulness of BWDM has been improved.

Table.1 Comparison of test case generation time

	Time
Average of 10 subjects	14m29s
BWDM	2.6s

6. Conclusion

In order to improve the usefulness of BWDM, an automatic test case generation tool for the VDM++ specification, two extensions have been made to the existing BWDM to solve the problems that only supports integer types and that it cannot generate test cases for operation and function definitions with recursive structure.

The application examples of the extended BWDM are shown, and the test case generation function for definitions using enumerated type and definitions with recursive structure for expanding the range of supported types is confirmed. We have confirmed that the extended BWDM can reduce test case generation time by about 14 minutes compared to manually generating test cases from VDM++ specifications. In addition, we have confirmed that the extended BWDM can eliminate human errors.

Therefore, it can be said that the usefulness of BWDM has been improved by the extensions in this paper.

The future issues are the following.

- Support for types other than integer and enumerated type
- Test case generation for input values with more recursive calls
- Test case generation for mutually recursive functions

References

1. Overture Project. Manuals. http://overturetool.org/documentation/manuals.html. Accessed: 2023-12-13.
2. T. Katayama, F. Hirakoba, Y. Kita, H. Yamaba, K. Aburada, and N. Okazaki. Application of Pirwise Tsting into BWDM which is a Test Case Generation tool for the VDM++ Specification. Journal of Robotics, Networking and Artificial Life, Vol.6, No.3, pp.143-147, 2019.
3. T. Muto, T. Katayama, Y. Kita, H. Yamaba, K. Aburada, and N. Okazaki. Expansion of Application Scope and Addition of a Function for Operations into BWDM which is an Automatic Test Cases Generation Tool for VDM++ Specification. Journal of Robotics, Networking and Artificial, Vol.9. No.3, pp.255-262, 2022.
4. Z3:Packagecom.microsoft.z3. https://z3prover.github.io/api/html/namespacecom_1_1 microsoft_1_1z3.html. Accessed: 2023-12-13.
5. Z3. https://github.com/Z3Prover/z3. Accessed: 2023-12-14.

Authors Introduction

Mr. Shota Takakura

Shota Takakura received the Bachelor's degree in engineering (computer science and systems engineering) from the University of Miyazaki, Japan in 2023. He is currently a Master's student in Graduate School of Engineering at the University of Miyazaki, Japan. His research interests include software testing, software quality, and formal method.

.

Dr. Tetsuro Katayama

He received a Ph.D. degree in engineering from Kyushu University, Fukuoka, Japan, in 1996. From 1996 to 2000, he has been a Research Associate at the Graduate School of Information Science, Nara Institute of Science and Technology, Japan. Since 2000 he has been an Associate Professor at the Faculty of Engineering, Miyazaki University, Japan. He is currently a Professor with the Faculty of Engineering, University of Miyazaki, Japan. His research interests include software testing and quality. He is a member of the IPSJ, IEICE, and JSSST.

Dr. Yoshihiro Kita

Yoshihiro Kita received a Ph.D. degree in systems engineering from the University of Miyazaki, Japan, in 2011. He is currently an Associate Professor with the Faculty of Information Systems, University of Nagasaki, Japan. His research interests include software testing and biometrics authentication.

Dr. Hisaaki Yamaba

He received the B.S. and M.S. degrees in chemical engineering from the Tokyo Institute of Technology, Japan, in 1988 and 1990, respectively, and the Ph D. degree in systems engineering from the University of Miyazaki, Japan in 2011. He is currently an Assistant Professor with the Faculty of Engineering, University of Miyazaki, Japan. His research interests include network security and user authentication. He is a member of SICE and SCEJ.

Dr. Kentaro Aburada

He received the B.S., M.S, and Ph.D. degrees in computer science and system engineering from the University of Miyazaki, Japan, in 2003, 2005, and 2009, respectively. He is currently an Associate Professor with the Faculty of Engineering, University of Miyazaki, Japan. His research interests include computer networks and security. He is a member of IPSJ and IEICE.

Dr. Naonobu Okazaki

He received his B.S, M.S., and Ph.D. degrees in electrical and communication engineering from Tohoku University, Japan, in 1986, 1988 and 1992, respectively. He joined the Information Technology Research and Development Center, Mitsubishi Electric Corporation in 1991. He is currently a Professor with the Faculty of Engineering, University of Miyazaki since 2002. His research interests include mobile network and network security. He is a member of IPSJ, IEICE and IEEE

Proposal of ASLA Which Is a Segmentation and Labeling Tool for Document Images Based on Deep Learning

Kanta Kakinoki*, Tetsuro Katayama*, Yoshihiro Kita†,
Hisaaki Yamaba*, Kentaro Aburada*, and Naonobu Okazaki*

* Department of Computer Science and Systems Engineering, Faculty of Engineering, University of Miyazaki,
1-1 Gakuen-kibanadai nishi, Miyazaki, 889-2192 Japan
†Department of Information Security, Faculty of Information Systems, Siebold Campus, University of Nagasaki
1-1-1 Manabino, Nagayo-cho, Nishi-Sonogi-gun, Nagasaki, 851-2195 Japan
E-mail: kakinoki@earth.cs.miyazaki-u.ac.jp, kat@cs.miyazaki-u.ac.jp, kita@sun.ac.jp,
yamaba@cs.miyazaki-u.ac.jp, aburada@cs.miyazaki-u.ac.jp, oka@cs.miyazaki-u.ac.jp

Abstract

This paper proposes a prototype of ASLA, segmentation and labeling tool for document images based on deep learning, to reduce the time required for region segmentation and label generation. To evaluate the usefulness of ASLA, we have compared the time required for region segmentation and label generation using ASLA and by hand, and then confirmed the reduction in time. We also have confirmed that the rule-based region redividing method achieves a high recall and precision.

Keywords: Region segmentation, Labeling, Document image, Rule-based region redividing

1. Introduction

Because storage space is needed to store paper documents, maintaining and managing their condition are costly. Electronic documents are used to solve this problem [1]. The use of electronic documents enables the reduction of paper and the cost of the maintenance and management. Therefore, the use of electronic documents is becoming more widespread.

The current situation of the electronic documents is only a substitute for paper. As a new way to utilize electronic documents, we focus on dividing electronic documents into regions by their elements and generating keywords and sentences as labels from the contents of the elements. In addition, the movement of the reader's line of sight when they read the electronic document is recorded as coordinate information. By realizing these features, it is possible to improve the efficiency of sales activities by capturing the level of interest and concern of customers, and to strengthen compliance by checking whether important matters are properly explained to subscribers of products or services.

To achieve the above, the following two tasks are required.

- Region segmentation that divides the components of an electronic document into some regions.
- Label generation that analyzes the contents of a region and then generates labels according to the contents of the region.

However, these tasks are time-consuming and labor-intensive in the manual. This paper proposes a prototype of ASLA (Automatic Segmentation and Labeling tool using AI), segmentation and labeling tool for document images based on deep learning, to reduce the time required for region segmentation and label generation.

2. ASLA

2.1. *Functions*

ASLA takes a single column or two columns Japanese document image as input. ASLA outputs segmentation image and class identification on the input image. Table 1 shows the definitions of class and region in ASLA. The output is a segmentation image and an analysis result file. The segmentation image is a JPEG format image. Each region of the image is color-coded and visualized as red for the text class region, blue for the figure class region, and green for the table class region. The analysis result file is a CSV (Comma Separated Value) format file. In that file, the top-left x-coordinate, top-left y-coordinate, bottom-right x-coordinate, bottom-right y-coordinate, class name, and label pair in each region are recorded per line.

Kanta Kakinoki, Tetsuro Katayama, Yoshihiro Kita, Hisaaki Yamaba, Kentaro Aburada, Naonobu Okazaki

Table 1 Definition of classes and regions in ASLA.

Class Name	Target	Definition of regions
Text	Text and title, etc.	Smallest rectangle surrounding each element
Figure	Figure	Smallest rectangle surrounding the figure
Table	Table	Smallest rectangle surrounding the table

2.2. Dataset Preparation

ASLA uses Cascade R-CNN [2] as the object detector for region segmentation. In addition, ASLA uses LayoutLMv3 [3] that is a pre-trained model specific to document AI tasks as a training model for the object detector. To improve the performance of region segmentation for Japanese document images, we fine-tune the model by using a dataset of Japanese document images prepared in advance.

The ASLA prototype uses 103 document images of the doctoral dissertation of the Graduate School of Educational and Developmental Sciences of Nagoya University provided by the Nagoya University Academic Repository [4] and 87 document images of the paper presented at the 2021 Kyushu Branch Joint Conference of the Institutes of Electrical and Related Engineers [5] as document images dataset for fine-tuning.

2.3. Structure

Fig. 1 shows the structure of ASLA. It consists of the following three processing parts.

- Region Segmentation Part:
  Region Segmentation Part at first generates a trained model by fine-tuning LayoutLMv3 with a previously prepared dataset. Next, it divides region by using Cascade R-CNN from the input document image with the generated trained model. This generates the top-left and bottom-right xy-coordinates and class names for each region. Thereafter, it redivides the region to modify the misalignment of regions. The rule-based region redividing is described later in Section 2.4. Finally, based on the results of region segmentation, a segmentation image is output. Since this, text class regions are referred to as text region, figure class regions as figure region, and table class regions as table region, respectively.
- Label Generation Part:
  Label Generation Part generates labels that are all nouns included in the text region for the text region. For the figure and table regions, it generates a label that is the text of each caption. Here, it extracts characters from Japanese text using Tesseract that is an OCR tool, and it generates labels from the extracted characters using MeCab which is a morphological analysis tool.
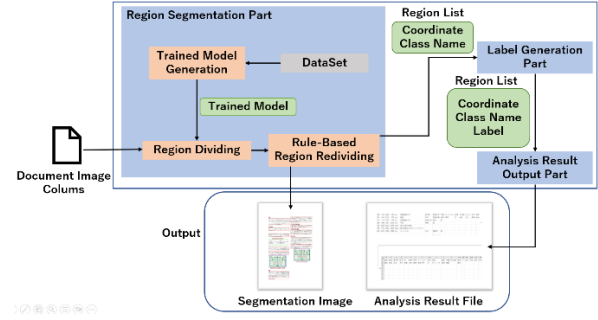
Fig. 1 The structure of ASLA.
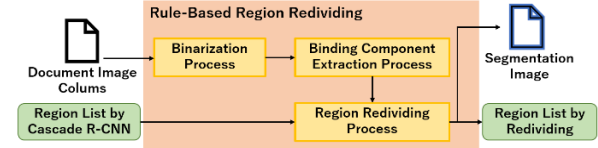
Fig. 2 The Process of rule-based region redividing.

- Analysis Result Output Part:
  Analysis Result Output Part outputs an analysis result file as a CSV format file based on the analysis results obtained by Region Segmentation Part and Label Generation Part.

2.4. Rule-Based Region Redividing

There are several studies that use object detectors to divide regions on document images [3], [6], [7]. The purpose is to roughly divide each region in a document image, such as layout analysis of a document image. Therefore, the existence of small misalignments between the divided regions and the actual regions is not a major problem. However, region segmentation in ASLA extracts the characters within a region after the region is divided. Therefore, if there is even a small misalignment, some characters cannot be extracted. This causes the problem that correct labels cannot be generated. Another problem with segmentation by the object detectors exists that it is difficult to detect small regions [6], [7].

Therefore, ASLA performs rule-based region redividing. This eliminates the misalignment of regions with region segmentation by the object detector, and also divides smaller regions that were not detected before. Fig. 2 shows the flow of the rule-based region redividing. In the region redividing process, each binding component extracted by the binding component extraction process determines which region is the same as the region that was divided using the object detector. Alternatively, binding components that are far from any region are divided as a new region. Fig. 3 shows an example of the region redividing process in rule-based region redividing. In Fig. 3, the region redividing process redivides the region to eliminate misalignment of text region segmented by Cascade R-CNN.
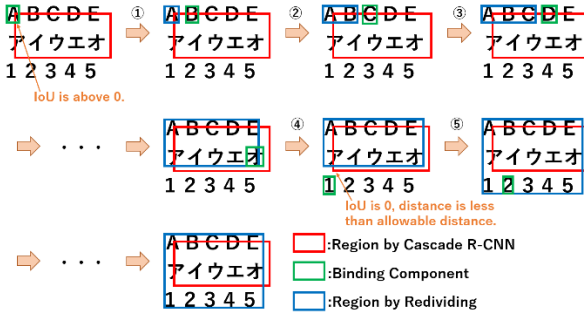
Fig. 3 An example of the region redividing process in rule-based region redividing.
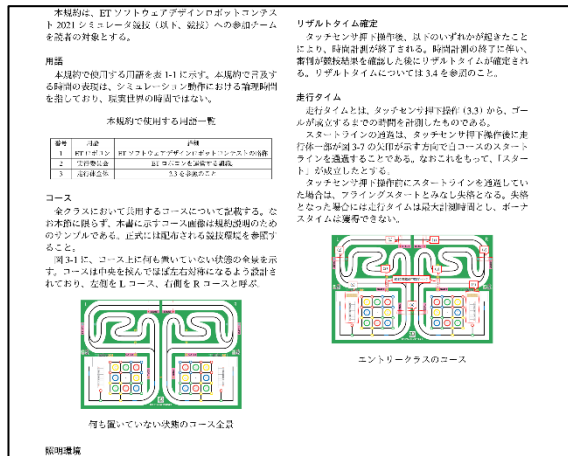
Fig. 4 Input document image sample
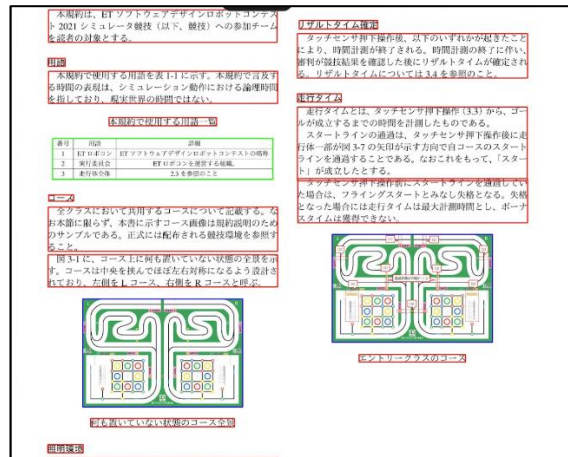
Fig. 5 Segmentation image

2.5. *Input/Output Example*

Fig. 4 shows an example of a document image to be analyzed by ASLA. Fig. 5 shows the segmentation image generated by applying the document image in Fig. 4 as input for ASLA. Fig. 6 shows an image of an analysis result file. From Fig. 5, we have confirmed that ASLA can divide the document image into text, figure, and table regions, as defined in Table 1.

Fig. 6. Result output file image

Table 2. Comparison of the time required for ASLA and manual region segmentation.

	Single column	Two columns	Average
Manual (Average)	2m6s	4m8s	3m7s
ASLA	5s	6s	6s

Table 3. Comparison of the time required for ASLA and manual labeling.

	Time
Manual (Average)	8m40s
ASLA	5s

3. Evaluation

We evaluate the usefulness of ASLA developed in this paper. It is evaluated in terms of the execution time and performance of region segmentation.

3.1. *Evaluation of the Execution Time*

We compare the time required for region segmentation by using ASLA and manually. Two document images are used for region segmentation. They are single column document image and two columns document image. In the manual region segmentation, five subjects divide regions on document images, and their working time is measured. Table 2 shows the comparison of the time required for region segmentation by using ASLA and manually. From Table 2, we have confirmed that ASLA reduces the time required for region segmentation by about 3 minutes (97%) per document image.

We compare the time required for label generation by using ASLA and manually. In the manual label generation, five subjects generate labels on document images that had been divided regions in advance, and their working time is measured. This labeling task involves extracting all nouns within each region. Table 3 shows a comparison of the time required for ASLA and manual label generation. From Table 3, we have confirmed that ASLA reduces the time required for label generation by about 8 minutes (99%) per document image. Based on the above, it can be said that ASLA helps reduce the time required for label generation.

Kanta Kakinoki, Tetsuro Katayama, Yoshihiro Kita, Hisaaki Yamaba, Kentaro Aburada, Naonobu Okazaki

Table 4 Comparison of precision and recall for Cascade R-CNN and Mask R-CNN with and without rule-based region redividing, respectively.

	IoU	Precision			Recall		
		Text	Figure	Table	Text	Figure	Table
Only Mask R-CNN	0.8	0.768	0.846	0.4	0.651	0.846	0.455
	0.9	0.348	0.154	0.2	0.295	0.154	0.182
Mask R-CNN and rule-based region redividing	0.8	0.906	0.929	0.9	0.955	1.0	1.0
	0.9	0.877	0.929	**0.9**	0.924	**1.0**	**1.0**
Only Cascade R-CNN	0.8	0.808	0.846	0.4	0.691	0.918	0.818
	0.9	0.337	0.317	0.2	0.305	0.308	0.182
Cascade R-CNN and rule-based region redividing	0.8	0.945	0.942	0.9	0.955	1.0	1.0
	0.9	**0.936**	**0.942**	**0.9**	**0.955**	**1.0**	**1.0**

3.2. *Evaluation of Region Segmentation Performance*

In addition to ASLA, there are other studies that divide region on document images [6], [7]. These studies use Mask R-CNN [8] as an object detector for region segmentation. Here, we evaluate the performance of region segmentation for each of the Cascade R-CNN used in ASLA and Mask R-CNN. The dataset used for performance evaluation consists of 8 document images. These document images consist of 4 single column document images and 4 two columns document images. The evaluation is based on the calculation of IoU for the regions and class names defined in Table 1 to calculate the precision and recall. Furthermore, to evaluate the usefulness of rule-based region redividing, we calculate the precision and recall with and without rule-based region redividing.

Table 4 shows the calculated precision and recall. From Table 4, we have confirmed that the combination of Cascade R-CNN and rule-based region redividing used in Region Segmentation Part of ASLA has the highest values for both precision and recall. We have confirmed that the usefulness of ASLA for region segmentation. In addition, the rule-based region redividing also has achieved the highest precision and recall when the IoU threshold is set to a high value of 0.9. We have confirmed that the usefulness of rule-based region redividing.

4. Conclusion

This paper has proposed a prototype of ASLA that is a deep learning tool for region segmentation and label generation, in order to reduce the time required to divide document images and generate labels. The proposed ASLA performs region segmentation and class identification on the input document image, generates labels each region, and outputs the segmentation image and the analysis result file.

After applying the document images to ASLA, we have confirmed that ASLA works correctly. Furthermore, we have compared the time required for region segmentation with the use of ASLA and manually. As a result, we have confirmed that the time required for region segmentation per document image could be reduced by about 3 minutes (97%). We also have compared the time required for label generation by using ASLA and manually. As a result, we have confirmed that the time required for label generation per document image could be reduced by about 8 minutes (99%). Furthermore, we have evaluated the performance of region segmentation. As a result, we have confirmed that the rule-based region redividing achieves high precision and recall even with a high IoU threshold value.

The future issues are as follows.

- Support for document images with non-white backgrounds.
- Automatic generation of sentences or nouns that represent the contents of figures or tables.
- Handling of text regions that do not make sense as Japanese.

References

1. Hirohito Shibata, "Paper vs. Electronic Media: Work Efficiency and Environmental Impact" NIP & Digital Fabrication Conference, Vol.27, pp.7-10, 2011.
2. Zhaowei Cai and Nuno Vasconcelos, "Cascade R-CNN: Delving into high quality object detection," Proc. of 2018 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), pp. 6154-6162, 2018.
3. Yupan Huang, Tengchao Lv, Lei Cui, Yutong Lu, and Furu Wei, "LayoutLMv3: Pre training for Document AI with Unified Text and Image Masking", MM '22: Proceedings of the 30th ACM International Conference on Multimedia, pp. 4083-4091, 2022.
4. "Nagoya Repository", https://nagoya.repo.nii.ac.jp/ (Accessed 2023-12-14)
5. "The 74th Joint Conference of Electrical, Electronics and Information Engineers in Kyushu", https://sites.google.com/jceee-kyushu.jp/2021/ (Accessed 2023-12-14)
6. Canhui Xu, Cao Shi, Hengyue Bi, Chuanqi Liu, Yongfeng Yuan, Haoyan Guo, and Yinong Chen: "A page object detection method based on Mask R-CNN", IEEE Access, Vol.9, pp.143448-143456, 2017.
7. Sanket Biswas, Pau Riba, Josep Llados, Umapada Pal, "Beyond document object detection: instance-level segmentation of complex layouts", IJDAR 24, pp.269-281, 2021.
8. Kaiming He, Georgia Gkioxari, Piotr Dollar, and Ross B. Girshick: "Mask R-CNN", International Conference on Computer Vision(ICCV), pp.2961-2969, 2017.

Authors Introduction

Mr. Kanta Kakinoki

He received the Bachelor's degree in engineering (computer science and systems engineering) from the University of Miyazaki, Japan in 2023. He is currently a Master's student in Graduate School of Engineering at the University of Miyazaki, Japan. His research interests include software development support method and image processing.

Dr. Tetsuro Katayama

He received a Ph.D. degree in engineering from Kyushu University, Fukuoka, Japan, in 1996. From 1996 to 2000, he has been a Research Associate at the Graduate School of Information Science, Nara Institute of Science and Technology, Japan. Since 2000 he has been an Associate Professor at the Faculty of Engineering, Miyazaki University, Japan. He is currently a Professor with the Faculty of Engineering, University of Miyazaki, Japan. His research interests include software testing and quality. He is a member of the IPSJ, IEICE, and JSSST.

Yoshihiro Kita

He received a Ph.D. degree in systems engineering from the University of Miyazaki, Japan, in 2011. He is currently an Associate Professor with the Faculty of Information Systems, University of Nagasaki, Japan. His research interests include software testing and biometrics authentication.

Dr. Hisaaki Yamaba

He received the B.S. and M.S. degrees in chemical engineering from the Tokyo Institute of Technology, Japan, in 1988 and 1990, respectively, and the Ph D. degree in systems engineering from the University of Miyazaki, Japan in 2011. He is currently an Assistant Professor with the Faculty of Engineering, University of Miyazaki, Japan. His research interests include network security and user authentication. He is a member of SICE and SCEJ.

Kentaro Aburada

He received the B.S., M.S, and Ph.D. degrees in computer science and system engineering from the University of Miyazaki, Japan, in 2003, 2005, and 2009, respectively. He is currently an Associate Professor with the Faculty of Engineering, University of Miyazaki, Japan. His research interests include computer networks and security. He is a member of IPSJ and IEICE.

Dr. Naonobu Okazaki

He received his B.S, M.S., and Ph.D. degrees in electrical and communication engineering from Tohoku University, Japan, in 1986, 1988 and 1992, respectively. He joined the Information Technology Research and Development Center, Mitsubishi Electric Corporation in 1991. He is currently a Professor with the Faculty of Engineering, University of Miyazaki since 2002. His research interests include mobile network and network security. He is a member of IPSJ, IEICE and IEEE.

A Design of a Modular Mobile Robot for Rescue Operations

Baris Celiker
Institute of Information Technology and Intelligent Systems, Kazan Federal University, Kazan, Russia

Shifa Sulaiman
Laboratory of Intelligent Robotics Systems (LIRS), Institute of Information Technology and Intelligent Systems, Kazan Federal University, Kazan, Russia

Tatyana Tsoy
Laboratory of Intelligent Robotics Systems (LIRS), Institute of Information Technology and Intelligent Systems, Kazan Federal University, Kazan, Russia
Email: Bariss.be@gmail.com, shifa_1910@it.kfu.ru, tt@it.kfu.ru

Abstract

Modular robotics is one of the subfields of mobile robotics, which is emerging as a new trend in various sectors. Modular mobile robots can be reconfigured to perform a wide variety of tasks. In this paper, applications of modular mobile robots in various sectors such as industry, space, surgery, rescue and entertainment tasks are discussed. Based on the study, an improved design of a modular mobile robot for navigating through different terrains during a rescue operation is presented. Simulation study of the robot is included to demonstrate a motion capability of the modular mobile robot.

Keywords: Modular robots, Applications of modular robots, Design, Simulation study

1. Introduction

Robots were known to assist people in a wide variety of tasks such as industrial [1], [2] household [3], rescue [4], [5], medical [6], [7] space [8], defense [9], and entertainment [10] fields. The emergence of artificial intelligence elevated a performance of robots [11]. Most industrial operations are currently performed by robotic systems due to its better performance capabilities compared to human operators. Robots are proved to be more precise and accurate in carrying out a task compared to human operators. Employers are forced to build unmanned workplaces due to a variety of factors such as increasing wages, absence of skilled employees, dangerous work conditions, etc. Hence, most industries are replacing human operators with multi-tasking robotic systems. However, some of the challenging tasks inside industrial and service sectors are nowadays carried out in a collaboration mode. Such robots are known as collaborative robots [12].

Medical, service and space sectors have started to explore the capabilities of mobile robots [13]. The major reason is the multifunctional capabilities of mobile robots especially modular mobile robots. These kinds of robots can be controlled remotely to use in places or environments which are not accessible for a human operator to work. Service sectors are employing modular mobile robots to invigilate the conditions of patients and logistics tasks inside hospitals [14]. Modular robots are also used in space explorations and rescue operations. Ground and aerial modular robots are nowadays employed in places of natural calamities such as earthquakes [15], landslides and [16] and floods [17]. Modular robots can reach any location due to its shape shifting, flexible design and motion capabilities.

2. Applications of mobile and modular robot

Mobile manipulation combines the advantages of manipulation and navigation. Basiri *et al.* [18] developed a mobile robot with a UR5e [19] arm for performing activities in a construction field. The four-wheeled mobile platform was 78 cm long and 55 cm tall. The platform can support a payload of 25 kg. Bricks for construction purposes were identified and moved using a vision-based technique. Lio [20] was a service mobile robot developed as a care assistant to patients admitted in a hospital. It consisted of a 6 degrees-of-freedom (dof) arm and a mobile base with 4 wheels.

Mobile manipulators can be used in cooperative and collaborative tasks to compensate a growing demand for a workforce in industrial sectors. Robots in the industrial sector must follow safety regulations and have been proved to save a significant amount of time and money. Collaborative robots were developed for carrying out

activities combinedly by human operators and robots in a shared configuration space [21]. A team of mobile manipulators used for carrying payload was presented in [22]. Each robot consisted of a 2 dof mobile base and a 2 dof planar arm for transferring an object from one place to another. Another set of mobile manipulators used for a collaborative 3D printing process was illustrated in [23]. Experimental validation proved that using collaborative robots increased a performance of a collaborative task.

Modular robots can perform a variety of tasks and adapt to the environment quickly. Connect-R [24] was a modular robot developed for implementing in an industrial environment. Each module has the ability to rotate around other modules, link to other modules, and extend based on the available space. A reconfigurable modular robot named Deformable Modular Robot (Datom) was presented in [25]. Each module of the robot was connected to other modules using a Face-centered cubic lattice pattern. The robot has a capability to separate from neighbouring modules during a task. FreeSN [26] was a modular reconfigurable robot consisted of strut and node modules. A node module was constructed using a low carbon steel spherical shell and a magnetic connector was used as a joining mechanism. SMORES-EP [27] was a modular robot used in construction sector. It can self-assemble to generate 3 different topologies based on a task. A docking control strategy was used for a successful docking of the robot. A self-reconfigurable mobile robot with 2 dof with a docking mechanism was demonstrated in [28]. It consisted of a hybrid mobile base with wheels and tracks. An image processing technique was used for locating a docking station for the charging of the modular robot.

Modular robots are also used in household and entertainment sectors. Roombots [29] were a group of modular robots designed to employ both as a payload carrier in houses and also act as toys for children. Roombots can act as adaptive and reconfigurable mobile robots and used for supporting furniture during relocation. The reconfigurable modular robot consisted of 12 modules with 36 dof. It can autonomously change shape, grasp, and move furniture. KAIRO-II, a modular robot developed for rescue operations was given in [30]. It consisted of 6 drive modules interconnected using 5 joint modules. KAIRO -II can ascend short stairs and steps shorter than 55 cm. Modular robots are more efficient in carrying out multiple tasks compared to mobile robots. Nowadays modular robots are used for navigating through different terrains using different modes of base motions. Unmanned ground and aerial modular mobile robots are used in rescue applications. These mobile robots can reach a specified location and invigilate the conditions of the environment using sensors attached to them. A design of a modular robot capable of navigating through uneven terrains was proposed in [31]. A hybrid wheel base for the robot was developed with both wheels and tracks. The robot consisted of a vertical translation unit, which was used for moving the robot in uneven terrains. Design of a homogenous modular robot consisted of revolute and prismatic joints was demonstrated in [32]. It consisted of a rotary plate docking mechanism that allows the robot to change configurations during a task.

Robots are used for carrying out surgeries and operations in the medical field. Versatile modular robotic arms are used in robotic surgeries, which can be assembled and dissembled based on task requirements. Robotic capsules are used for carrying out an endoscopy to detect internal diseases. Cameras and biochemical sensors used in these devices can detect internal body issues. Numerous distinct terrestrial modular robotic systems are employed for carrying out space related tasks. However, due to extreme terrestrial conditions, only few robots have been in active employment currently. However, modular robots are preferred for space related tasks since a reconfigurable robot can carry out multiple tasks [33], [34]. Mobile modular robots have been also used in military operations to reach complex terrains [35].

3. Advantages of modular robots

Generally, the cost of a robot manufactured is directly proportional to the size of the robot due to material requirements. Due to the high cost of industrial robots, most of the industries are not able to employ such robots for execution of tasks. However, modular robots are compact and cost-effective in terms of their size and materials.

Capabilities of mobile robots available in current market are not adequate to carry out complex tasks. In comparison to mobile robots, modular robots are more effective and successful when dealing with emergency situations. Modular robots are employed for rescue operations and in narrow pathways, where normal sized mobile robots cannot navigate. Another major capability of a modular robot to transform the shapes and patterns of individual modules.

One of the most important requirements for accomplishing a logistics-related operation is the capability to provide required speed of navigation. The ordered item is anticipated to be delivered as soon as possible. Traditional mobile robots cannot traverse through complex terrains due to mobile base mechanism limitations. However, a modular robot can transform its shape and pattern to control speeds during navigations.

One of the major issues faced by mobile robots is the limited space for navigation. Modular robots are proved to be traversable through narrow and curved pathways without any difficulty. This is mostly due to their shape shifting capabilities and ability to adapt to different

environments. Hence, modular robots can be employed remotely to reach disaster zones for monitoring situations.

4. A design of a mobile robot

A design process of a modular robot is very complex since modular robots must be able to shift shapes based on requirements. Specifications of a modular robot play a crucial role in completing a task successfully. A traditional modular robot has 2 or more modules based on the requirements. Multiple modules can be manipulated for shifting shapes and configuration of the robot. In some circumstances, each module of a modular robot can be operated as a single robot, with no connectivity to other modules.

In this paper, a design of a modular mobile robot, which can be used for search and rescue operations was presented. The specific design (Fig. 1) allowed the robot to enter areas inaccessible to human rescue operators. It can carry necessary equipment, medicines, and has capability to convey information to communicate with online remote operators. A modular robot with two modules each with 4 wheels was designed in Tinkercad software [36], which is significantly more simple in modelling and operating than popular Gazebo [37], [38] or Webots [39], [40] robotic simulators. The design of each module consisted of a mobile robot with 4 wheels. One module with 4 wheels can be moved above the other. An 8-wheeled single-decker vehicle can be created by lining up two distinct vehicles in a row with a rail system in between the wheels. Self-rotating pallets were used for rotating one module and placing it on top of another module. The two modules can also be used separately depending on the task requirements. In such cases, GPS technology is used to detect positions of each module and reach a location for docking.
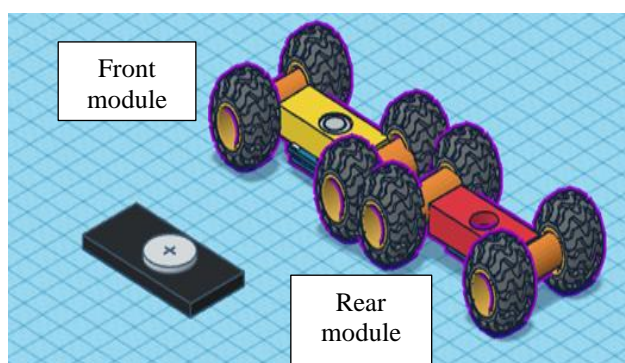
Fig. 1 Modular mobile robot with 2 modules

Each module of the robot consisted of a hydrogen fuel battery, a Raspberry PI 4 as processor and a WIFI module (Fig, 2). The batteries presented in the robot were charged using a coil. Raspberry PI 4 was used to control an overall process of the modular mobile robot. WIFI module was used for receiving and sending information. The robot was charged using a wireless charging method also. A

rectifier and a filter were equipped inside the robot to manipulate incoming and outgoing currents. Magnetic resonance technique can also be used for charging the robot.
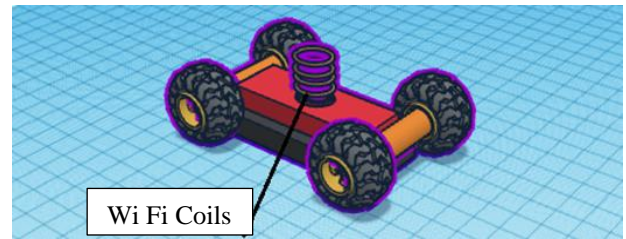
Fig. 2 WIFI equipped on modular robot

5. Simulation study

A simulation study was conducted for analysing motion capabilities of the modular robot. A task was assigned to the robot to reach a goal location from an initial location in a stacked configuration. The initial configuration of the modular robot is shown in Fig. 1. The rear module of the modular robot was sliding on to the top of front module using the wheel rotations as shown in Fig. 3. The final configuration of the robot when rear module reached the top of front module is shown in Fig. 4. The rear module can revert back to its initial pose using the same motion shown in the Fig. 4. The designed modular robot can be traversed using 8 wheels or 4 wheels by keeping rear module on the top based on the space constraints.
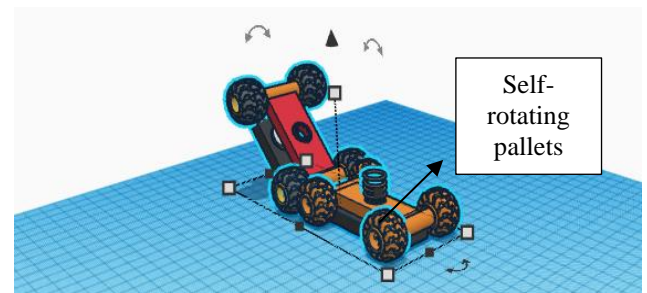
Fig. 3 Intermediate position of rear and front modules
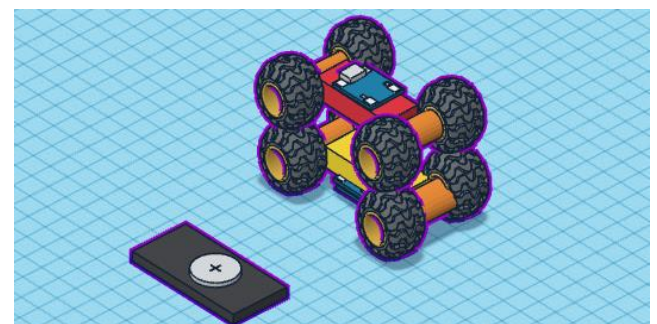
Fig. 4 Stacked configuration of the modular robot

The modular robot was able to reach the goal location in a stacked configuration. The motion of the robot was smooth on a flat terrain.

6. Conclusion

In this paper, applications and advantages of available designs of modular robots in various sectors were reviewed. Based on the study, a design of a modular robot with two modules was introduced. The modules can traverse individually or in combined pattern based on the space constraints and assigned tasks. A simulation study was carried out to analyze a motion capability of the designed modular robot. The nature of motion obtained was smooth and able to complete the task within limited time period. The motion capabilities of the designed modular robot in complex terrains and 3D spaces will be analyzed in future works. The energy capacity of the modular robot will be analyzed during a task for determining the operation runtime of the robot.

Acknowledgements

References

1. M. Mstafin, E. Chebotareva, E. Magid, H. Li and E. Martinez-Garcia, Features of Interaction Between a Human and a Gestures-Controlled Collaborative Robot in an Assembly Task: Pilot Experiments, International Conference on Artificial Life and Robotics (ICAROB 2023), 2023, pp. 158-162.\

2. R. Sultanov, S. Sulaiman, H. Li, R. Meshcheryakov and E. Magid, A Review on Collaborative Robots in Industrial and Service Sectors, in 2022 International Siberian Conference on Control and Communications (SIBCON), IEEE, 2022, pp. 1-7.

3. J. Zhong, C. Ling, A. Cangelosi, A. Lotfi and X. Liu, On the gap between domestic robotic applications and computational intelligence, Electronics, vol. 10(7), 2021, p. 793.

4. E. Magid, Simulation Tools for Urban Search and Rescue Robotics, International Conference on Artificial Life and Robotics (ICAROB 2023), 2023, pp. 4-11.

5. A. Dobrokvashina, S. Sulaiman, T. Gamberov, K.-H. Hsia and E. Magid, New Features Implementation for Servosila Engineer Model in Gazebo Simulator for ROS Noetic, in Proceedings of the International Conference on Artificial Life Science and Technology, vol. 28, 2023, pp. 153-156.

6. B. Abbyasov, A. Dobrokvashina, R. Lavrenov, E. Kharisova, T. Tsoy, L. Gavrilova, S. Bulatov, E. Maslak, N. Schiefermeier-Mach and E. Magid, Ultrasound sensor modeling in Gazebo simulator for diagnostics of abdomen pathologies, The 15th Siberian Conference on Control and Communications (SIBCON 2021), 2021, № 9438910.

7. E. Magid, A. Zakiev, T. Tsoy, R. Lavrenov and A. Rizvanov, Automating pandemic mitigation, Advanced Robotics, vol. 35(9), 2021, pp. 572-589.

8. Z. Jiang, X. Cao, X. Huang, H. Li and M. Ceccarelli, Progress and development trend of space intelligent robot technology, Space: Science & Technology, 2022.

9. M. Payal, P. Dixit, T. V. M. Sairam and N. Goyal, Robotics, AI, and the IoT in defense systems, AI and IoT-Based Intelligent Automation in Robotics, 2021, pp. 109-128.

10. K. J. Morris, V. Samonin, J. Baltes, J. Anderson and M. C. Lau, A robust interactive entertainment robot for robot magic performances, Applied Intelligence, vol. 49, 2019, pp. 3834-3844.

11. E. Magid, A. Pashkin, N. Simakov, B. Abbyasov, J. Suthakorn, M. Svinin, M and F. Matsuno, Artificial intelligence based framework for robotic search and rescue operations conducted jointly by international teams. Smart Innovation, Systems and Technologies, vol. 154, 2019, pp. 15-26.

12. A. Dobrokvashina, A. Zagirov, E. Chebotareva, K.-H. Hsia, E. Magid and S. Sulaiman, Human robot interaction in collaborative manufacturing scenarios: prospective cases. Siberian Conference on Control and Communications, (SIBCON 2022), 2022, pp. 1-6.

13. F. Rubio, F. Valero, and C. Llopis-Albert, A review of mobile robots: Concepts, methods, theoretical framework, and applications, International Journal of Advanced Robotic Systems, vol. 16(2), 2019.

14. R. Safin, R. Lavrenov, T. Tsoy, E. Magid, M. Svinin, S. D. Roy and S. K. Saha, Prioritizing tasks within a robotic transportation system for a smart hospital environment, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 12998, 2021, pp. 182-193.

15. E. Magid, K. Ozawa, T. Tsubouchi, E. Koyanagi and T. Yoshida, Rescue Robot Navigation: Static Stability Estimation in Random Step Environment, Lecture Notes In Computer Science, vol. 5325, 2008, pp. 305-316.

16. E. Magid, F. Matsuno, J. Suthakorn, M. Svinin, Y. Bai, T. Tsoy, R. Safin, R. Lavrenov, A. Zakiev, H. Nakanishi, M. Hatayama and T. Endo, E-ASIA Joint Research Program: development of an international collaborative informational system for emergency situations management of flood and land slide disaster areas, Artificial Life and Robotics, vol. 27(4), 2022, pp. 613-623.

17. Y. Bai, M. Svinin and E. Magid, Multi-Robot Control for Adaptive Caging and Tracking of a Flood Area, 59th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE), 2020, pp. 1452-1457.

18. M. Basiri, J. Gonçalves, J. Rosa, A. Vale and P. Lima, An autonomous mobile manipulator to build outdoor structures consisting of heterogeneous brick patterns, SN Applied Sciences, vol. 3, 2021, pp. 1-14.

19. R. Sultanov, S. Sulaiman, T. Tsoy and E. Chebotareva, Virtual Collaborative Cells Modelling for UR3 and UR5 Robots in Gazebo Simulator, International Conference on Artificial Life and Robotics (ICAROB 2023), 2023, pp. 150-153.

20. J. Mišeikis, P. Caroni, P. Duchamp, A. Gasser, R. Marko, N. Mišeikienė and H. Früh, Lio-a personal robot assistant for human-robot interaction and care applications, IEEE Robotics and Automation Letters, vol. 5(4), 2020, pp. 5339-5346.

21. R. Galin, A. Shiroky, E. Magid, R. Meshcheryakov and M. Mamchenko, Effective Functioning of a Mixed Heterogeneous Team in a Collaborative Robotic System. Informatics and Automation (SPIIRAS Proceedings), vol. 20(b), 2022, pp. 1224-1253.

22. C. P. Tang, R. Bhatt and V. Krovi, Decentralized kinematic control of payload by a system of mobile manipulators,

in IEEE International Conference on Robotics and Automation, vol. 3, April 2004, pp. 2462-2467.

23. X. Zhang, M. Li, J. H. Lim, Y. Weng, Y. W. D. Tay, H. Pham and Q. C. Pham, Large-scale 3D printing by a team of mobile robots, Automation in Construction, vol. 95, 2018, pp. 98-106.

24. M. E. Sayed, J. O. Roberts, K. Donaldson, S. T. Mahon, F. Iqbal, B. Li, S. F. Aixela, G. Mastorakis, E. T. Jonasson, M P. Nemitz, S. Bernardini and A. A. Stokes, Modular robots for enabling operations in unstructured extreme environments, Advanced Intelligent Systems, vol. 4(5), 2022, pp. 2000227.

25. B. Piranda and J. Bourgeois, Datom: A Deformable modular robot for building self-reconfigurable programmable matter, in Distributed Autonomous Robotic Systems: 15th International Symposium, 2022, pp. 70-81.

26. Y. Tu, G. Liang and T. L. Lam, FreeSN: A freeform strut-node structured modular self-reconfigurable robot-design and implementation, International Conference on Robotics and Automation (ICRA), 2022, pp. 4239-4245.

27. C. Liu, Q. Lin, H. Kim and M. Yim, SMORES-EP, a modular robot with parallel self-assembly, Autonomous Robots, vol. 47(2), 2023, pp. 211-228.

28. S. S. Sohal, B. Sebastian and P. Ben-Tzvi, Autonomous Docking of Hybrid-Wheeled Modular Robots With an Integrated Active Genderless Docking Mechanism, Journal of Mechanisms and Robotics, vol. 14(1), 2022, p. 011010.

29. S. Hauser, M. Mutlu, P. A. Léziart, H. Khodr, A. Bernardino and A. J. Ijspeert, Roombots extended: Challenges in the next generation of self-reconfigurable modular robots and their application in adaptive and assistive furniture, Robotics and Autonomous Systems, vol. 127, 2020, p. 103467.

30. L. Pfotzer, S. Ruehl, G. Heppner, A. Roennau and R. Dillmann, KAIRO 3: A modular reconfigurable robot for search and rescue field missions, IEEE Int. Conf. on Robotics and Biomimetics, 2014, pp. 205-210.

31. P. Kumar, W. Saab and P. Ben-Tzvi, Design of a multi-directional hybrid-locomotion modular robot with feedforward stability control, in International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, August 2017, vol. 58189, pp. V05BT08A010.

32. J. Baca, S. G. M. Hossain, P. Dasgupta, C. A. Nelson and A. Dutta, Modred: Hardware design and reconfiguration planning for a high dexterity modular self-reconfigurable robot for extra-terrestrial exploration, Robotics and Autonomous Systems, vol. 62(7), 2014, pp. 1002-1015.

33. SpaceArchitect.org, 3D printed Mars habitat (3rd phase of NASA's centennial challenge), 2018, http://spacearchitect.org/portfolio-item/ 3d-printed-mars-habitat-3rd-phase-of-nasas-centennial-challenge/.

34. M. A. Post and J. Austin, Knowledge-based self-reconfiguration and self-aware demonstration for modular satellite assembly, in 10th International Workshop on Satellite Constellations & Formation Flying 2019, April 2019.

35. J. Whitman, R. Bhirangi, M. Travers and H. Choset, Modular robot design synthesis with deep reinforcement learning, in Proc. of the AAAI Conference on Artificial Intelligence, vol. 34(6), April 2020, pp. 10418-10425.

36. E. A. Juanda and F. Khairullah, Tinkercad Application Software to Optimize Teaching and Learning Process in Electronics and Microprocessors Subject, in 6th UPI International Conference on TVET 2020, 2021, pp.124-128.

37. A. Dobrokvashina, R. Lavrenov, E. Magid, Y. Bai and M. Svinin, How to Create a New Model of a Mobile Robot in ROS/Gazebo Environment: An Extended Tutorial, International Journal of Mechanical Engineering and Robotics Research, vol. 12(4), 2023, pp. 192-199.

38. D. Imameev, A. Zakiev, H. Li, E. Martinez-Garcia and E. Magid, Modelling Autonomous Perpendicular Parking Procedure for Car-like Robot Avrora Unior in Gazebo Simulator, 6th International Scientific Conference on Information, Control, and Communication Technologies (ICCT 2022), 2022, pp. 412-415.

39. A. Dobrokvashina, R. Lavrenov, E. Magid, Y. Bai, M. Svinin and R. Meshcheryakov, Servosila Engineer Crawler Robot Modelling in Webots Simulator, International Journal of Mechanical Engineering and Robotics Research, vol. 11(6), 2022, pp. 417-421.

40. A. Dobrokvashina, R. Lavrenov, Y. Bai, M. Svinin and E. Magid, Sensors modelling for Servosila Engineer crawler robot in Webots simulator, Moscow Workshop on Electronic and Networking Technologies (MWENT-2022), 2022, pp. 1-5.

Authors Biographies

Mr. Baris Celiker
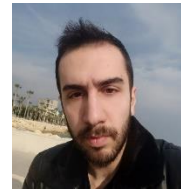
In 2021, he graduated from the Universty of Dokuz Eylul, Turkey. Since 2022 he has been a master's student at the Institute of Information Technology and Intelligent Systems (ITIS), Kazan Federal University (KFU).

Dr. Shifa Sulaiman
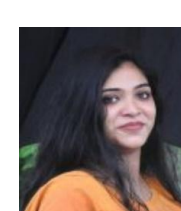
In 2013, she received her Master's degree in Machine Design from Mahatma Gandhi University, India. In 2013-2017 worked as an Assistant Professor in various Indian Engineering Institutes. She earned her PhD from National Institute of Technology, Calicut, India, specialized in Humanoid Robotics in 2023. She is currently working as a Research Associate at the Laboratory of Intelligent Robotic Systems (LIRS) at ITIS, KFU, Russia.

Ms. Tatyana Tsoy

In 2012 she graduated from International Area Studies Master's Degree Program at the University of Tsukuba. In 2023 completed a PhD in Computer science and information processes at ITIS, KFU, Russia.

Implementation of Bug1 and Bug2 Basic Path-Planning Algorithms for a TurtleBot 3 Robot in ROS Noetic

Ilya Spektor
HSE University, Moscow, Russian Federation

Aidar Zagirov
Intelligent Robotics Department, Kazan Federal University, 420008, Kazan, Russian Federation

Ramil Safin
Intelligent Robotics Department, Kazan Federal University, 420008, Kazan, Russian Federation

Evgeni Magid
Intelligent Robotics Department, Kazan Federal University, 420008, Kazan, Russian Federation
HSE University, Moscow, Russian Federation
E-mail: ivspektor@edu.hse.ru, ai.zag@it.kfu.ru, safin.ramil@it.kfu.ru, magid@it.kfu.ru

Abstract

Mobile robots typically operate in a dynamically changing unknown environments. The Bug family algorithms were developed to address path planning challenges for ground vehicles within 2D configuration spaces of unknown environments. These algorithms utilize local sensory data about obstacles encountered during navigation. This paper introduces the implementation of Bug1 and Bug2 local path planning algorithms in ROS Noetic. Traditional 2D Bug algorithms are extended into 3D Gazebo simulation environment. Performance evaluation of Bug1 and Bug2 were conducted using the TurtleBot3 Burger model in both simple convex and maze environments.

Keywords: BUG Algorithms, Path Planning, Navigation, Mobile Robots, ROS, Gazebo

1. Introduction

There are two main path planning models distinguished by the type of information they utilize. The first model, known as path planning with complete information or global navigation, assumes that a robot possesses comprehensive information about its surroundings. In contrast, the second model, known as path planning with incomplete information, introduces uncertainty about an environment [1]. It is essential to integrate both global and local path planning strategies to ensure secure robotic navigation. In this context, a primary objective of local planners is to adhere to an initial plan provided by global planners [2]. Thus, an implementation of a safe and reliable local navigation algorithm, tested in conditions close to real-world scenarios, holds a significant importance.

The Boundary-following and Ultimate Goal (BUG) family of algorithms is among the most renowned in the field of local path-planning. The BUG algorithms follow a simple cycle consisting of two states. In the first state, a robot pursues a directly or intricately calculated path to a target point until encountering an obstacle. The second step involves circumnavigating the obstacle until an exit condition is met.

This paper outlines challenges and potential solutions for implementing local path-planning algorithms within 3D environments, utilizing Bug1 and Bug2 as reference models. It marks an initial step toward an implementation of more complex BUG algorithms [3], [4]. The algorithms were implemented in ROS and evaluated in Gazebo simulator using TurtleBot3 Burger [5]. Custom-designed environments were used for testing while as a part of our ongoing work the algorithms are being validated employing a broad set of environments generated by an automatic tool for Gazebo simulator [6] and a number of standard warehouses' models [7].

2. Algorithms Overview

The Bug1 and Bug2 algorithms are a foundation for the entire BUG family, encompassing algorithms such as Class 1, 2 and 3 [8], Alg1 and Alg2 [9], [10], VisBug21 and VisBug22 [11], Wedgebug [12], Rev1 and Rev2 [13],

Ilya Spektor, Aidar Zagirov, Ramil Safin, Evgeni Magid

Distbug [14], TangentBug [15], CautiousBug [16] and others. Both algorithms operate by transitioning between two states: move-to-target and boundary-following (Fig. 1). These algorithms require a robot to be equipped with a touch sensor, rendering them useful for compact machinery. Next, we offer a brief overview of Bug1 and Bug2 algorithms. For a more in-depth analysis refer to the original work of Lumelsky and Stepanov [1].
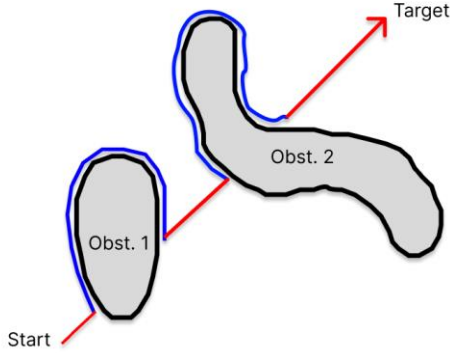
Fig. 1. Trajectory of Bug2 algorithm. Red lines depict the move-to-target state, while blue lines illustrate the boundary-following state.

We use the following notations for both algorithms:

- *Target* – a destination point that the robot should eventually reach.
- L_i (leave point) – a point defined when departing from an *i*-th obstacle to resume moving toward *Target*.
- H_i (hit point) – a point defined when an *i*-th obstacle is encountered.

2.1. Bug1 Algorithm

For the Bug1 algorithm we introduce variable Q_m that stores an encountered point with a minimal distance between an obstacle boundary and *Target*. Additionally, we will introduce three registers: R_1 to store coordinates of the point Q_m, R_2 for a length of the obstacle boundary from H_i to Q_m, and R_3 for a length of the obstacle boundary from Q_m to H_i. The following steps are executed at every point of a continuous path:

1) From point L_{i-1} move toward *Target* along a straight line until one of the following events occurs:
    (a) *Target* is reached: stop the algorithm.
    (b) An obstacle is encountered: define hit point H_i and proceed to step 2.
2) Follow the obstacle boundary using a local direction (local information about the obstacle's boundary):
    (a) If *Target* is reached: stop the algorithm.
    (b) Otherwise, traverse the boundary and return to H_i. Define new leave point $L_i = Q_m$ and proceed to step 3.
3) Apply a test for target reachability. If *Target* is not reachable, stop the algorithm. Otherwise, using

content of registers R_2 and R_3, determine a shorter way along the boundary to L_i, use it to get to L_i. Increment counter *i* and go back to step 1.

2.2. Bug2 Algorithm

Steps of Bug2 are executed at any point of a continuous path as follows:

1) From point L_{j-1}, move along a straight line (*Start, Target*) until one of the following events occurs:
    (a) If *Target* is reached: stop the algorithm.
    (b) An obstacle is encountered: define hit point H_j and proceed to step 2.
2) Follow an obstacle boundary:
    (a) If *Target* is reached: stop the algorithm.
    (b) If line (*Start, Target*) is met at point Q such that distance $d(Q) < d(H_j)$, and line $(Q, Target)$ does not cross the current obstacle at point Q. Define leave point $L_j = Q$. Increment counter *j* and go back to step 1.
    (c) If the robot returns to H_j and, hence, completes a closed curve (the obstacle boundary) without defining next hit point H_{j+1}, it means that *Target* cannot be reached, stop the algorithm.

3. Implementation

Several challenges were encountered during Bug1 and Bug2 algorithms' implementation in ROS/Gazebo. The first problem was a simulation of a touch sensor for the TurtleBot3 Burger robot. It is equipped with a 180 degrees field of view range sensor for safety and integrity reasons [5]. To simulate a touch sensor, we restricted the sensor's vision range to $\rho_v = 1.5R_b$, where R_b is a radius of a robot's body circumcircle. Touch sensors usually provide boolean values: *true* means that a robot is touching an obstacle, *false* otherwise. In our implementation we divided a field of view into five quadrants (of 36 degrees each) and register collision events within each quadrant independently. Another problem is that Bug1 and Bug2 do not limit robot's forward and angular velocities [1]. The most straightforward solution is to set a constant velocity. Potential further enhancements may incorporate a PID controller into the algorithm's pipeline.

Our implementation and validation were conducted using Ubuntu OS, employing robust capabilities of Robot Operating System (ROS) Noetic Ninjemys and the Gazebo simulator [17], [18]. The system configuration for the rigorous testing and validation is outlined in Table 1. For testing purposes, we selected the TurtleBot3 Burger (see Fig. 2) as the robot model [19]. The algorithms were initially implemented in Python, leveraging its popularity and the ease of prototyping and development. However, recognizing the importance of computational efficiency, we plan to reimplement the algorithms in C++.

Table 1. System configuration

Component	Specification
OS	Ubuntu 20.04 LTS
ROS	Noetic Ninjemys
Gazebo	11.11.0
Processor	Intel Core i5-11300H @ 3.10GHz
RAM	24 GB
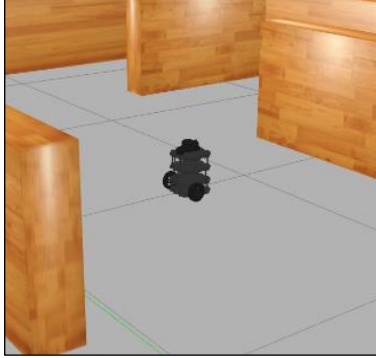
Fig. 2. TurtleBot3 Burger in the Gazebo environment.

There are three ROS nodes created to simulate a cycle of switching the states:

1) **Main**: operates continuously, responsible for calculating an optimal time to switch between the states based on predefined conditions.
2) **Go-to-Target**: orients and moves the robot toward the target.
3) **Follow-Wall**: facilitates the circumnavigation of obstacles. In the case of Bug1, there are two instances of this node – one for clockwise rotations and another for counterclockwise rotations.

4. Validation

To ensure the robustness and effectiveness of the Bug1 and Bug2 algorithm implementations, a comprehensive validation process was undertaken in Gazebo simulator using ROS navigation stack [19], [20]. The validation encompassed a diverse set of tests conducted in 3D simulation world featuring individual complex obstacles and a maze [6]. To ascertain the correctness of the target point reachability determination, a testing scenario was implemented for both Bug1 and Bug2 algorithms. The validation process involved 30 tests for each algorithm on maps with convex obstacles. Half of these tests featured reachable targets, while the remaining half featured unreachable targets.

Given the Bug1's requirement to fully circumnavigate obstacles, testing in maze environments was time-consuming. Thus, only 10 tests for each algorithm were conducted on the maze map, with an equal distribution of reachable and unreachable target scenarios. All target points were accurately identified for reachability in every

test scenario. When a point was deemed reachable, the robot successfully navigated its path to the target, affirming the efficacy of the Bug1 and Bug2 algorithms.

The outcomes of the Bug1 and Bug2 algorithms were visually demonstrated through simulations in Gazebo. Fig. 3 illustrates a resulting path in an environment with convex obstacles, achieved by the Bug1 algorithm. Fig. 4 presents a resulting path in the same environment by Bug2. Notably, a difference was significant, with Bug2 path displaying a more straightforward trajectory, which is observed in a substantial portion of the conducted tests.
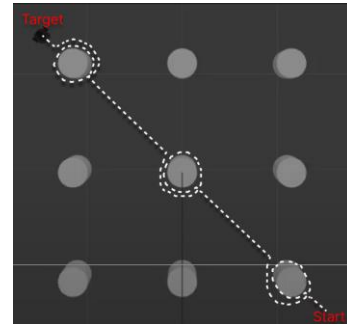
Fig. 3. Path of Bug1 in the environment with convex obstacles (white dashed line).
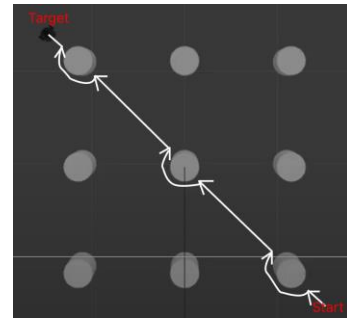
Fig. 4. Path of Bug2 in the environment with convex obstacles (white line).

5. Conclusion

In this work we presented the implementation of the Bug1 and Bug2 algorithms, delving into the challenges encountered while integrating local navigation algorithms into realistic ROS/Gazebo simulations. This article lays a ground for future implementations of more complex BUG family algorithms.

Acknowledgements

References

1. V. J. Lumelsky and A. A. Stepanov, Path-planning strategies for a point mobile automaton moving amidst

unknown obstacles of arbitrary shape, In Autonomous robot vehicles, Vol. 2, 1990, pp. 403-430.

2. A. Apurin, B. Abbyasov, E. A. Martínez-García and E. Magid, Comparison of ROS Local Planners for a Holonomic Robot in Gazebo Simulator, In International Conference on Interactive Collaborative Robotics, Cham: Springer Nature Switzerland, 2023, pp. 116-126.

3. J. Ng and T. Bräunl, A Practical Comparison of Robot Path Planning Algorithms given only Local Information, Centre for Intelligent Information Processing Systems School of Electrical and Electronic Engineering The University of Western Australia, 2005.

4. J. Ng and T. Bräunl, Performance comparison of bug navigation algorithms, Journal of Intelligent and Robotic Systems, Vol. 50(1), 2007, pp. 73-84.

5. R. Safin, T. Tsoy, R. Lavrenov, I. Afanasyev and E. Magid, Modern Methods of Map Construction Using Optical Sensors Fusion, International Conference on Artificial Life and Robotics (ICAROB 2023), 2023, pp. 166-169.

6. A. Iskhakova, B. Abbyasov, T. Tsoy, T. Mironchuk, M. Svinin and E. Magid, LIRS-Mazegen: An Easy-to-Use Blender Extension for Modeling Maze-Like Environments for Gazebo Simulator, Frontiers in Robotics and Electromechanics, Vol. 329, 2023, pp. 147-161.

7. A. Khazetdinov, A. Zakiev, T. Tsoy, R. Lavrenov and K.-H. Hsia, Standard-complaint Gazebo warehouse modelling and validation, Proceedings of 13th International Conference on Developments in eSystems Engineering (DeSE), 2020, p. 218-221.

8. H. Noborio, Several path-planning algorithms of a mobile robot for an uncertain workspace and their evaluation, Proceedings of the IEEE International Workshop on Intelligent Motion Control, Istanbul, Turkey, 1990, pp.289-294.

9. A. Sankaranarayanan and M. Vidyasagar, A New Path Planning Algorithm For Moving A Point Object Amidst Unknown Obstacles In A Plane, IEEE Conference on Robotics and Automation, 1990, pp. 1930-1936.

10. A. Sankaranarayanan and M. Vidyasagar, Path Planning For Moving A Point Object Amidst Unknown Obstacles In A Plane: A New Algorithm And A General Theory For Algorithm Development, Proceedings of the 29th Conference on Decision and Control, 1991, pp. 1111-1119.

11. V. Lumelsky and T. Skewis, Incorporating range sensing in the robot navigation function, IEEE Transactions on Systems, Man, and Cybernetics, Vol. 20(5), 1990, pp. 1058-1069.

12. S. L. Laubach and J. W. Burdick, An autonomous sensor-based path-planner for planetary microrovers, In Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No. 99CH36288C), Vol. 1, 1999, pp. 347-354.

13. Y. Horiuchi and H. Noborio, 2001, May. Evaluation of path length made in sensor-based path-planning with the alternative following, In Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation (Cat. No. 01CH37164), Vol. 2, 2001, pp. 1728-1735.

14. I. Kamon and E. Rivlin, Sensory-based motion planning with global proofs, IEEE Transactions on Robotics and Automation, Vol. 13(6), 1997, pp. 814-822.

15. I. Kamon, E. Rimon and E. Rivlin, TangentBug: A Range-Sensor-Based Navigation Algorithm. The International Journal of Robotics Research, Vol. 17(9), 1998, pp. 934-953.

16. E. Magid and E. Rivlin, CAUTIOUSBUG: A Competitive Algorithm for Sensory-Based Robot Navigation.

Proceedings of 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2004, pp. 2757-2762.

17. R. Sultanov, S. Sulaiman, T. Tsoy and E. Chebotareva, Virtual Collaborative Cells Modelling for UR3 and UR5 Robots in Gazebo Simulator, International Conference on Artificial Life and Robotics (ICAROB 2023), Vol. 28, 2023, pp. 149-152.

18. A. Apurin, A. Dobrokvashina, B. Abbyasov, T. Tsoy, E. A. Martinez-Garcia and E. Magid, LIRS-ArtBul: Design, modelling and construction of an omnidirectional chassis for a modular multipurpose robotic platform, Lecture Notes in Computer Science, Vol. 13719, 2022, p. 70-80.

19. F. U. Pereira, P. M. de A. Brasil, Marco Antonio De Souza Leite Cuadros and Anselmo Rafael Cukla, Analysis of Local Trajectory Planners for Mobile Robot with Robot Operating System, IEEE Latin America Transactions, Vol. 20(1), 2022, pp. 92-99.

20. A. Dobrokvashina, R. Lavrenov, T. Tsoy, E. A. Martinez-Garcia, Y. Bai, Navigation stack for the crawler robot Servosila Engineer, Proceedings of the IEEE 16th Conference on Industrial Electronics and Applications (ICIEA), 2021, p. 1907-1912.

Authors Introduction

Mr. Spektor Ilya

Spektor Ilya is a fourth-year bachelor student in Applied Mathematics' at Tikhonov Moscow Institute of Electronics and Mathematics, HSE University, Russian Federation.

Mr. Aidar Zagirov

Aidar Zagirov is a master student in intelligent robotics at the Institute of Information Technology and Intelligent Systems, Kazan Federal University (KFU), Russia. He works as a research assistant at the Laboratory of Intelligent Robotic Systems at KFU.

Mr. Ramil Safin

Ramil Safin is a Ph.D. student and a research assistant at the Laboratory of Intelligent Robotic Systems, Institute of Information Technology and Intelligent Systems, Kazan Federal University, Russia. He teaches «Fundamentals of technical vision», «Robotic systems sensors».

Professor Evgeni Magid
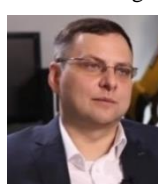
A Professor, a Head of Intelligent Robotics Department and a Head of Laboratory of Intelligent Robotic Systems (LIRS) at Kazan Federal University, Russia. Professor at HSE University, Russia. Senior IEEE member. He earned his Ph.D. degree from University of Tsukuba, Japan. He authors over 200 publications.

Implementation of Alg1 and Alg2 Path Planning Algorithms for Mobile Robots Using ROS Noetic

Anastasia Yankova
HSE University, Moscow, Russian Federation

Timur Gamberov
Intelligent Robotics Department, Kazan Federal University, 420008, Kazan, Russian Federation

Tatyana Tsoy
Intelligent Robotics Department, Kazan Federal University, 420008, Kazan, Russian Federation
E-mail: aayankova@edu.hse.ru, TRGamberov@stud.kpfu.ru, tt@it.kfu.ru

Abstract

Two standard approaches for a robot path planning include a global and a local navigation. The later does not require to store an environment model in a robot memory. This paper presents implementations of two local navigation algorithms, Alg1 and Alg2, with a robot having no prior information about an environment and obstacles. It calculates a path in a real time, continuously changing its states depending on correspondent conditions. The algorithms were implemented for an existing differential drive robot Turtlebot3 Burger using Robot Operation System (ROS). Virtual experiments were performed in the Gazebo simulator employing a simple 3D environment with only convex obstacles and a small 3D maze.

Keywords: path planning, mobile robots, local navigation, Alg1, Alg2, Gazebo, ROS Noetic

1. Introduction

Mobile robots and autonomous navigation are gradually integrated in various aspects of human activities, from occasional operations in dangerous scenarios to daily social interactions. The former include firefighting services [1], [2], urban search and rescue operations [3], [4], and special military operations [5], while the later improve processes in education [6], [7], manufacturing [8], [9], medicine [10], [11], agriculture [12], [13], rehabilitation [14] and service tasks [15], [16]. These tasks require advanced sensory-based autonomous navigation in various environments [17], [18]. Autonomous navigation allows a robot to decide on its motion and actions, based on onboard sensory data about its environmental and current location [19].

There are two types of path-planning: a global navigation and a local navigation [20]. In the global planning approach, a mobile robot has a well-defined map of an environment in which the robot can build its path. In more realistic settings, a robot deals with uncertain maps and relies on its sensors to plan a path [21]. which is called the local navigation. In the local planning approach a robot is being placed at a starting position and must reach a target or report if it cannot be reached while no other information is known to the robot in advance [22]. In this case, the robot uses local sensory data [20] to detect obstacles within its radius of vision [23], which allows the robot to encounter an obstacle only when it hits the obstacle in most algorithms [24].

Boundary-following and Ultimate Goal (BUG) family algorithms were designed to solve the local navigation problem without generating a full map of an environment [22]. Following a BUG family algorithm, a robot could operate in a broad variety of environments [25] and (by an algorithm design) attempts to construct a shortest path toward its destination [26]. BUG algorithms have two modes of motion: moving towards a target and following an obstacle boundary. A robot goes towards the target until it hits any obstacle. Then it starts to follow the boundary until a straight path towards the target becomes clear again [27]. A condition that defines if the path is clear differs depending on a particular algorithm.

In BUG model a robot is considered as a point object [22] in a 2D-map. In this paper a path-planning sensory-based navigation was simulated in a 3D-environment with a real mobile robot model, which employs Alg1 [28] and Alg2 [29] algorithms. The algorithms were implemented using robot operating system (ROS) [30] and evaluated using Gazebo simulator [31].

2. Brief description of the algorithms

Alg1 and Alg2 algorithms belong to BUG family. Two states of a robot under BUG strategy are: 1) go to the target; 2) follow the boundary. Most algorithms employ a line from start point to target, which is called *M-line*. The robot switches from state 1 to state 2 when it hits an obstacle at a point that is defined as a *Hit-point* (*H-point*). The switch from state 1 to state 2 occurs when the robot decides to abandon the current obstacle at a point that is defined as a *Leave-point* (*L-point*). Both H-point and L-point are defined differently by particular algorithms.

2.1. Alg1 algorithm

Alg1 algorithm improves basic Bug2 algorithm [24] in a sense of excluding multiple traces of long segments of a path. It collects H-points (H_i) and L-points (L_j) of previous iterations and uses this information to generate shorter paths by changing a local direction to the opposite. The algorithm works as follows [32]:

0) Initialize iteration i to 0, define *M-line* as line connecting start S and target T points.
1) Increment i and follow *M-line* toward T until either:
   - T is reached. Stop.
   - An obstacle is hit. Define this point as H_i. Go to step 2.
2) Keeping the obstacle on the right, follow the obstacle boundary. Do this until one of the following occurs:
   - T is reached. Stop.
   - A point y is found such that:
     - it is on *M-line* and
     - $d(y, T) < d(x, T)$ for all x ever visited by the robot along *M-line* and
     - The robot can move towards T at y.
   Define this point as L_i and go to step 1.
   - A previously defined point H_j or L_j is encountered such that $j<i$. Change the local direction one and return to H_i. When H_i is reached, follow the obstacle boundary keeping the wall on the left. This rule cannot be applied again until L_i is defined.
   - The robot returns to H_i. T is unreachable. Stop.

2.2. Alg2 algorithm

Alg2 algorithm upgrades Alg1 algorithm by abandoning *M-line* concept. The leaving condition is that

the robot became closer to T than before. The algorithm operates as follows [30]:

0) Initialize iteration i to 0 and $Q = d(S, T)$.
1) Increment i and proceed in the direction of T whilst continuously updating Q to $d(x, T)$ if $Q < d(x, T)$, where x is a current position. Q should now represent the closest to T point where the robot has ever been. Repeat this until one of the following occurs:
   - T is reached. Stop.
   - An obstacle is hit. Define this point H_i. Go to step 2.
2) Keeping the obstacle on the right, follow the obstacle boundary continuously updating Q to $d(x, T)$ if $Q < d(x, T)$. Do this until one of the following occurs:
   - T is reached. Stop.
   - A point y is found such that:
     - $d(y, T) < Q$ and
     - The robot can move towards T at y.
   Define this point as L_i and go to step 1.
   - A previously defined point H_j or L_j is encountered such that $j<i$. Change the local direction and return to H_i. When H_i is reached, follow the obstacle boundary keeping the wall on the left. This rule cannot be applied again until L_i is defined.
   - The robot returns to H_i. T is unreachable. Stop.

3. Implementation details

Ubuntu operating system and ROS Noetic Ninjemys were used. Gazebo simulator was employed for debugging and verifying algorithms' implementation with Turtlebot3 Burger robot model (Fig. 1) from an open source software kit [33].

Alg1 and Alg2 were implemented in Python3 programming language and arranged as a package with two files, containing Alg1 and Alg2 respectively. A main service uses services for going to a point and a wall following in a clockwise and counter clock-wise order.

The following libraries and modules were used for all the services:

- *rospy* – a pure Python client library for ROS;
- *geometry_msgs* module used for generating and sending messages for setting robot's position;
- *sensor_msgs* module to register laser range finder's measurements;
- *gazebo_msgs* module to define and set a robot's current state;
- *tf* module for angles' operations;
- *nav_msgs* module for odometry;
- *std_srv* module for *ros services*.
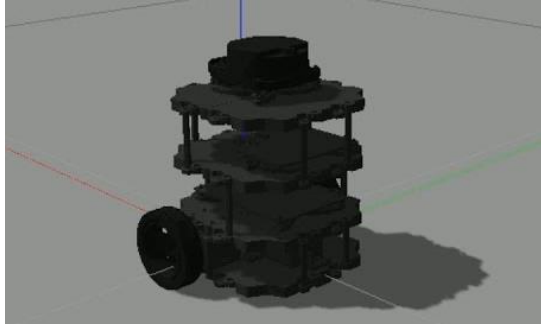
Fig. 1. Turtlebot3 Burger in the Gazebo environment

A main difficulty that was encountered while implementing the algorithms was a definition of suitable constants that allow comparing distance measurements. For example, one of the problems was to define a threshold ε>0 that could be employed to define two distinct robot positions. This way we define that if a Euclidean distance *Dist* between two points p_k and p_m is less than ε, it means that the two points coincide with each other:

$$Dist(p_k, p_m) \leq \varepsilon \iff p_k = p_m \tag{1}$$

A value of ε was chosen empirically to fit all types of environments.

Another problem relates to the construction of the robot that causes stuck because the robot sensor does not consider the robot's wheels, which poke out from the mobile base. In some cases, the robot does not register an obstacle when its wheels already have hit a convex corner. A possible solution is an increase of a threshold δ that helps defining a *H-point* as:

$$Dist(p_k, p_{obs}) \leq \delta \implies p_{obs} = H_i \tag{2}$$

where p_{obs} is a point on a boundary of a currently hit obstacle, p_k is a current position of the robot, H_i is a newly defined *H*-point. Yet, this may cause the algorithms' failure since the robot may miss a *H-point* and thus skips a switch of its state into the boundary following mode.

Finally, one more difficulty appeared only for Alg2 algorithm at step 1 when it checks whether the robot became closer to *T* than before ($Q < d(x, T)$). While in a mathematical sense this comparison is performed constantly, in practice a particular small time step *Δt* between checking a new value of *x* should be selected. The value of *Δt* was found empirically so that it successfully works for both employed types of maps: separate convex obstacles and mazes.

4. Validation

To validate the implementation of Alg1 and Alg2 algorithms 45 tests within two different types of environments were conducted. While there exist several popular tools for environment construction for Gazebo worlds, ranging from semiautonomous generators of single environments from 2D images [18],[34] to autonomous generators of multiple environments, including maze-like environments [35] and random step environment generators [36] using ergonomic graphical user interfaces, it was decided to construct two environments manually in order to ensure interesting cases for algorithms' testing. The first constructed environment was bounded by an external non-traversable black wall with five green towers and contained separate convex obstacles in a form of nine identical columns. The second environment was a maze. 25 and 20 test cases within the convex environment and the maze were conducted, correspondingly. For each test cases four runs were conducted. In total, 90,5% of the tests were successful while seven tests in the convex environment and ten tests in the maze failed due to the problems that were stated in Section 3. Additional tests for the target reachability were performed successfully in both environment. Table 1 presents the system configuration.

Table 1. System configuration

Parameter	Characteristics
Operation System details	Ubuntu 20.0.4
Memory	8.00 GB
Processor	Intel(R) Core (TM) i7-4510U CPU @ 2.00GHz 2.60 GHz

Fig. 2 presents a path constructed by Alg1 algorithm in the convex environment. Due to the environment's simplicity, a path of Alg2 algorithm was almost identical. Fig. 3 and Fig. 4 demonstrate paths with the same *S, T* locations that were constructed in the maze environment by Alg1 and Alg2 respectively. In this particular case Alg2 outperformed Alg1; it is wrong to state that Alg2 always outperforms Alg1 in every case as a result depends on the *S, T* and the environment selection.
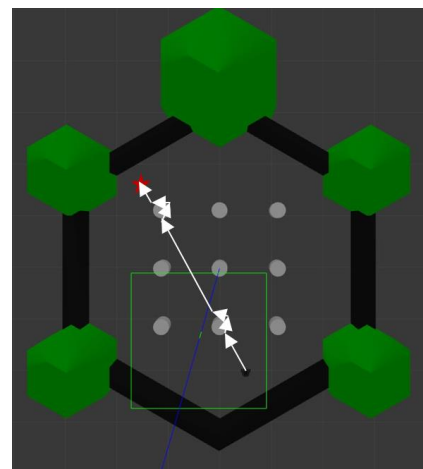
Fig. 2. A path of Alg1 in the convex environment is shown by white lines. The arrows depict the direction of motion, the red star marks the target.

Anastasia Yankova, Timur Gamberov, Tatyana Tsoy
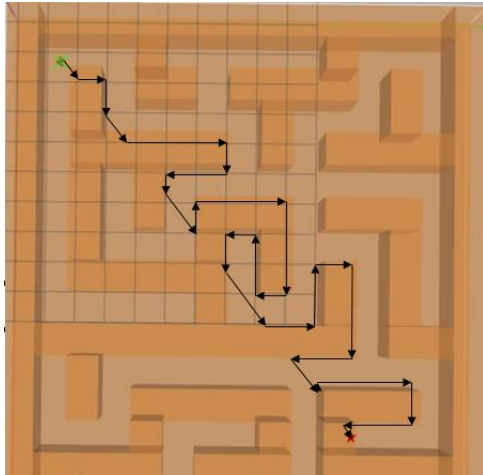
Fig. 3. A path of Alg1 in the maze
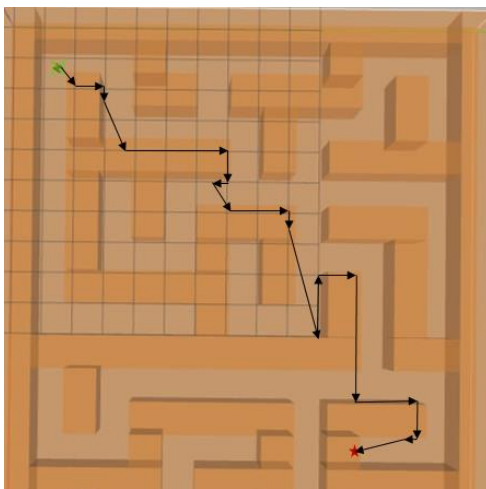
Fig. 4. A path of Alg2 in the maze

5. Conclusions

The paper presented the implementation of Alg1 and Alg2 algorithms in Python programming language for Turtlebot3 Burger robot model using ROS Noetic. The conducted tests successfully validated the implemented algorithms in the simple convex environment and in the maze. A number of difficulties that were encountered while implementing the algorithms are discussed.

Acknowledgment

References

1. J. Zhu, W. Li, D. Lin, H. Cheng and G. Zhao, Intelligent fire monitor for fire robot based on infrared image feedback control. Fire Technology, vol. 56, no. 5, 2020, pp. 2089–2109.

2. R. Bogue, The role of robots in firefighting. Industrial Robot: the international journal of robotics research and application, 48(2), 2021, pp.174-178.

3. E. Magid, K. Ozawa, T. Tsubouchi, E. Koyanagi and T. Yoshida, Rescue Robot Navigation: Static Stability Estimation in Random Step Environment. Lecture Notes in Computer Science, vol. 5325, 2008, p. 305-316.

4. E. Magid, Simulation Tools for Urban Search and Rescue Robotics. International Conference on Artificial Life and Robotics (ICAROB 2023), 2023, pp. 4-11.

5. C. Chen, S. Wang, L. Li, S. Ke, C. Wang and X. Bu, Intelligent covert satellite communication for military robot swarm. IEEE Access, vol. 8, 2019, pp. 5363–5382.

6. T. Tsoy, L. Sabirova, M. Abramsky and E. Magid, Establishing effective teaching for robotics: a comparison study of bachelor students participated in introduction to robotics course. In Int. Conf. on Artificial ALife and Robotics, 2018, pp. 212–215.

7. E. Chebotareva and M. Mustafin, Android Based Educational Mobile Robot Design and Pilot Evaluations. International Conference on Artificial Life and Robotics (ICAROB 2023), 2023, pp. 146-149.

8. A. Dobrokvashina, S. Sulaiman, A. Zagirov, E. Chebotareva, K.-H. Hsia and E. Magid, Human robot interaction in collaborative manufacturing scenarios: Prospective cases. Siberian Conference on Control and Communications, (SIBCON 2022), November 2022, p. 1-6.

9. R. Galin and R. Meshcheryakov, Automation and robotics in the context of Industry 4.0: the shift to collaborative robots. In IOP Conference Series: Materials Science and Engineering, vol. 537, no. 3, May 2019, p. 032073.

10. E. Magid, A. Zakiev, T. Tsoy, R. Lavrenov and A. Rizvanov, Automating pandemic mitigation. Advanced Robotics, vol. 35, no. 9, 2021, pp. 572–589.

11. R. Safin, R. Lavrenov, K.-H. Hsia, E. Maslak, N. Schiefermeier-Mach and E. Magid, Modelling a TurtleBot3 Based Delivery System for a Smart Hospital in Gazebo. The 15th Siberian Conference on Control and Communications (SIBCON 2021), 2021, № 9438875.

12. J. Samaniego, E. López-González, E. Magid, and E. A. Martínez-García, Path-Tracking in Double Spiraliform Sowing Fields with Agricultural Robotics. 2023, Available at SSRN 4589037.

13. M.S.A. Mahmud, M.S.Z. Abidin, A.A. Emmanuel and H.S. Hasan, Robotics and automation in agriculture: present and future applications. Applications of Modelling and Simulation, vol. 4, 2020, pp.130-140.

14. C. Eicher, M. Haesner, M. Spranger, O. Kuzmicheva, A. Gräser and E. Steinhagen-Thiessen, Usability and acceptability by a younger and older user group regarding a mobile robot-supported gait rehabilitation system. Assistive technology, 31(1), 2019, p. 25-33.

15. D. Ryumin, I. Kagirov, A. Axyonov, N. Pavlyuk, A. Saveliev, I. Kipyatkova, M. Zelezny, I. Mporas and A. Karpov, A multimodal user interface for an assistive robotic shopping cart. Electronics, vol. 9, no. 12, 2020, p. 2093.

16. E. Chebotareva, R. Safn, K.-H. Hsia, A. Carballo and E. Magid, Person-following algorithm based on laser range finder and monocular camera data fusion for a wheeled autonomous mobile robot, in International Conference on Interactive Collaborative Robotics (ICR 2020). Lecture Notes in Computer Science, vol 12336, 2020.

17. A. Eryomin, R. Safin, T. Tsoy, R. Lavrenov, and E. Magid, Optical Sensors Fusion Approaches for Map Construction: A Review of Recent Studies. In Journal of Robotics,

Networking and Artificial Life, vol. 10(2), 2023, pp. 127-130.

18. R. Lavrenov, A. Zakiev and E. Magid, Automatic mapping and filtering tool: From a sensor-based occupancy grid to a 3D Gazebo octomap. International Conference on Mechanical, System and Control Engineering (ICMSC), 2017, pp. 190-195.

19. M. Mustafin, T. Tsoy, E.A. Martınez-Garcıa, R. Meshcheryakov and E. Magid, Modelling mobile robot navigation in 3D environments: camera-based stairs recognition in Gazebo. Moscow Workshop on Electronic and Networking Technologies (MWENT-2022), 2022, pp. 1-6.

20. E. Magid and E. Rivlin, CAUTIOUSBUG: A Competitive Algorithm for Sensory-Based Robot Navigation. Proceedings of 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2003, pp. 2757-2762.

21. I. Kamon and E. Rivlin, Sensory-based motion planning with global proofs. IEEE Transactions on Robotics and Automation, vol.13, 1997, pp. 814-822.

22. N. James and T. Bräunl, Performance comparison of bug navigation algorithms. Journal of Intelligent and Robotic Systems 50, no. 1, 2007, pp. 73- 84.

23. V. Lumelsky and T. Skewis, Incorporating range sensing in the robot navigation function. IEEE Transactions on Systems, Man, and Cybernetics, 1990, pp. 1058-1069.

24. V. Lumelsky and A. Stepanov, Path-planning strategies for a point mobile automaton moving amidst unknown obstacles of arbitrary shape. In Autonomous robot vehicles, 1990, pp. 363-390.

25. H. Noborio, Evaluation of Path Length Made in SensorBased Path-Planning with the Alternative Following. Proceedings of the 2001 IEEE International Conference on Robotics 8 Automation, 2001, pp. 1728-1735.

26. H. Noborio, Several Path-Planning Algorithms of a Mobile Robot for an Uncertain Workspace and their Evaluation. Proc. of the IEEE Intelligent Motion Contro, 1990, pp. 289-294.

27. A. Yufka and O. Parlaktuna Performance comparison of bug algorithms for mobile robots. In Proceedings of the 5th international advanced technologies symposium, 2009, pp. 13-15.

28. A. Sankaranarayanan and M. Vidyasagar, A New Path Planning Algorithm For Moving A Point Object Amidst Unknown Obstacles In A Plane. IEEE Conference on Robotics and Automation, 1990, pp. 1930-1936.

29. A. Sankaranarayanan and M. Vidyasagar, Path Planning For Moving A Point Object Amidst Unknown Obstacles In A Plane: A New Algorithm And A General Theory For Algorithm Development. Proc. of the 29th Conference on Decision and Control, 1991, pp. 1111-1119.

30. A. Apurin, B. Abbyasov, E.A.Martínez-García and E. Magid Comparison of ROS Local Planners for a Holonomic Robot in Gazebo Simulator. In International Conference on Interactive Collaborative Robotics, Cham: Springer Nature Switzerland, 2023, pp. 116-126.

31. A. Dobrokvashina, R. Lavrenov, E. Magid, Y. Bai and M. Svinin. How to Create a New Model of a Mobile Robot in ROS/Gazebo Environment: An Extended Tutorial. International Journal of Mechanical Engineering and Robotics Research, 12(4), 2023, pp. 192-199.

32. N. James and T. Bräunl, A Practical Comparison of Robot Path Planning Algorithms given only Local Information. Centre for Intelligent Information Processing Systems School of Electrical and Electronic Engineering The University of Western Australia, 2005.

33. F. U. Pereira, P. M. de A. Brasil, M. A. De S. L. Cuadros and A. R. Cukla, Analysis of Local Trajectory Planners for Mobile Robot with Robot Operating System. IEEE Latin America Transactions, vol. 20(1), 2022, pp. 92-99.

34. B. Abbyasov, R. Lavrenov, A. Zakiev, K. Yakovlev, M. Svinin and E. Magid, Automatic Tool for Gazebo World Construction: From a Grayscale Image to a 3D Solid Model. International Conference on Robotics and Automation (ICRA), 2020, p. 7226-7232.

35. A. Iskhakova, B. Abbyasov, T. Tsoy, T. Mironchuk, M. Svinin and E. Magid, LIRS-Mazegen: An Easy-to-Use Blender Extension for Modeling Maze-Like Environments for Gazebo Simulator. Frontiers in Robotics and Electromechanics, vol. 329, 2023, pp. 147-161.

36. R. Gabdrahmanov, T. Tsoy, Y. Bai, M. Svinin and E. Magid, Automatic Generation of Random Step Environment Models for Gazebo Simulator. Lecture Notes in Networks and Systems, vol. 324, 2021, p. 408-420.

Authors Introduction

Ms. Anastasia Yankova

Anastasia Yankova is a master student in System analysis and mathematical technologies at HSE University, Russia. Previously she obtained a bachelor degree in information systems at ITMO University, Russia.

Mr. Timur Gamberov
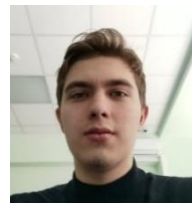
Timur Gamberov is a bachelor student at Institute of Information Technology and Intelligent Systems, Kazan Federal University, Russia.

Ms. Tatyana Tsoy

In 2012 she graduated from International Area Studies Master's Degree Program at the University of Tsukuba. In 2023 completed a PhD in Computer science and information processes at ITIS, KFU, Russia.

Implementation of VisBug-21 and VisBug-22 Path Planning Algorithms Using ROS Noetic

Viktoriia Mirzoian
HSE University, Moscow, Russian Federation

Maxim Mustafin
Intelligent Robotics Department, Kazan Federal University, 420008, Kazan, Russian Federation

Evgeni Magid
Intelligent Robotics Department, Kazan Federal University, 420008, Kazan, Russian Federation
HSE University, Moscow, Russian Federation
E-mail: vimirzoyan@edu.hse.ru, mustafin@it.kfu.ru, magid@it.kfu.ru

Abstract

Local navigation algorithms are crucial for autonomous robots operating in unknown environments where a presence of obstacles and dynamic changes pose significant challenges. A focus of these algorithms is to enable a real-time path calculation, allowing a robot to adapt its states dynamically based on corresponding environmental conditions, despite an absence of prior knowledge about surroundings. This paper presents an implementation of the VisBug-21 and VisBug-22 algorithms, designed to address challenges of a local navigation. The algorithms were implemented for a differential drive robot Turtlebot3 Burger using Robot Operation System (ROS). Virtual experiments were performed in the Gazebo simulator employing a simple 3D environment that contained only convex obstacles and a small 3D maze.

Keywords: Path-Planning, Mobile Robots, Local Navigation, VisBug-21, VisBug-22, Gazebo, ROS Noetic

1. Introduction

Robotics researchers face a significant challenge when it comes to efficient navigation [1]. A current focus in robot motion planning concentrates around two different models, each based on distinct assumptions about available data for planning [2]. The first model, known as a path planning with complete information or the *piano movers problem*, assumes perfect knowledge about a robot and obstacles [3]. This paper is concerned with the second model, which is a path planning with incomplete information [4]. In this model, some degree of uncertainty exists, and missing data are obtained from local sources, e.g., laser range finders or cameras. One advantage of this model is the ability to incorporate sensory feedback, transforming a motion planning into a continuous dynamic process. It also eliminates a need for an analytic representation of obstacle boundaries [5]. A key issue in the motion planning with incomplete information is how to integrate sensory data into a planning function [6], [7].

This paper describes how to implement two mathematically described algorithms that can guide a mobile robot through complex environments with various obstacles, using basic distance data provided by a range finder or stereo vision. To achieve this goal, a simplified model of a vision sensor [8] is proposed, which functions similarly to a rangefinder, providing the robot with coordinates of obstacles' boundaries within a limited radius of its visibility [9].

2. Algorithms' overview

BUG family of algorithms provides efficient and effective ways for robots to navigate through populated with obstacles environments and reach their target locations [10]. These algorithms are widely used in robotics research and are valuable tools for path planning and navigation applications. VisBug-21 and VisBug-22 algorithms are an enhanced version of the basic BUG algorithm [8], which allow the robot making more informed decisions by introducing a vision control.

2.1. *VisBUG-21*

Incorporating vision into Bug2 algorithm is to mentally reconstruct a segment of Bug2 path visible to the robot at each step [11]. Using this reconstructed path, the robot identifies a farthest point and moves directly towards it. It is important to note that a segment continuity is a crucial factor in this process, rather than remembering the entire segment itself [12], [13]. The algorithm is comprised of two key components: a main body that plans a motion along a path responsible for generating the

Viktoria Mirzoian, Maksim Mustafin, Evgeni Magid

next intermediate target T_{i+1}, and a procedure "Compute T_{i+1}" that tests its reachability [8]. Typically, the main body only operates in step S1, but step S2 is executed when the robot is navigating a (locally) convex boundary of an obstacle, preventing it from utilizing its vision to define next intermediate target T_{i+1}.

Main body:

1. Move towards current target T_i, execute "Compute T_i-21" and check:
   a. If start point C = target point T, stop the procedure.
   b. If T is unreachable, stop the procedure.
   c. If C = T_i, go to step 2.
2. Move along the obstacle boundary, execute "Compute T_i-21" and check:
   a. If C = T, stop the procedure.
   b. If T is unreachable, stop the procedure.
   c. If C $\neq$ T_i, go to step 1.

Procedure Compute Ti-21 builds a path that BUG2 algorithm would build, and then selects a farthest point on a section of a path that it sees. This point becomes a next current target.

2.2. *VisBug-22*

It is possible to create a different method, similar to BUG2 mechanism but with a more opportunistic approach compared to VisBug-21. Instead of strictly following a BUG2 path, the robot can deviate from it if it spots more promising opportunities, while still ensuring a convergence. This alternative process is referred to as VisBug-22. A design of this algorithm bears a strong resemblance to VisBug-21 algorithm, albeit with one notable difference - the robot no longer prioritizes making sure that all intermediate targets T_j are located on the BUG2 path [11]. Instead, it strives to select intermediate targets that are as close as possible to the target T while still being on the same line. This leads to a distinct convergence mechanism and a behavior that sets it apart from VisBug-21 algorithm. The algorithm is made up of two parts: a main body, which is the same as the main body of VisBug-21, and a procedure named «Compute T_i-22». This procedure calculates a next intermediate target T based on a current position C of the robot and also checks if the target can be reached.
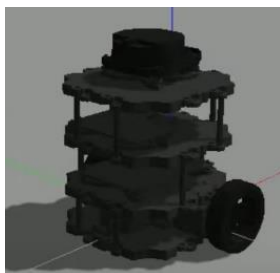
Fig.1. Turtlebot3 Burger in the Gazebo environment

Procedure Compute T_i-22 keeps identifying points along the BUG2 path or a quasi- BUG2 path segment, until a better point (in terms of its proximity to T), S', is identified on the line Start-Target. Then S' becomes the starting point of another quasi- BUG2 path segment, and the process repeats. As a result, unlike algorithms Bug2 and VisBug-21, where each defined hit point has its matching leave point, in VisBug-22 no such matching necessarily occurs.

3. Configuration

An environment configuration consists of Robot Operation System (ROS), Gazebo, and the TurtleBot. ROS was chosen due to its modular framework that allows developers breaking down complex algorithms into smaller, reusable software components known as nodes. This modularity enables easier development, testing, and maintenance of robot motion algorithms [14]. Additionally, ROS provides a wide range of sensor drivers and libraries that make it easier to integrate different sensors into a robot's system. This allows robots gathering data from various sources, such as cameras, LiDAR, or depth sensors, which is crucial for navigating in unknown environments [15].

The benefit of using Gazebo for programming algorithms for robot movement in unknown conditions is a realistic physics simulation environment for robots [16], [17] allowing developers to test and refine their algorithms in various unknown conditions without a need for physical prototypes and without damage risks. This enables faster development and iteration cycles. Gazebo provides a highly customizable and extensible platform where developers can create virtual environments with different terrain types, lighting conditions, or weather effects to simulate distinct unknown conditions [18]. This flexibility allows for comprehensive testing and robustness evaluation of algorithms in various scenarios.
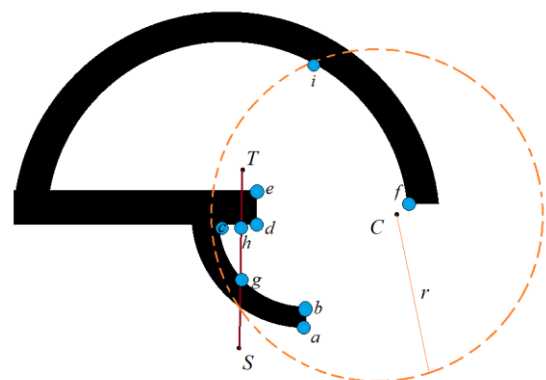
Fig. 2. Schematic representation of the search for "interesting" points

TurtleBot 3 Burger virtual model of a robot was employed for experiments in Gazebo simulator (Fig. 1) due to its simplicity that makes it accessible for users of
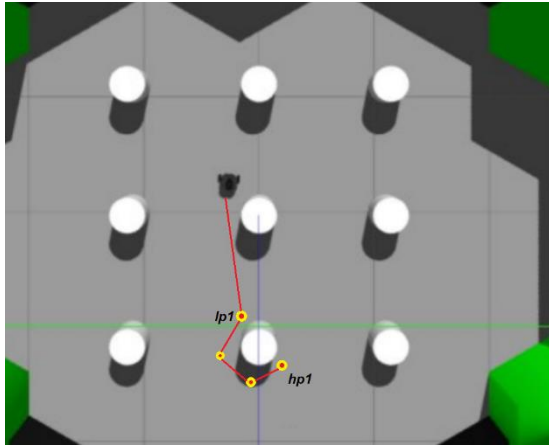
Fig. 3. VisBug-21 path in the map with convex pillars

all levels of expertise, including beginners [19], round shape and differential drive chassis that ease local navigation. The robot's small size allows navigating through tight spaces, making it suitable for various environments, including confined ones.

4. Implementation details

Python 3 was utilized to implement VisBug-21 and VisBug-22. The package includes two separate files containing the algorithms for VisBug-21 and VisBug-22, representing distinct processes. Additionally, the main service offers user services for point navigation, clockwise and counterclockwise wall following. To ensure the functionality of all services, the following libraries and modules were employed [20], [21]:

- *rospy*: is a Python client library specifically designed for ROS.
- *sensor_msgs* module: records laser measurements.
- *geometry_msgs* module: generates and transmitts messages pertaining to setting the robot's position and utilizing points.
- *gazebo_msgs* module: defines and configures the robot's current state.
- *nav_msgs* module: odometry use.
- *tf* module: allows angle transformations.
- *std_srv* module: involved in running ROS services.

A main difficulty in implementing the algorithm was a vision system design for the robot. The original articles [8], [11] described only a visual sensor range concept. It was decided to model the vision system using a rangefinder package of ROS. The algorithm contains a method that serves as a callback for processing laser scan messages. The purpose of this function is to extract specific ranges of data from the laser scan message and store them as global variables. It defines global variables for each range of 360 degrees around the robot, populates a range dictionary with minimum range values from specific ranges and assigns a maximal vision range value if no valid readings are available.
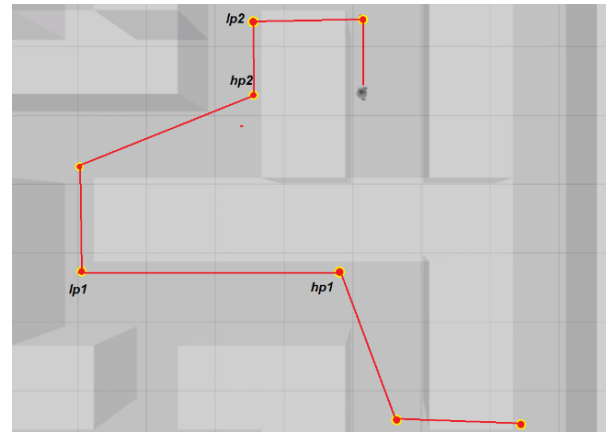
Fig. 4. VisBug-21 path in the maze

Another difficulty was related to a practical identification of "interesting" points that were described in [1], [2]. At the current location C of the robot (Fig. 2), the algorithm processes a grid of its sensor's vision within its radius r (an area inside the orange circle). The robot rotates 360 degrees in place and searches for interesting points (marked with blue dots) that form three groups:

- corners of obstacles (indicated as *a, b, c, d, e, f*);
- visible points of obstacles' boundaries at distance r (indicated as *i*);
- visible intersections of M-line and obstacles (*g, h*).

5. Validation

The algorithms' implementation was validated in two types of environments: a world of identical pillars of a cylindric shape and a maze. Virtual tests in Gazebo were carried out for reachable and unreachable targets. 20 (Start, Target) pairs were selected in the first environment and 25 pairs in the second. Each test was run multiple times. While most tests were successful, 9.5% of the tests failed due to situations at the boundaries of the obstacles where the robot's wheels were not considered by the sensors and thus the robot got stuck.

Table 1 presents system configuration that was used for the validation.

Table 1. System configuration

Parameter	Characteristics
Operation System details	Ubuntu 20.0.4
Memory	16.00 GB
Processor	AMD Ryzen 5 4600H with Radeon Graphics 3.00 GHz
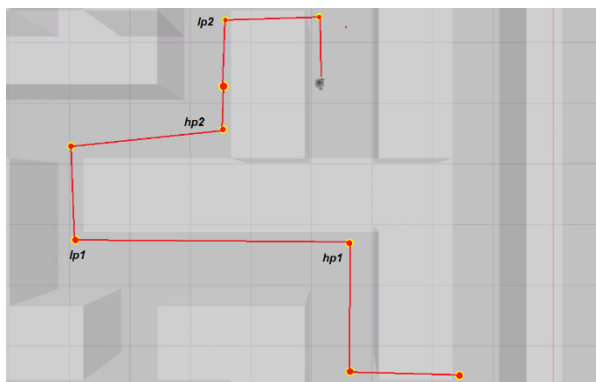
Fig. 5. VisBug-22 path in the maze

Fig. 3 present a path constructed by VisBug-21 algorithm in the convex environment. Fig. 4 and Fig. 5 demonstrate paths with the same S, T selection in the maze environment by VisBug-21 and VisBug-22 respectively.

6. Conclusions

The paper presented the implementation of VisBug-21 and VisBug-22 algorithms in Python programming language for Turtlebot3 Burger robot model using ROS Noetic. The implementation was validated in two different environments of Gazebo simulator and demonstrated a success in 90.5% of the test cases.

Acknowledgment

References

1. G. Huang, Visual-inertial navigation: A concise review. In 2019 international conference on robotics and automation (ICRA), May 2019, pp. 9572-9582.
2. J. Kerr and K. Nickels, Robot operating systems: Bridging the gap between human and robot Proceedings. 44th IEEE Southeastern Symposium on System Theory (SSST), 2012.
3. G. E. Jan, K. Y. Chang and I. Parberry, Optimal path planning for mobile robot navigation. IEEE/ASME Transactions on mechatronics, 13(4), 2008, pp. 451-460.
4. E. Magid and E. Rivlin, CAUTIOUSBUG: A Competitive Algorithm for Sensory-Based Robot Navigation. Proceedings of 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2003, pp. 2757-2762.
5. I. Kamon and E. Rivlin, Sensory-based motion planning with global proofs. IEEE Transactions on Robotics and Automation, vol. 13, 1997, pp. 814-822.
6. E. Magid, Simulation Tools for Urban Search and Rescue Robotics. International Conference on Artificial Life and Robotics (ICAROB 2023), 2023, pp. 4-11.
7. I. Mavrin, T. Tsoy and E. Magid, Modified E3 exploration algorithm for unknown environments with obstacles. 13th

8. Asian Control Conference (ASCC), May 2022, pp. 1413-1418.
8. V. Lumelsky and A. Stepanov, Path-planning strategies for a point mobile automaton moving amidst unknown obstacles of arbitrary shape. In Autonomous robot vehicles, 1990, pp 363-390.
9. H. Noborio, Evaluation of Path Length Made in SensorBased Path-Planning with the Alternative Following. Proceedings of the 2001 IEEE International Conference on Robotics 8 Automation, 2001, pp. 1728-1735.
10. H. Noborio, Several Path-Planning Algorithms of a Mobile Robot for an Uncertain Workspace and their Evaluation. Proc. of the IEEE Intelligent Motion Control, 1990, pp. 289-294.
11. V. Lumelsky and T. Skewis, Incorporating range sensing in the robot navigation function. IEEE Transactions on Systems, Man, and Cybernetics, 1990, pp. 1058-1069.
12. A. Yufka and O. Parlaktuna, Performance comparison of bug algorithms for mobile robots. In Proceedings of the 5th international advanced technologies symposium, 2009, pp. 13-15.
13. A. Sankaranarayanan and M. Vidyasagar, Path Planning For Moving A Point Object Amidst Unknown Obstacles In A Plane: A New Algorithm and A General Theory for Algorithm Development. Proc. of the 29th Conference on Decision and Control, 1991, pp. 1111-1119.
14. N. James and T. Bräunl, Performance comparison of bug navigation algorithms. Journal of Intelligent and Robotic Systems 50, no. 1, 2007, pp. 73- 84.
15. A. Sankaranarayanan and M. Vidyasagar, A New Path Planning Algorithm for Moving A Point Object Amidst Unknown Obstacles in A Plane. IEEE Conference on Robotics and Automation, 1990, pp. 1930-1936.
16. A. Iskhakova, B. Abbyasov, T. Tsoy, T. Mironchuk, M. Svinin and E. Magid, LIRS-Mazegen: An Easy-to-Use Blender Extension for Modeling Maze-Like Environments for Gazebo Simulator. Frontiers in Robotics and Electromechanics, vol.329, 2023, pp. 147-161.
17. A. Khazetdinov, A. Zakiev, T. Tsoy, R. Lavrenov and K.-H. Hsia, Standard-complaint Gazebo warehouse modelling and validation. Proceedings of 13th International Conference on Developments in eSystems Engineering (DeSE), 2020, p. 218-221.
18. R. Gabdrahmanov, T. Tsoy, Y. Bai, M. Svinin and E. Magid, Automatic Generation of Random Step Environment Models for Gazebo Simulator. Lecture Notes in Networks and Systems, №324, 2021, pp. 408-420.
19. E. Chebotareva, T. Tsoy, B. Abbyasov, J. Mustafina, E. A. Martinez-Garcia, Y. Bai and M. Svinin, On the Problems of SLAM Simulation for Mobile Robots in the Arctic Conditions. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 12336, 2020, p. 34-44.
20. F. Duan, W. Li and Y. Tan, The Framework and Fundamental Use of ROS. In Intelligent Robot: Implementation and Applications, 2023, pp. 43-69.
21. A. Dobrokvashina, R. Lavrenov, T. Tsoy, E. A. Martinez-Garcia and Y. Bai, Navigation stack for the crawler robot Servosila Engineer. Proceedings of the IEEE 16th Conference on Industrial Electronics and Applications (ICIEA), 2021, p. 1907-1912.

Authors Introduction

Ms.Viktoria Mirzoian

She is a master student of The Internet of Things and cyber physical systems program at HSE University, Russian Federation. She got a bachelor's degree in information systems and technologies at Saint Petersburg Mining University, Russia.

Mr. Maksim Mustafin

He is a master student of Intelligent Robotics program at Institute of Information Technology and Intelligent Systems (ITIS), Kazan Federal University (KFU), Russia. In 2023 he obtained a bachelor degree at ITIS, KFU.

Professor Evgeni Magid

He is a Head of Intelligent Robotics Department and a Head of Laboratory of Intelligent Robotic Systems (LIRS) at Kazan Federal University, Russia. Professor at HSE University, Russia. Senior IEEE member. Previously he worked at University of Bristol, UK; Carnegie Mellon University, USA; University of Tsukuba, Japan; National Institute of Advanced Industrial Science and Technology, Japan. He earned his Ph.D. degree from University of Tsukuba, Japan. He authors over 200 publications.

DistBug path planning algorithm package for ROS Noetic

Alexander Pak
Higher School of Economics University, Moscow, Russian Federation

Alexander Eremin
Intelligent Robotics Department, Kazan Federal University, Kazan, Russian Federation

Tatyana Tsoy
Intelligent Robotics Department, Kazan Federal University, Kazan, Russian Federation
E-mail: adpak_1@edu.hse.ru, aneremin@it.kfu.ru, tt@it.kfu.ru

Abstract

Algorithms of path-planning in an unknown environment play an important role in robotics. They do not require a prior information about obstacles' locations around a robot and allow calculating a path in a real time. This article presents an implementation of a sensory-based DistBug algorithm, which operates reactively using range data for immediate decision-making without constructing a world model. The algorithm was programmed in Python using robot operating system (ROS) and validated in the Gazebo simulator. For virtual experiments Turtlebot 3 Burger mobile robot was employed. The experiments were conducted in two types of environment: an environment with convex obstacles and a maze. The paper demonstrates analysis of experiments using several standard criteria of a path quality estimation.

Keywords: BUG algorithm, path planning, sensor-based navigation.

1. Introduction

A basic motion planning issue is to compute a path from a given starting point to a specified destination point [1]. In robotics, there are two approaches for path planning: global and local [2]. Global planning typically relies on preliminarily known data about an environment [3]. Local path planning uses measurements from sensors and knowledge of a target point [4] while the robot moves in an unknown environment avoiding obstacles [5].

The algorithms of the BUG family use only local sensory information and odometry data to control movement and address a challenge of local navigation without a need to create a comprehensive map of an environment [6]. BUG algorithms have two motion modes: moving toward a target and following a boundary of an obstacle [7]. The robot moves in a straight line towards the target until it reaches a minimum distance to an obstacle [8]. Further, the robot follows a boundary of the obstacle [9]. The algorithm uses *leaving conditions* to determine when to exit the obstacle boundary and head directly toward the target again [10]. The robot receives measurements from its sensors to detect nearby obstacles and plan a subsequent trajectory [11].

In this article, we present an implementation of the DistBug [12] algorithm using robot operating system

(ROS) [13]. The algorithm was tested in the Gazebo simulator [14]. The navigation was performed in a 3D environment for the TurtleBot 3 Burger robot model [15]. The robot measures a distance to an obstacle with a 2D LiDAR of 3.5 m range and 360 degrees field of view.

2. A brief description of the DistBug algorithm

The sensory-based DistBug algorithm was designed to reach a goal in an unknown environment or indicate that the goal is unreachable [12]. The robot does not construct a model of the world and relies on sensory data to make decisions [16]. The method consists of two behavior models, represented as movement modes: a straight-line motion towards the goal and an obstacle boundary following. A condition for leaving an obstacle is based on a free range in a goal direction when the target is on a line of sight or a distance to it gradually decreases. The algorithm utilizes sensory data with a limited range of view, and the robot moves in steps of a given size. The core concept of the algorithm is the following:

1) Moving from starting point S to target point T in a straight line:

- If T is reached, the navigation is successfully completed.

- If an obstacle is encountered, a collision point with the obstacle is determined, and the algorithm proceeds to step 2.

2) The robot follows a boundary of the obstacle, taking steps of a specified size. If the robot reaches *T*, the navigation is successfully completed. Otherwise, the robot leaves the obstacle and heads towards *T* from a point of exit, referred as a *leaving point*. The selection of the leaving point depends on achieving the following conditions:

- The robot can move towards *T* without colliding with the obstacle.
- $Free_{dist} > 0$, and ($Curr_{dist} - Free_{dist} \leq 0$ or $Curr_{dist} - Free_{dist} \leq Best_{dist}$), where $Free_{dist}$ is a distance within the line of sight from a current location to a nearest obstacle in a direction of *T*, $Curr_{dist}$ is a distance from the current location to *T*, and $Best_{dist}$ is a difference between the distance from the collision point with the obstacle to *T* and the step size.

3) If the robot cannot find a leaving point or it returns to a position it was at a previous step, then *T* is unreachable.

3. System setup

The algorithm was implemented in Noetic Ninjemys version of ROS. The algorithm uses ROS topics to obtain measurements from sensors and wheel odometry [17], and to control the robot. Messages are transmitted using standard types predefined in ROS: *sensor_msgs:: PointCloud*, *geometry_msgs::Twist*, *nav_msgs::Odometry*, and *geometry_msgs::Point*.

The algorithm was tested in the ROS-compatible Gazebo 11 simulator [18]. The Gazebo simulator creates realistic environments that allow conducting experiments with virtual robot models. The robot can be equipped with any of the standard sensors, and Gazebo has a support for third-party plug-ins. The Robot Visualization tool (RViz, [19]) is used to visualize a robot's position, measurements from sensors, and a path trajectory. The tool allows setting a target point as a 2D Nav Goal. A message is sent to an appropriate topic and accepted by a local planner.

The TurtleBot3 robot model (Fig. 1) with an open-source software was selected. The TurtleBot3 Burger [20] model was used, which has necessary sensors for navigating an environment under DistBug protocol. The robot's design allows making a transition from a 3D configuration space to a 2D one, which simplifies a path calculation [21].

4. Implementation details

DistBug algorithm was implemented in Python programming language, version 3.7. The implementation relies on ROS libraries and modules, including *rospy* (a Python client library for ROS), *geometry_msgs* (contains a type of messages for exchanging a position of the robot), *sensor_msgs* (contains types of laser measurement messages and motion commands), *gazebo_msgs* (contains types that determine a robot state), *tf* (manages transformations between coordinate systems), *nav_msgs* (contains the odometry data type), *std_srv* (contains packages for creating services in ROS) [22].

Fig. 1. Turtlebot3 Burger model in the Gazebo.

The algorithm's implementation contains three nodes. The first and second nodes regulate a robot motion directly to the target and the clockwise wall following motion, respectively. Each node implements a ROS server to activate and deactivate the mode. The third node launches a certain mode via a request to a corresponding server. The server choice depends on measurements received from the laser rangefinder. If the rangefinder indicates an occupied space ahead of the robot while the robot is moving towards the target, the robot changes the motion mode. Similarly, if a space in a direction of the target is free while the robot is moving along an obstacle boundary, the robot switches the mode. When the stop condition occurs, the algorithm turns off all nodes. To evaluate a performance the algorithm logs its operation time, all trajectory points (to calculate a path length), and a number of encountered obstacles during a path search.

To configure the algorithm, the following parameters were introduced: a timeout, a stopping threshold, a range of the rangefinder (i.e., radius of the robot's vision), and a robot's step size. The stopping threshold was empirically defined to determine a point where the robot stops when a certain coordinate value is reached. The timeout allows avoiding robot's stuck cases: despite a static environment, in some cases the TurtleBot 3 robot can get stuck at an obstacle boundary if one of its wheel touches an outer corner of the obstacle. This occurs due

to a rangefinder's placement that may cause a misregistration of occupied cells of the environment as free cells. If the robot gets stuck, a timer is activated and after the timeout expires, the algorithm stops with an error message. The rangefinder range and the step size are user provided parameters. The range should be less than the step size. A small step size causes a better precision of a computed path with regard to a theoretical path of an infinitely smell step and affects how often the obstacle detour check is be performed. With a large step size a robot may miss the target. The rangefinder measurements were divided into sectors by their direction. The value of a sector was determined by a minimum distance to an obstacle in the range.

5. Validation

To evaluate the implementation, tests were conducted in two types of environment: an environment with convex 3D obstacles of a cylindrical shape and a maze. The tests included reachable and unreachable goals. In the latter case the algorithm returned an error after reaching the stop state. The test results were recorded into a file that contained tracked metrics for each (start S, target T) pair.
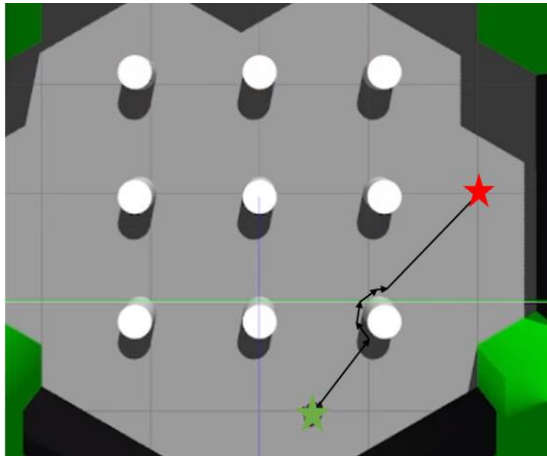
Fig. 2. A path of the DistBug in the simple map. The green star marks the start, the red star is the target.

We made 20 selection of (S,T) for the simple map (Fig. 2) and 15 pairs for the maze (Fig. 3). For each pair a number of runs was performed and they demonstrated approximately 97% success. The failures were cause by the problem mentioned in Section 4. Fig. 2 and Fig. 3 present the resulted paths of the robot using DistBug in the simple map and in the maze respectively. The virtual experiments were conducted using a PC with Ubuntu 20.0.4(LTS) operation system, 6.00 GB memory and AMD Ryzen 5 5500U processor with 2.10 GHz Radeon Graphics. While the environments for testing were constructed manually, a part of our ongoing work is to test the algorithm in automatically constructed environments for Gazebo, e.g., [23], [24], [25].
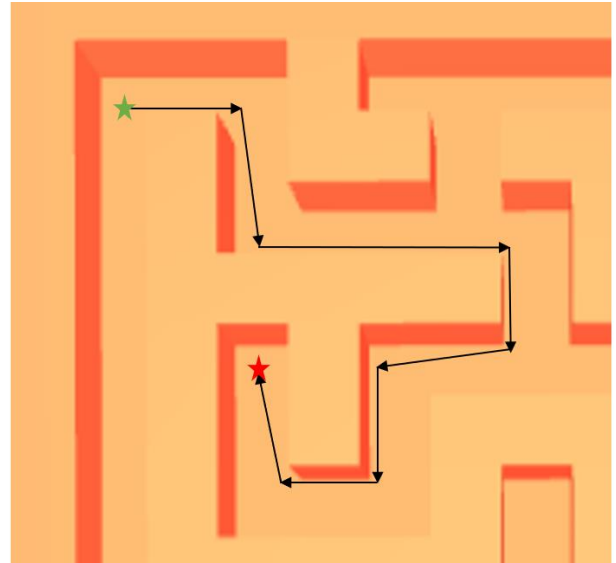
Fig. 3. A path of the DistBug in the maze

6. Conclusions

The paper presented an implementation of DistBug algorithm in the ROS Noetic environment. The evaluation was performed in two different types of environment and with a set of different target points. After reaching the stop condition, the control node logged the metrics, which included a flag of arrival to the target, a trajectory of the robot, a total time of the algorithm execution, and a number of encountered on the way obstacles. The tests demonstrated 97% of success. The implemented method can be adapted for other robot models and environments by empirically adjusting the timeout and the stopping threshold, while setting corresponding to the robot sensor range and the step size.

Acknowledgements

References

1. A. Zakiev, R. Lavrenov, E. Magid and V. Indelman, Path planning for Indoor Partially Unknown Environment Exploration and Mapping. ICAROB 2018: Proceedings of the 2018 International Conference on Artificial Life and Robotics, 2018, p. 399-402.
2. B. K. Patle et al, A review: On path planning strategies for navigation of mobile robot. Defence Technology, vol. 15, 2019, pp. 582-606.
3. E. Magid, R. Lavrenov and A. Khasianov, Modified spline-based path planning for autonomous ground vehicle. Proceedings of the 14th International Conference on Informatics in Control, Automation and Robotics (ICINCO 2017), vol. 2, 2017, p. 132-141.
4. H. Zhang, W. Lin and A. Chen, Path planning for the mobile robot: A review. Symmetry, vol. 10, 2018, pp. 450.

5. V. J. Lumelsky and T. Skewis, Incorporating range sensing in the robot navigation function. IEEE transactions on Systems, Man, and Cybernetics, vol. 20, 1990, pp. 1058-1069.

6. A. Sankaranarayanan and M. Vidyasagar, A new path planning algorithm for moving a point object amidst unknown obstacles in a plane. IEEE International Conference on Robotics and Automation, IEEE, 1990, pp. 1930-1936.

7. V. J. Lumelsky and A. A. Stepanov, Path-planning strategies for a point mobile automaton moving amidst unknown obstacles of arbitrary shape. Algorithmica, vol. 2, 1987, pp. 403-430.

8. E. Magid and E. Rivlin, CAUTIOUSBUG: A competitive algorithm for sensory-based robot navigation. 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No. 04CH37566), IEEE, vol. 3, 2004, pp. 2757-2762.

9. H. Noborio, Several path-planning algorithms of a mobile robot for an uncertain workspace and their evaluation. Proc. of the IEEE Int. Work. Intel. Motion Control, Istanbul, Turkey, vol. 1, 1990, pp. 289-294.

10. A. Yufka and O. Parlaktuna, Performance comparison of bug algorithms for mobile robots. Proceedings of the 5th international advanced technologies symposium, Karabuk, Turkey, 2009, pp. 13-15.

11. Y. Horiuchi and H. Noborio, Evaluation of path length made in sensor-based path-planning with the alternative following. Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation (Cat. No. 01CH37164), IEEE, vol. 2, 2001, pp. 1728-1735.

12. I. Kamon and E. Rivlin, Sensory-based motion planning with global proofs. IEEE transactions on Robotics and Automation, vol. 13, 1997, pp. 814-822.

13. A. Apurin, B. Abbyasov, E.A.Martínez-García and E. Magid, Comparison of ROS Local Planners for a Holonomic Robot in Gazebo Simulator. International Conference on Interactive Collaborative Robotics, Cham: Springer Nature Switzerland, 2023, pp. 116-126.

14. A. Dobrokvashina, R. Lavrenov, E. Magid, Y. Bai and M. Svinin, How to Create a New Model of a Mobile Robot in ROS/Gazebo Environment: An Extended Tutorial. International Journal of Mechanical Engineering and Robotics Research, vol. 12, 2023, pp. 192-199.

15. R. Safin, R. Lavrenov, K.-H. Hsia, E. Maslak, N. Schiefermeier-Mach and E. Magid, Modelling a TurtleBot3 Based Delivery System for a Smart Hospital in Gazebo. The 15th Siberian Conference on Control and Communications (SIBCON 2021), 2021, № 9438875.

16. R. Safin, T. Tsoy, R. Lavrenov, I. Afanasyev and E. Magid, Modern Methods of Map Construction Using Optical Sensors Fusion. International Conference on Artificial Life and Robotics (ICAROB 2023), 2023, pp. 167-170.

17. P. Marin-Plaza et al, Global and local path planning study in a ROS-based research platform for autonomous vehicles, Journal of Advanced Transportation, 2018, pp. 1-10.

18. B. Abbyasov, K. Kononov, T. Tsoy, E. A. Martinez-Garcia and E. Magid, Experience in Efficient Real Office Environment Modelling in Gazebo: a Tutorial. International Conference on Artificial Life and Robotics (ICAROB 2022), 2022, p. 673-677.

19. H. Maghfiroh and H. P. Santoso, Online Navigation of Self-Balancing Robot using Gazebo and RVIZ. Journal of Robotics and Control (JRC), 2(5), 2021.

20. Fabio Ugalde Pereira, Pedro Medeiros de Assis Brasil, Marco Antonio De Souza Leite Cuadros and Anselmo Rafael Cukla, Analysis of Local Trajectory Planners for Mobile Robot with Robot Operating System, IEEE Latin America Transactions, vol. 20, 2022, pp. 92-99.

21. N. James and T. Bräunl, Performance comparison of bug navigation algorithms, Journal of Intelligent and Robotic Systems, vol. 50, 2007, pp. 73-84.

22. A. Dobrokvashina, S. Sulaiman, T. Gamberov, K.-H. Hsia and E. Magid, New Features Implementation for Servosila Engineer Model in Gazebo Simulator for ROS Noetic. International Conference on Artificial Life and Robotics (ICAROB 2023), 2023, pp. 154-157.

23. R. Lavrenov, A. Zakiev and E. Magid, Automatic mapping and filtering tool: From a sensor-based occupancy grid to a 3D Gazebo octomap. International Conference on Mechanical, System and Control Engineering (ICMSC), 2017, p. 190-195.

24. A. Iskhakova, B. Abbyasov, T. Tsoy, T. Mironchuk, M. Svinin and E. Magid, LIRS-Mazegen: An Easy-to-Use Blender Extension for Modeling Maze-Like Environments for Gazebo Simulator. Frontiers in Robotics and Electromechanics, Vol. 329, 2023, pp. 147-161.

25. B. Abbyasov, R. Lavrenov, A. Zakiev, K. Yakovlev, M. Svinin and E. Magid, Automatic Tool for Gazebo World Construction: From a Grayscale Image to a 3D Solid Model. International Conference on Robotics and Automation (ICRA), 2020, p. 7226-7232.

Authors Introduction

Mr. Alexander Pak
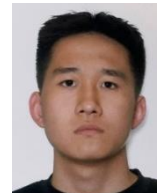
He is a third-year bachelor student of Computer Science and Engineering at HSE University, Russia.

Mr. Alexander Eryomin

He received a BSc degree in Software Engineering from the School of Space and Information Technology at the Siberian Federal University in 2022. Currently, he is a second-year student of Master program in Intelligent Robotics at the Institute of Information Technology and Intelligent Systems at the Kazan Federal University, Russia.

Ms. Tatyana Tsoy

In 2012 she graduated from the University of Tsukuba. Since 2018 she has been a PhD student in Robotics at the Institute of Information Technology and Intelligent Systems of Kazan Federal University, Russia.

On sensor modeling in Gazebo simulator

Niez Yuldashev
Intelligent Robotics Department, Kazan Federal University, 420008, Kazan, Russian Federation

Alexandra Dobrokvashina
Intelligent Robotics Department, Kazan Federal University, 420008, Kazan, Russian Federation

Roman Lavrenov
Intelligent Robotics Department, Kazan Federal University, 420008, Kazan, Russian Federation
Email: NBYuldashev@stud.kpfu.ru, dobrokvashina@it.kfu.ru, lavrenov@it.kfu.ru

Abstract

Sensor modeling in the Gazebo simulator is fundamental to robotics advancement. This review explores sensor modeling intricacies, methodologies, and applications, while emphasizing a critical role of a precise sensor modeling. Application scenarios demonstrate a sensor modeling's broad utility across fields including medical diagnostics, autonomous navigation, and industrial automation. Differences in research focus, methodology, and implementation underline a varied nature of sensor modeling studies. Key challenges include a need for more detailed world models. The paper guides a research in sensor modeling and identifies crucial questions.

Keywords: Overview, Gazebo, Sensor modeling, Sensor simulation.

1. Introduction

In contemporary robotics and autonomous systems, simulation plays an important role [1]. It offers researchers and engineers the means to create and test robots in virtual environments, reducing costs and ensuring safety during the development of robotic systems [2]. However, for simulations to produce realistic results, precise modeling of sensors, such as cameras, lidars, infrared, and ultrasonic sensors, is paramount. These sensors are instrumental in perceiving the robot's surroundings and interacting with them. Accurate sensor models in simulators have become fundamental for the successful development and testing of perception, navigation, and control algorithms for robots.

In this review article, we delve into a significant issue within the realm of robotic simulation, focusing on sensor modeling in the Gazebo simulator [3]. Among robotic simulators, Gazebo stands out for its powerful functionality and flexibility. It allows for the creation of complex 3D robot models and their environments, along with the simulation of a wide range of sensors [4].

The objective of our article is to provide readers with an overview of contemporary methodologies and approaches to sensor modeling within the Gazebo simulator. We will explore the challenges researchers address in this domain, the methods employed for precise sensor modeling, and the profound impact these models have on the development of robotic systems. Moreover, we will discuss the primary challenges researchers encounter and outline potential directions for future research in this rapidly evolving field.

2. Methodologies

The choice of a specific method and methodology for modeling sensors in the Gazebo simulator depends on several key factors. Initially, this is the type of sensor that is required to simulate, since different sensors have different characteristics and requirements. For example, the modeling of the camera requires accounting for optical parameters [5], while the modeling of the ultrasonic sensor can be based on acoustic properties [6]. The objectives of the study also play an important role in choosing a methodology. If the main goal is to assess the accuracy of the sensor, then the methods focused on comparison with real dimensions can be the most suitable [7]. If the goal is to develop and test new algorithms for processing data from the sensor, then modeling taking into account noise and distortion can be more relevant [8]. The level of detail also plays a key role. Highly detailed models of sensors may require large

computing resources and time for simulation, while simplified models can be less accurate, but more effective in terms of computing resources [9].

Some of the common methodologies include: Physical modeling, including taking into account the physical properties of the sensor and its interaction with the environment [10]. Methods of machine learning that allow training sensors based on real data and apply it in simulations [11]. Geometric modeling, focused on the exact reconstruction of the geometry of the environment and objects with which the sensor interacts [12]. Statistical modeling methods: include statistical approaches to take into account various sources of noise and uncertainty in the measurements of sensors [13]. This may include methods for filtering, smoothing and assessing the state that allow you to take into account the statistical characteristics of measurements.

Table 1 provides an overview of recent and noteworthy publications related to sensor modeling within the Gazebo simulator. This table showcases articles that have leveraged Gazebo for simulating sensors, highlighting the various contexts and applications where these models have been employed.

3. Common aspects

One of the key tasks of research involving sensor modeling is to evaluate and analyze the accuracy of various sensors, such as cameras, lidars, ultrasonic and infrared sensors, in a virtual environment. This allows researchers to understand how the sensors interact with the environment and how accurate the data they provide [5], [6].

Research in this area often involves the creation and testing of data processing algorithms for sensors. These algorithms may include computer vision methods, image processing, machine learning, and other techniques used to analyze data collected from sensors [14].

Gazebo's sensor modeling research aims to create virtual scenarios that are as close to real-world conditions as possible. Researchers strive to accurately model the objects with which sensors interact and whose condition can affect the simulated sensor readings [6], [15].

Researchers often use sensor simulations to validate the algorithms of their robots. This allows them to test the validity of their designs under different conditions. Research may also involve integrating sensor modeling into real robotic platforms. This helps researchers adapt the results of their work for use in actual robots [15], [16].

In summary, research related to sensor modeling in the Gazebo simulator aims to improve the understanding of sensor behavior in robotics systems, develop efficient data processing algorithms, and create realistic simulation environments for testing and training in the field of autonomous systems and robotics.

4. Distinctions

Each scientific paper can focus on different aspects of sensor modeling. For instance, one paper may investigate camera modeling for robot navigation [17], while another may emphasize lidar sensors for obstacle detection [16]. Consequently, the research topic can substantially differ.

Researchers can use various methods and techniques for sensor modeling. For example, some may develop their own sensor models [18], [19], while others might utilize standard tools and libraries available in Gazebo [20]. This can impact the accuracy of modeling and the scope of the results.

Studies may also differ in terms of research objectives. Some papers may aim to optimize sensor performance [21], while others seek to develop new data processing algorithms or assess sensor reliability in various scenarios [22].

Different studies may be oriented toward various types of robotic systems. For example, research may target mobile robots, autonomous vehicles [23], robot manipulators [6], [24], or drones [25], and each of these system types may have unique characteristics in sensor modeling.

Scientific papers can be directed toward diverse applications of sensor modeling, such as medical diagnostics [26], autonomous navigation [17], [27], and more.

These differences in research themes, methods, and objectives make each scientific paper unique and specialized. Collectively, they contribute to the advancement of sensor modeling in the Gazebo simulator, providing new knowledge and tools for developers and researchers in the field of robotics.

5. Application scenarios

Modeling sensors in the Gazebo simulator finds applications in various fields of robotics and autonomous systems.

One of the key applications of sensor modeling in Gazebo is the development and testing of autonomous transportation systems, such as self-driving cars [30]. Virtual simulations enable engineers to assess navigation and control algorithms under diverse conditions, ensuring safety before these vehicles hit the roads.

Table 1. Relevant publications in Gazebo sensors simulation.

Sensor modeling context	Sensor type	Application	Reference
A software package for ultrasound sensor modeling in the Gazebo simulator	Ultrasound sensor	Medical Robotics	[6]
Incorporating sensors into the model	LRF, IMU sensor, cameras	USAR, AMR	[15]
Incorporating sensors into the model	LIDAR, IMU, GPS/GLONASS, Side sonars	Autonomous Vehicles	[23]
It is explained how to incorporate sensors into models in Gazebo	LRF, IMU sensor, cameras	Robotics Education, AMR	[28]
Shows physical integration of the main types of sensors in UAV domain both for navigation and collision avoidance	LIDAR, sonar, radar, video	Autonomous Vehicles, UAV	[16]
Automatic tool allows creating realistic landscapes in Gazebo simulation based on results of real world sensor-based exploration	LIDAR, RGBD-camera	UAV, UGV	[12]
The simulated tactile sensor can produce high-resolution images from depth-maps captured by a simulated optical sensor	GelSight tactile sensor	Tactile Sensing Visualization	[18]
Simulating the operation of LIDAR sensors in real-time, considering LASER beam propagation and attenuation of LASER energy in adverse weather conditions such as fog and rain	LIDAR	Automotive	[7]
An algorithm for detection of obstacles and lines was implemented for the IR sensors	IR	Robotics Education	[20]
The scalability investigation of LIDAR-type sensor simulation in ROS-supported Gazebo	LIDAR	AMR	[29]
Evaluate a variety of visual and LiDAR SLAM algorithms in a simulated environment	LIDAR, RGB-D-camera	Mobile Robots	[21]
Simulation of the RFID readers-antennas mounted on robots reading RFID tags in the environment	RFID	UAV	[19]
A method to simulate the ARVA transceiver	ARVA	Search And Rescue, UAV	[10]
Test the control-level and sensor-level algorithms in simulation environment	LIDAR	UAV	[25]

In the medical domain, Gazebo's sensor modeling capabilities can be used to create virtual training environments [31]. These environments allow medical professionals to practice operating medical robots and other devices [6]. Additionally, it is beneficial for research in the field of medical robotics, including surgical robots and rehabilitation devices.

Gazebo is frequently used in the educational field to teach students and researchers the fundamentals of robotics [28]. Sensor modeling permits students to experiment with various sensor types and data processing algorithms, enhancing their understanding of robotics principles.

Sensor modeling in Gazebo is invaluable for simulating complex urban search and rescue (USAR) scenarios [15]. It allows researchers and first responders to test various robotic systems that can navigate disaster-stricken areas, locate survivors, and assess structural damage, all while avoiding risks to human life.

These applications showcase the versatility of sensor modeling in the Gazebo simulator. They represent only a fraction of the possibilities offered by sensor modeling, with its utility extending to various challenges and

objectives in the realms of robotics and autonomous systems.

6. Main challenges and future directions

While sensor modeling in the Gazebo simulator has seen remarkable advancements, several challenges persist, and promising future directions are emerging. Here, we outline some of the key challenges and areas of potential growth.

Many studies have identified the need for further refinement of the world models used in simulations [6], [15]. Realism of sensor simulations often depends on the complexity of the environment being simulated. Future research may focus on enhancing world models to better mimic real-world scenarios, including factors like lighting, weather, and dynamic objects.

Some articles focused on the complexity of sensory synthesis [32]. As robotic systems become more sophisticated, sensor fusion, the integration of data from multiple sensors, poses a growing challenge. Future developments may revolve around creating more

advanced sensor fusion models to handle the increasing complexity of robotic applications.

Balancing realism and computational efficiency remains a fundamental challenge. Simulations with high levels of detail can be computationally intensive, limiting real-time performance [29], [33]. Researchers may explore novel techniques to achieve a harmonious balance between realism and efficiency.

Leveraging AI and machine learning for sensor modeling is a burgeoning area [34]. Future research may delve into integrating AI techniques for sensor data interpretation, enabling robots to learn from simulated sensor inputs.

As robots increasingly interact with humans, modeling sensors for human-awareness becomes vital. Future work may focus on simulating sensors that enable robots to perceive and respond to human actions and intentions.

With the rise of autonomy in robotics, future directions include the development of sensor models that support high-level decision-making in autonomous systems.

In summary, sensor modeling in Gazebo continues to evolve, aiming for realism, efficiency, and applicability. Addressing these challenges and exploring these future directions will contribute to the continued growth of robotics and autonomous systems.

7. Conclusion

In conclusion, this review article has explored the intricate world of sensor modeling within the Gazebo simulator. We've delved into the aspects of sensor modeling, discussed current research directions, and identified key research questions in this evolving field. Through an examination of various methodologies, application scenarios, and the distinctiveness of each research endeavor, we've gained a comprehensive understanding of how sensor modeling contributes to the advancement of robotic systems. As the field of robotics continues to expand, the importance of accurately modeling sensors in virtual environments has become increasingly clear. This review serves as a foundational resource for further research and innovations in the domain of sensor modeling within the Gazebo simulator.

Acknowledgements

References

1. K. Liu, D. Negrut, The Role of Physics-Based Simulators in Robotics, Annual Review of Control, Robotics, and Autonomous Systems, 4, 2021, pp 35-58.
2. K. Fernandez, The use of computer graphic simulation in the development of robotic systems, Acta Astronautica, 17(1), 1988, pp. 115-122.
3. N. Koenig, A. Howard, Design and use paradigms for Gazebo, an open-source multi-robot simulator, International Conference on Intelligent Robots and Systems (IROS), 3, 2004, pp. 2149-2154.
4. R. Lavrenov, A. Zakiev, Tool for 3D Gazebo map construction from arbitrary images and laser scans, International Conference on the Developments in eSystems Engineering, 2017, pp. 256-261.
5. T. Tsoy, R. Safin, E. Martinez-Garcia, S. Roy, S. Saha, E. Magid, Exhaustive simulation approach for a virtual camera calibration evaluation in Gazebo, International Conference on Automation, Robotics and Applications (ICARA), 2022, pp. 233-238.
6. B. Abbyasov, A. Dobrokvashina, R. Lavrenov, E. Kharisova, T. Tsoy, L. Gavrilova, S. Bulatov, E. Maslak, N. Schiefermeier-Mach, E. Magid, Ultrasound sensor modeling in Gazebo simulator for diagnostics of abdomen pathologies, International Siberian Conference on Control and Communications (SIBCON), 2021, pp. 1-6.
7. M. Hadj-Bachir, P. Souza, LIDAR sensor simulation in adverse weather condition for driving assistance development, 2019.
8. A. Stefanov, Distortion analysis of underwater acoustic sensor networks, International Conference on New Technologies, Mobility and Security (NTMS), 2015, pp. 1-4.
9. K. Demarco, M. West, A. Howard, A computationally-efficient 2D imaging sonar model for underwater robotics simulations in Gazebo, OCEANS 2015 - MTS/IEEE Washington, 10, 2015, pp. 1-7.
10. W. Zhao, J. Queralta, T. Westerlund, Sim-to-Real Transfer in Deep Reinforcement Learning for Robotics: a Survey, IEEE Symposium Series on Computational Intelligence (SSCI), 2020, pp. 737-744.
11. R. Lavrenov, A. Zakiev, E. Magid, Automatic mapping and filtering tool: From a sensor-based occupancy grid to a 3D Gazebo octomap, International Conference on Mechanical, System and Control Engineering (ICMSC), 2017, pp. 190-195.
12. N. Hirsenkorn, T. Hanke, A. Rauch, B. Dehlink, R. Rasshofer, E. Biebl, Virtual sensor models for real-time applications, Advances in Radio Science, vol. 14, 2016, pp. 31-37.
13. A. Buyval, I. Afanasyev, E. Magid, Comparative analysis of ROS-based Monocular SLAM methods for indoor navigation, International Conference on Machine Vision (ICMV), 2017.
14. A. Dobrokvashina, R. Lavrenov, E. Martinez-Garcia, Y. Bai, Improving model of crawler robot Servosila "Engineer" for simulation in ROS/Gazebo, International Conference on Developments in eSystems Engineering (DeSE), 2020, pp. 212-217.
15. J. García, J. Molina, Simulation in real conditions of navigation and obstacle avoidance with PX4/Gazebo platform, IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), 2019, pp. 979-984.

16. M. Mustafin, T. Tsoy, E. Martınez-Garcıa, R. Meshcheryakov, E. Magid, Modelling mobile robot navigation in 3D environments: camera-based stairs recognition in Gazebo, Moscow Workshop on Electronic and Networking Technologies (MWENT), 2022, pp. 1-6.

17. D. Gomes, P. Paoletti, S. Luo, Generation of GelSight Tactile Images for Sim2Real Learning, IEEE Robotics and Automation Letters, 6(2), 2021, pp. 4177-4184.

18. A. Alajami, R. Pous, G. Moreno, Simulation of RFID Systems in ROS-Gazebo, International Conference on RFID Technology and Applications (RFID-TA), 2022, pp. 113-116.

19. A. Rafael, C. Santos, D. Duque, S. Fernandes, A. Sousa, L. Reis, Development of an AlphaBot2 Simulator for RPi Camera and Infrared Sensors, Robot 2019: Fourth Iberian Robotics Conference, 2019, pp. 502-514.

20. R. Giubilato, A. Masili, S. Chiodini, M. Pertile, S. Debei, Simulation Framework for Mobile Robots in Planetary-Like Environments, International Workshop on Metrology for AeroSpace (MetroAeroSpace), 2020, pp. 594-599.

21. R. Lavrenov, F. Matsuno, E. Magid, Modified spline-based navigation: Guaranteed safety for obstacle avoidance, Interactive Collaborative Robotics, 2017, pp. 123-133.

22. I. Shimchik, A. Sagitov, I. Afanasyev, F. Matsuno, E. Magid, Golf cart prototype development and navigation simulation using ROS and Gazebo, MATEC Web of Conferences, vol. 75, 2016.

23. A. Klimchik, E. Magid, S. Caro, K. Waiyakan, A. Pashkevich, Stiffness of serial and quasi-serial manipulators: comparison analysis, MATEC Web of Conferences, vol. 75, 2016.

24. J. Zhu, C. Xu, A comprehensive simulation testbench for aerial robot in dynamic scenario using gazebo-ros, Chinese Automation Congress (CAC), 2017, pp. 7664-7669.

25. A. Liu, F. Tendick, K. Cleary, C. Kaufmann, A Survey of Surgical Simulation: Applications, Technology, and Education, Presence: Teleoperators and Virtual Environments, 12, 2003, pp. 599-614.

26. Y. Alborzi, B. Jalal, E. Najafi, ROS-based SLAM and Navigation for a Gazebo-Simulated Autonomous Quadrotor, International Conference on Research and Education in Mechatronics (REM), 2020, pp. 1-5.

27. A. Dobrokvashina, R. Lavrenov, E. Magid, Y. Bai, M. Svinin, How to Create a New Model of a Mobile Robot in ROS/Gazebo Environment: An Extended Tutorial, International Journal of Mechanical Engineering and Robotics Research, 12(4), 2023, pp. 192-199.

28. A. Saglam, Y. Papelis, Scalability of sensor simulation in ROS-Gazebo platform with and without using GPU, Spring Simulation Conference (SpringSim), 2020, pp. 1–11.

29. P. Kaur, S. Taghavi, Z. Tian, W. Shi, A Survey on Simulators for Testing Self-Driving Cars, International Conference on Connected and Autonomous Driving (MetroCAD), 2021, pp. 62-70.

30. M. Tiryaki, Ö. Erin, M. Sitti, A Realistic Simulation Environment for MRI-Based Robust Control of Untethered Magnetic Robots With Intra-Operational Imaging, IEEE Robotics and Automation Letters, 5(3), 2020, pp. 4501-4508.

31. E. Cheborateva, R. Safin, K. Hsia, A. Carballo, E. Magid, Person-Following Algorithm Based on Laser Range Finder and Monocular Camera Data Fusion for a Wheeled Autonomous Mobile Robot, Interactive Collaborative Robotics, 2020, pp. 21-33.

32. J. Borrego, R. Figueiredo, A. Dehban, P. Moreno, A. Bernardino, J. Santos-Victor, A generic visual perception domain randomisation framework for Gazebo, IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC), 2018, pp. 237-242.

33. K. Chin, T. Hellebrekers, C. Majidi, Machine Learning for Soft Robotic Sensing and Control, Advanced Intelligent Systems, 2(6), 2020.

34. M. Mustafin, E. Chebotareva, E. Magid, H. Li, E. Martinez-Garcia, Features of Interaction Between a Human and a Gestures-Controlled Collaborative Robot in an Assembly Task: Pilot Experiments, International Conference on Artificial Life and Robotics (ICAROB), 28, 2023, pp. 157-161.

Authors Introduction

Mr. Niez Yuldashev

Currently he is a master student at Institute of Information Technology and Intelligent Systems, Kazan Federal University, Russia.

Ms. Alexandra Dobrokvashina
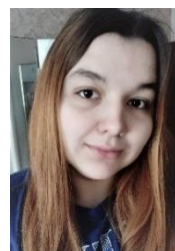
She received a master degree at Institute of Information Technology and Intelligent Systems, Kazan Federal University in 2021. Currently she works as a research assistant at the Laboratory of Intelligent Robotic Systems (LIRS) at Kazan Federal University, Russia.

Assistant Professor Roman Lavrenov
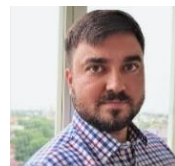
Currently he works as an Assistant Professor at Kazan Federal University. He obtained his Ph.D. at Kazan Federal University in 2020. He teaches the following master's courses: «Robotic Operating System», «Industrial robotics» and others. Author of an online course and methodological guide on the Robotic Operating System (ROS). His main interests are path planning, robot operating system, and autonomous mobile robots.

An Overview of Kinect Based Gesture Recognition Methods

Alexander Alexeev

Institute of Information Technology and Intelligent Systems, Kazan Federal University, Kazan, Russia

Tatyana Tsoy

Laboratory of Intelligent Robotics Systems (LIRS), Institute of Information Technology and Intelligent Systems, Kazan Federal University, Kazan, Russia

Edgar A. Martínez-García

Robotics Laboratory, Department of Industrial Engineering and Manufacturing, The Autonomous University of Ciudad Juarez, Chihuahua, Mexico

Evgeni Magid

Laboratory of Intelligent Robotics Systems (LIRS), Institute of Information Technology and Intelligent Systems, Kazan Federal University, Kazan, Russia
HSE University, Moscow, Russia
E-mail: san73rus@gmail.com, tt@it.kfu.ru, edmartin@uacj.mx, magid@it.kfu.ru

Abstract

Visual sensors play an important role in a broad variety of robotic systems applications. Even though Kinect technology appeared over 10 years ago, Kinect sensors are still actively employed by researchers around the world. This paper presents an overview of Kinect and Kinect 2 sensors' applications in a human gesture based control. We analyzed existing research papers to estimate a popularity of particular feature extraction and gesture recognition methods, recommendations on a distance between an object of interest and a sensor, reported accuracy and latency of the sensor. Our analysis is supposed to facilitate a selection of a suitable combination of methods for a particular application of Kinect sensor in gesture recognition while considering its performance.

Keywords: Microsoft Kinect, robotics, gesture control, gesture recognition.

1. Introduction

Robotics has achieved a significant progress, enabling robots to autonomously plan routes [1] and interact with humans in medicine [2], education [3], and rescue operations [4]. Controlling robots with a voice and gestures is the most natural way of interaction [5] due to daily habits of using these methods by humans. Control using gestures can be implemented with vision sensors [6], electromyography methods (that track human muscle contractions [7]), employing a touch screen, an accelerometer or other sensors [8]. This paper overviews gesture control implementations in robotics using Kinect sensor. Kinect is originally a Microsoft motion controller for gaming that uses computer vision methods for a gesture based control. Later Microsoft enabled a custom application development with Software Development Kit for Windows [9]. This study summarize information about Microsoft Kinect use in robotics, compares accuracy of different gesture recognition methods and emphasizes a reported latency of Kinect in various tasks.

2. Kinect in Robotics

Kinect sensor (Fig. 1a) is a device introduced by Microsoft in 2010 for XBOX 360 gaming console control. A combination of an infrared sensor and RBG-D camera makes Kinect suitable for human body tracking purposes. Software Development Kit [10] allows developing a custom software for a variety of tasks including gesture based robot control. The introduced by Microsoft in 2013 Kinect v2 was (Fig. 1b) a second generation of Kinect device that also combines RGB-D and infrared sensors. Similar to its predecessor, Kinect v2 is intended for a gesture-based control. However, due to technical improvements such as adding a wide angle time of flight (ToF) camera [11] and upscaling a color camera resolution to 1080p [12], it exhibits a better performance compared to the first generation, which is demonstrated in Section 4.

Alexander Alexeev, Tatyana Tsoy, Edgar A. Martínez-García, Evgeni Magid

a. Kinect

b. Kinect v2

Fig. 1. Microsoft Kinect first generation and Kinect v2.

Kinect and Kinect v2 capture color and depth data of a scene and are broadly used in computer vision for object recognition including gesture recognition. The distinctive feature of Kinect family is its versatility compared to counterpart sensors. Leap Motion Controller is also used in robotics for gesture-based robot control [13], [14]. Yet, this device is designed to recognize only hand movements, whereas Kinect does not have such limitations and can be employed for a full-body gesture recognition [15] and a face tracking [16]. In addition, Kinect v2 contains an embedded microphone that enables a voice control [17], which can augment a gesture-based control. Leap Motion Controllers could be used in pairs, e.g., to recognize each hand separately [18], which may increase the system's load. Similarly, Kinect sensors can form complicated systems for tracking and motion capture system purposes [19], [20]. However, in some cases Kinect sensor may not be a right solution, e.g., authors in [21] complained that Kinect is not compatible with Linux systems as Kinect's Software Development Kit has only Windows OS support. To solve this issue, the researchers employed Asus Xtion Pro Live with an RGB-D camera similar to Kinect.

In robotics, Kinect could be used to control different types of real robots. In [17] authors presented a Kinect-based industrial robot control and demonstrated that Kinect could be successfully integrated into simulators. In [22] Kinect was applied for a remote control of a mechanical arm that replicates operator-defined gestures. A gesture-based control using Kinect could also be implemented in mobile robotics, where Kinect serves as a robot control panel [23].

3. Gesture Recognition Methods

Choosing a gesture recognition method for Kinect based control is important, as it directly affects an overall system performance. If a gesture recognition method is not accurate enough, not suitable for a particular task or

a recognition system demonstrates significant delays, a different method of control should be considered. In this section, we summarize information about methods of Kinect based control and emphasize their accuracy.

Table 1 presents an analysis of studies that used various combinations of feature extraction and gesture recognition methods. The first column of the table describes a particular task; the second column *Dist* specifies a recommended optimal distance between an object of interest (gesture producer) and a sensor in meters; the third and the forth – a feature extraction *FEM* and a gesture recognition methods *GRM*, respectively. The fifth column *Acc* shows an accuracy of each approach in percents (as reported by its authors); the sixth *Skeleton* emphasizes if an approach employs a skeleton extraction; the last one *Rf* refers to a corresponding paper. N/A in the table denotes that a paper does not contain corresponding data.

As Table 1 shows, a combination of the Histogram of Oriented Gradients (HOG) for feature extraction and Support Vector Machine (SVM) for gesture recognition ([24], [33]) is a popular approach among researchers. The same combination was used at an earlier stage of [26] research, however, a deep learning AlexNet model eventually was chosen due to a better performance. In [31] the authors emphasized an influence of a distance at which gestures were recognized by Kinect on a recognition accuracy. The highest accuracy of 87% was achieved at a distance of 3 meters from the Kinect, while the lowest accuracy of gesture recognition of 25% was obtained at a distance of 5 meters from the sensor. The importance of a classifier selection when controlling a robot with user-defined gestures was confirmed in [23], where the SURF and FLANN methods were employed, to extract arbitrary points from an image and to recognize gestures. During the experiments, it was discovered that the FLANN library failed to achieve required results in recognizing gestures as it could not match corresponding features in a variety of gestures.

Majority of studies in Table 1 used a method of a skeleton construction, where a skeleton consists of joints of a human body or hands only. This approach allows a more accurate extraction of special points, which are further required for a gesture recognition. An alternative approach built a 3D model of a hand [34]; yet the authors did not provide accuracy indicators for the gesture recognition, and thus a usability of the method for a real gesture recognition system is questionable.

4. Kinect Latency

Latency has a significant importance when operating technical devices. Latency can be caused by hardware

Table 1. Gesture Recognition Methods

Task	Dist, m	FEM	GRM	Acc, %	Skeleton	Rf
Static hand gesture recognition	N/A	Histogram of Oriented Gradients	Support Vector Machine	98.3	No	[24]
Hand gesture recognition	1.2–3.5	Skeleton Extraction	Artificial Neural Network	97.8	Yes	[25]
Touchless visualization of 3D medical images	2.5-3.5	Histogram of Oriented Gradients	AlexNet (CNN)	96.5	No	[26]
Hand gesture recognition	N/A	MediaPipe Palm Detector	MediaPipe Gesture Recognition	95.7	Yes	[27]
Sign language recognition for Arabic speakers	N/A	Used (not specified which one)	RandomForest Classifier + Ada-Boosting	93.7	Yes	[28]
Dynamic hand gesture recognition	N/A	N/A	Dynamic Time Warping	92	Kinect embedded	[29]
Full body gesture recognition	N/A	CNN	Fast Dynamic Time Warping + CNN	90.8	Yes	[30]
Smart home control	3 m	Self-developed	Self-developed	87	Yes	[31]
Sign language recognition	N/A	Histogram of Oriented Gradients	Dynamic Time Warping	86	Yes	[32]
Hand gesture recognition	2.5-3 m	Histogram of Oriented Gradients	Dynamic Time Warping	76.7	Yes	[33]

and software factors. Moreover, a task-dependent software delay sums up with a hardware delay, which leads to a decrease in device performance. This section discusses a latency between a user input and an expected system output, which differ for a particular task.

Table 2. Latency Measurements of Kinect

Task	Latency, ms	Ref
Kinect		
Skeleton detection	106-500	[35]
Human fall incident detection	300	[36]
Contactless hand tracking for surgical robot	89-576	[37]
Skeleton position estimation	100	[38]
Skeleton detection	200-400	[39]
Kinect v2		
Human gait analysis	29-127	[40]
Human gait analysis	200	[41]
Skeleton detection	66-100	[42]
IRB4400 industrial robot control	20 (only a hardware latency)	[43]
Rat behavior tracking	80-126	[44]
3D content capture	65-67	[45]

Table 2 presents ab information about latency measurements of both generations of Kinect devices employed in various tasks. Note that reference papers of Table 1 and Table 2 are different since Table 1 papers did not specify a latency as they considered a different aspect of Kinect usage.

In [41] authors emphasized a 20 ms hardware latency as a minimum latency level of the device itself, which cannot be reduced. While official manufacture stated latency data for the first generation of Kinect is not available, the study [36] reports the latency value of 100 ms. Based on the surveyed in Table 2 papers we concluded that the average latency for the first generation Kinect is 267 ms, while the average latency for Kinect v2 is 123 ms. However, even though the Kinect v2's average latency significantly improved over the first generation, its effectiveness still depends on a specific task in which it is used and could achieve up to 200 ms.

5. Conclusions

This paper presents an overview of Microsoft Kinect sensor applications for gesture based control in robotics. Kinect first generation and Kinect v2 related research papers were analyzed in terms of range and accuracy in a number of typical tasks that require a sensor with RBG-D capabilities. Based on the survey we concluded that the average overall latency of the first generation Kinect is 267 ms, while the average latency for Kinect v2 is 123 ms. However, even though the Kinect v2's average latency has improved over the first generation, its effectiveness naturally depends on a specific task and could achieve up to 200 ms.

References

1. Y. Bai, On Motion Planning and Control for Partially Differentially Flat Systems. Robotica, vol. 39, no. 4, 2021, pp. 718-734.
2. B. Abbyasov, Ultrasound sensor modeling in Gazebo simulator for diagnostics of abdomen pathologies. 2021 International Siberian Conference on Control and Communications (SIBCON), Kazan, 2021, pp. 1-6.
3. L. Gavrilova, Facilitating a preparatory stage of real-world experiments in a humanoid robot assisted English language teaching using Gazebo simulator. 2020 13th International Conference on Developments in eSystems Engineering (DeSE), Liverpool, December 2020, pp. 222-227.
4. E. Magid, Artificial intelligence based framework for robotic search and rescue operations conducted jointly by international teams. Proceedings of 14th International Conference on Electromechanics and Robotics "Zavalishin's Readings" ER (ZR), 2019, pp. 15-26.
5. J. Berger and L. Shuang, Review of Interfaces for Industrial Human-Robot Interaction. Current Robotics Reports, vol. 1, 2020, pp. 27-34.
6. M. Mustafin, Features of Interaction Between a Human and a Gestures-Controlled Collaborative Robot in an Assembly Task: Pilot Experiments. The 2023 International Conference on Artificial Life and Robotics (ICAROB2023), February 2023, pp. 157-161.
7. X. Zhang, A Framework for Hand Gesture Recognition Based on Accelerometer and EMG Sensors. IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, vol. 41, no. 6, 2011, pp. 1064-1076.
8. S. Berman and H. Stern, Sensors for Gesture Recognition Systems. IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and Reviews, vol. 42, no. 3, 2012, pp. 277-290.
9. Kinect – Starting to Develop with Kinect, Available at: https://learn.microsoft.com/en-us/archive/msdn-magazine/2012/june/kinect-starting-to-develop-with-kinect (accessed 29 November 2023).
10. Kinect for Windows, Available at: https://web.archive.org/web/20120103103154/http://www.kinectforwindows.org/ (accessed 10 May 2023).
11. A. Corti, A metrological characterization of the Kinect V2 time-of-flight camera. Robotics and Autonomous Systems, vol. 75, 2016, pp. 584-594.
12. L. Jamhoury, Understanding Kinect V2 Joints and Coordinate System, 2018, Available at: https://lisajamhoury.medium.com/understanding-kinect-v2-joints-and-coordinate-system-4f4b90b9df16 (accessed 29 November 2023).
13. A. Latif, Design and Control Of Autonomous Robot Using Gesture Based Intuitive Interaction. International Journal of Advanced Computer Systems and Software Engineering, vol. 1, no. 4, 2020, pp. 7-11.
14. D. A. Katsay, T. D. Isupova, and A. D. Biryukov, Non-contact control system for anthropomorphic robot. II International scientific and technical conference "Prom-Engineering", May 2016. pp. 441-444.
15. Y. Ma, Deep Learning-Based Upper Limb Functional Assessment Using a Single Kinect v2 Sensor. Sensors, vol. 20, no. 7, 2020,1903 p.
16. L. Vladutu and C. Marian, Framework for Posture and Face Recognition using Kinect an ambient-intelligence method. WSEAS Transactions on Computers, vol. 22, 2023, pp. 13-24.
17. W. Kaczmarek, Industrial Robot Control by Means of Gestures and Voice Commands in Off-Line and On-Line Mode. Sensors, vol. 20, no. 21, 2020, 6358 p.
18. W. Qi, Multi-Sensor Guided Hand Gesture Recognition for a Teleoperated Robot Using a Recurrent Neural Network. IEEE Robotics and Automation Letters, vol. 6, no. 3, 2021, pp. 6039-6045.
19. B. Gabbasov, I. Danilov, I. Afanasyev, and E. Magid, Toward a human-like biped robot gait: Biomechanical analysis of human locomotion recorded by Kinect-based motion capture system. Proceedings of the 10th International Symposium on Mechatronics and its Applications (ISMA15), 2015, pp. 1-6.
20. B. Gabbasov, I. Danilov, I. Afanasyev, and E. Magid, ZMP trajectory from human body locomotion dynamics evaluated by Kinect-based motion capture system. Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISAPP), 2016, pp. 160-166.
21. N. Pavón-Pulido, IoT Architecture for Smart Control of an Exoskeleton Robot in Rehabilitation by Using a Natural User Interface Based on Gestures. Journal of Medical Systems, vol. 44, 2020, pp. 1-10.
22. P. Tu and C. Huang, Mechanical arm teleoperation control system by dynamic hand gesture recognition based on kinect device. The Journal of Engineering, 2019, vol. 2019, no. 23, 2019, pp. 9110-9113.
23. N. Nikiforov, Pilot studies on Avrora Unior car-like robot control using gestures. Electromechanics and Robotics: Proceedings of 16th International Conference on Electromechanics and Robotics" Zavalishin's Readings" (ER (ZR) 2021), April 2021, pp. 271-283.
24. T. Ma and M. Guo, Research on Kinect-based gesture recognition. 2019 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), IEEE, September 2019, pp. 1-5.
25. B. Ganguly, Kinect Sensor Based Single Person Hand Gesture Recognition for Man–Machine Interaction. Computational Advancement in Communication Circuits and Systems: Proceedings of ICCACCS 2018, 2020, pp. 139-144.
26. J. Liu, An Improved Kinect-Based Real-Time Gesture Recognition Using Deep Convolutional Neural Networks for Touchless Visualization of Hepatic Anatomical Mode. Journal of Image and Graphics, vol. 7, no. 2, 2019, pp. 45-49.
27. M. Harris, Applying hand gesture recognition for user guide application using MediaPipe. 2nd International Seminar of Science and Applied Technology (ISSAT 2021), Atlantis Press, 2021, pp. 101-108.
28. B. Hisham and A. Hamouda, Supervised learning classifiers for Arabic gestures recognition using Kinect V2. SN Applied Sciences, vol. 1, 2019, pp. 1-21.
29. W. Lv, Gesture recognition in somatosensory game via Kinect sensor. Internet Technology Letters, 2021, 6 p.

30. M. Pfitscher, Activity gesture recognition on kinect sensor using convolutional neural networks and FastDTW for the MSRC-12 dataset. Intelligent Systems Design and Applications: 18th International Conference on Intelligent Systems Design and Applications (ISDA 2018), vol. 1, Springer International Publishing, December 2020, pp. 230-239.

31. H. Fakhrurroja, Hand State Combination as Gesture Recognition using Kinect v2 Sensor for Smart Home Control Systems. 2019 IEEE International Conference on Internet of Things and Intelligence System (IoTaIS), IEEE, November 2019, pp. 74-78.

32. C. Sun, T. Zhang, and C. Xu, Latent support vector machine modeling for sign language recognition with Kinect. ACM Transactions on Intelligent Systems and Technology (TIST), vol. 6, no. 2, 2015, pp. 1-20.

33. K. N. Krisandria, B. S. B. Dewantara, and D. Pramadihanto, HOG-based Hand Gesture Recognition Using Kinect. 2019 International Electronics Symposium (IES), IEEE, September 2019, pp. 254-259.

34. E. E. Shelomtsev, Solution to the problem of gesture recognition for controlling the behavior of a robot. Youth and modern information technologies: collection of papers from the 11th International Scientific and Practical Conference of Students, Postgraduates and Young Scientists, Tomsk Polytechnic University, November 2013, pp. 278-280.

35. M. A. Livingston, Performance measurements for the Microsoft Kinect skeleton. 2012 IEEE Virtual Reality Workshops (VRW), IEEE, March 2012, pp. 119-120.

36. A. A. M. Bigy, Recognition of postures and Freezing of Gait in Parkinson's disease patients using Microsoft Kinect sensor. 2015 7th International IEEE/EMBS Conference on Neural Engineering (NER), IEEE, April 2015, pp. 731-734.

37. Y. Kim, Kinect technology for hand tracking control of surgical robots: technical and surgical skill comparison to current robotic masters. Surgical endoscopy, vol. 28, 2014, pp. 1993-2000.

38. S. Feng, R. Murray-Smith, and A. Ramsay, Position stabilisation and lag reduction with Gaussian processes in sensor fusion system for user performance improvement. International Journal of Machine Learning and Cybernetics, vol. 8, 2017, pp. 1167-1184.

39. S. Park, An approach to the reduction of time delay in a networked performance. Proceedings of the Asia-Pacific Advanced Network, vol. 36, 2013, pp. 91-96.

40. D. Guffanti, The accuracy of the Microsoft Kinect V2 sensor for human gait analysis. A different approach for comparison with the ground truth. Sensors, vol. 20, no. 16, 2020, 4405 p.

41. A. Amini, K. Banitsas, and S. Hosseinzadeh, A New Technique for Foot-Off and Foot Contact Detection in a Gait Cycle Based on the Knee Joint Angle Using Microsoft Kinect v2. 2017 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI), IEEE, February 2017, pp. 153-156.

42. T. Kure and S. Kasahara, Lighterbody: RNN based anticipated virtual body makes you feel lighter. 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), IEEE, 2021, pp. 561-562.

43. A. Aalerud and G. Hovland, Evaluation of perception latencies in a human-robot collaborative environment. 2020 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2020, pp. 5018-5023.

44. J. Taufatofua, P. Pounds, and J. Wiles, PiRat: An autonomous framework for studying social behaviour in rats and robots. 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, October 2018, pp. 7601-7608.

45. K. Zhou, RGBD-based Real-time Volumetric Reconstruction System: Architecture Design and Implementation. 2022 IEEE International Conference on Visual Communications and Image Processing (VCIP), IEEE, December 2022, pp. 1-5.

Authors Introduction

Mr. Alexander Alexeev
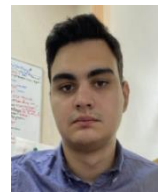
He received his Bachelor's degree in Regional Studies in 2022 from the Intitute of International Relations, Kazan Federal University (KFU), Russia. He is currently a master student in Institute of Information Technologies and Intelligent Systems (ITIS), KFU, Russia

Professor Edgar A. Martínez-García

He is a full Professor at the Universidad Autónoma de Ciudad Juárez, Mexico; founder and Head of the Robotics Laboratory; leader of the Mechatronics academic body at the Institute of Engineering and Technology, since 2007. He obtained his Ph.D. degree in Robotics Engineering from the University of Tsukuba, Japan (2005).

Ms. Tatyana Tsoy

In 2012 she graduated from International Area Studies Master's Degree Program at the University of Tsukuba. In 2023 completed a PhD in Computer science and information processes at IT IS, KFU.

Professor Evgeni Magid

He is a Head of Intelligent Robotics Department, a founder, and a Head of Laboratory of Intelligent Robotic Systems (LIRS) at Kazan Federal University, Russia. Senior member of IEEE. He obtained Ph.D. degree in Robotics Engineering from the University of Tsukuba, Japan.

An investigation on the impact of human-robot interactions during an autonomous obstacle avoidance task

Riham Salman

Institute of Information Technology and Intelligent Systems, Kazan Federal University, Kazan, Russia

Shifa Sulaiman

Laboratory of Intelligent Robotics Systems (LIRS), Institute of Information Technology and Intelligent Systems, Kazan Federal University, Kazan, Russia

Renata Islamova

Laboratory of Intelligent Robotics Systems (LIRS), Institute of Information Technology and Intelligent Systems, Kazan Federal University, Kazan, Russia

Tatyana Tsoy

Laboratory of Intelligent Robotics Systems (LIRS), Institute of Information Technology and Intelligent Systems, Kazan Federal University, Kazan, Russia
Email: RikSalman@stud.kpfu.ru, shifa_1910@it.kfu.ru, rislamova@it.kfu.ru, tt@it.kfu.ru

Abstract

The purpose of this research is to investigate how an interaction between humans and robots influences a safety of an autonomous obstacle avoidance task. The research collected and analyzed data from surveys and interviews using a combination of quantitative and qualitative methodologies. The findings contributed to our understanding of a complex interplay between a human-robot interaction, a perception, and a robot navigation safety. Based on these findings, the study proposed a number of recommendations for improving both physical and psychological safety aspects of an autonomous robot navigation.

Keywords: Human-robot interaction, HRI, autonomous obstacle avoidance, navigation safety

1. Introduction

Autonomous robots are becoming an integral part of diverse human environments due to the advancements of robotic technologies from personal healthcare [1] to automated manufacturing fields [2], [3]. The progression is expected to lead to a future in which both humans and robots coexist and interact on a regular basis [4], [5]. However, the integration of autonomous robots into human environments is not without significant challenges, particularly when it comes to ensure safety during autonomous obstacle avoidance tasks [6], [7]. Autonomous obstacle avoidance [8] and human following [9] is a critical feature for any autonomous robot operating in a human-populated environment [10]. It is a capability that allows a robot to navigate through an environment avoiding static and dynamic obstacles, including humans. A robot which can precisely avoid obstacles is considered safe [11]. Understanding and addressing various aspects of safety require a comprehensive exploration of the relationship between humans and robots [12], [13]. This involves indulging

into the complex dynamics of human-robot interaction (HRI), which encompasses the objective behavior of the robot, subjective experiences and perceptions of the human [14], [15].

The objective of this research is to provide an in-depth examination of the relationship between humans and robots, focusing on its impact on the safety of autonomous obstacle avoidance. Specifically, the research will investigate how humans interact with robots that are capable of autonomous navigation. The investigation will contribute to a deeper understanding of this critical aspect of autonomous robotics, providing insights that could help improve the design and operation of future autonomous robots. It will also address gaps in the existing body of knowledge, particularly in the intersection of autonomous navigation, HRI, and perceived safety [16]. Surveys and interviews were conducted in a group of people, who had experience with robotic environments. The inferences obtained through these surveys and interviews were used to conclude about the current status of HRI in various fields [17].

2. Methodology

The methodology outlines the research approaches that were utilized to investigate the relationship between HRI and the safety of autonomous obstacle avoidance. The research began with a comprehensive review of existing literature from databases, journals, and conference proceedings related to HRI, robotic safety, and autonomous obstacle avoidance. Following the comprehensive literature review, the research focused on the development of an experimental setup, which consisted of a robot equipped with a capacity for autonomous navigation and obstacle avoidance. A restaurant delivering food to customers using a mobile robot (shown in Fig. 1(a) was selected for the survey. The robot can move to assigned tables and deliver food as shown in Fig. 1(b).
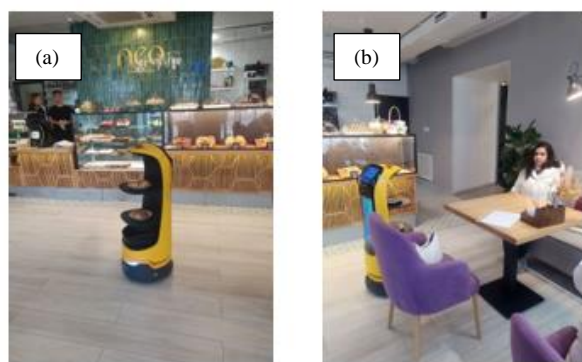
Fig. 1 Robot navigation (a) A food delivery robot (b) The robot delivering food to customers

The robot avoided collision with static and dynamic obstacles using IR sensors fixed on the base of the robot. Fig. 2 depicts the stages of a food delivery using the mobile robot. Fig. 2 represents the robot waiting for the food delivery near the restaurant kitchen. Fig. 2(a) and (b) represent a motion of the robot towards the assigned table and destination table respectively. A customer receiving the food items are shown in Fig. 2(d).

To obtain a human-centered perspective on the safety and comfort of autonomous robots, surveys and semi-structured interviews were conducted with individuals who had interacted with the robot. Questions were designed to explore various phases of human-robot interaction, including social acceptability, perceived safety, and trust in the robot's autonomous decision-making capabilities. The customers and employees working in the restaurant were participated in the survey. The gathered data were subjected to rigorous analyses using suitable statistical techniques. This critical process of data analysis allowed the transformation of raw data into meaningful information, thereby offering both quantitative and qualitative insights into the correlation between HRI and the safety of autonomous obstacle avoidance.

Fig. 2 Different phases of the robot work: (a) Robot waiting for collecting food items (b) Moving to towards a destination table (c) Robot reached the destination table (d) A customer taking the food item from the robot

The final step in the methodology involved the validation of the findings. This was done through peer-reviews and expert consultations, ensuring the research conclusions were robust, reliable, and could stand up to scrutiny. This step served as a quality check and ensured that the research had been conducted in a methodologically sound manner. Further, it also provided an opportunity for receiving recommendations for future studies, thereby adding to the overall value and impact of the research. This methodological approach aimed to provide a comprehensive, multi-dimensional understanding of the research topic, combining the strengths of both empirical and qualitative research methods.

3. Results/findings

In this section, we present and analyze the results gathered from our research and the insights obtained from the HRI studies. we conducted surveys and interviews with participants after they observed or interacted with the robot. A total of 60 participants attended the survey and expressed their opinions and suggestions regarding the coexisting of humans and robots inside a restaurant environment. The qualitative data suggested a generally positive perception of safety, comfort, and trust in the robot's autonomous capabilities. Approximately 90% of the participants reported that they felt safe with the robot's movements and obstacle avoidance capabilities. Interestingly, a correlation was observed between the

robot's behavior and participant responses. Smoother and slower robot movements were associated with increased comfort and trust ratings from the participants. About 80% of the participants conveyed that the robot food delivery was faster compared to human food delivery services. However, some of the participants are concerned about the precision of delivery and quoted about interaction difficulties they faced during food order updating.

The perception of risk and social acceptability associated with the autonomous robot were also included in the survey. We interpreted that in a populated environment, participants perceived the robot as less risky compared to a controlled environment, contrary to our expectations. This might be attributed to the social dynamics of a populated setting, where humans provide an additional layer of unpredictability that makes the robot's autonomous capabilities more evident and appreciated. The social acceptability of the robot was rated highly by 82% of participants, signifying the successful integration of social robotics principles in its design and operation. However, 27% of the customers preferred contactless delivery. 70% of the participants considered robot usage inside a restaurant as an eco-friendly option. 20% of the participants chose human workers in place of autonomous rover delivery due to safety concerns. Most participants concerned about the robot malfunctioning, and navigational errors.

4. Discussion

The study was designed to investigate the relationship between HRI and the safety of autonomous obstacle avoidance, contributing to the growing body of knowledge in this field. The multi-dimensional methodological approach, combining both empirical and qualitative research methods, was integral in producing a comprehensive understanding of the research topic. The survey results regarding autonomous navigation and obstacle avoidance revealed that the robot was successful in navigating the environments and avoiding obstacles with a high success rate.

Results from the HRI revealed a generally positive perception of the robot's autonomous abilities, showing trust in the robot's decision-making and overall satisfaction with their interaction. These findings provided valuable insights into the subjective aspects of robot safety, complementing the quantitative data obtained. Risk perception and social acceptability were also significant factors in our study. Results suggested that participants perceived the autonomous robot as safe and socially acceptable, even in densely populated environments.

When performing HRI activities, robots must be equipped with adaptive abilities to prevent accidents and task incompletions. There are a lot of static and dynamic obstacles present in a restaurant setting. In order to prevent collisions when moving from the food pickup location to the destination table, the robot needs to be employed with an effective collision avoidance approach. By utilizing proper sensors and controllers, the robot can steer clear of obstructions and prevent crashes during the motions. Reserving separate paths for robots can minimize chances of collisions with obstacles. To help users comprehend how the robot is moving, the robot can also show path details on a screen or verbally during its motion.

The study's findings also had several theoretical implications. They reinforced the theories of autonomy, social robotics, proxemics, risk perception, and planned behavior [15]. The robot's autonomous performance aligned well with the theory of autonomy in robotics. The feedback from participants suggested that social robotics theories were applicable in real-world environments, reinforcing the necessity of developing robots that respect human social norms.

5. Conclusion

In conclusion, the research undertaken successfully addressed the primary goals of investigating the relationship between HRI and the safety of autonomous obstacle avoidance. The investigation revealed that the interaction between humans and robots is multi-faceted. It emphasized that the concept of safety in autonomous obstacle avoidance is not limited to the physical ability of the robot to navigate without causing collisions or accidents. It also heavily depends on the subjective feelings of safety, comfort, and trust experienced by the humans who interact with or are in the vicinity of these robots. Therefore, the importance of considering both objective and subjective factors when evaluating the safety of autonomous robots in environments shared with humans was strongly underscored. This research also pointed towards potential areas for future study, demonstrating that the quest for knowledge in this domain is far from exhausted. For instance, it suggested the need for further exploration into how different design elements and behavioral parameters of the robot might influence human perception and interaction. This could include aspects such as the robot's size, color, speed, motion patterns, and communication methods. Research into these directions could yield interesting results that might further refine the approach to designing and deploying autonomous robots in human environments.

Acknowledgements

Riham Salman, Shifa Sulaiman, Renata Islamova, Tatyana Tsoy

References

1. E. Magid, A. Zakiev, T. Tsoy, R. Lavrenov and A. Rizvanov, Automating pandemic mitigation, Advanced Robotics, 35(9), 2021, pp. 572-589.

2. M. Shamout, R. Ben-Abdallah, M. Alshurideh, H. Alzoubi, B. A. Kurdi and S. Hamadneh, A conceptual model for the adoption of autonomous robots in supply chain and logistics industry, Uncertain Supply Chain Management, 10(2), 2022, pp. 577-592.

3. M. Mustafin, E. Chebotareva, E. Magid, H. Li and E. Martinez-Garcia, Features of Interaction Between a Human and a Gestures-Controlled Collaborative Robot in an Assembly Task: Pilot Experiments, International Conference on Artificial Life and Robotics (ICAROB 2023), 2023, pp. 158-162.

4. K. Merckaert, B. Convens, C. J. Wu, A. Roncone, M. M. Nicotra and B. Vanderborght, Real-time motion control of robotic manipulators for safe human–robot coexistence, Robotics and Computer-Integrated Manufacturing, 2022, pp. 73.

5. S. R. Schepp, J. Thumm, S. B. Liu and M. Althoff, Sara: A tool for safe human-robot coexistence and collaboration through reachability analysis, in 2022 International Conference on Robotics and Automation (ICRA), 2022, pp. 4312-4317.

6. A. Dobrokvashina, S. Sulaiman, A. Zagirov, E. Chebotareva, K.-H. Hsia and E. Magid, Human robot interaction in collaborative manufacturing scenarios: Prospective cases, in 2022 International Siberian Conference on Control and Communications (SIBCON), 2022, pp. 1-6.

7. A. C. Simões, A. Pinto, J. Santos, S. Pinheiro and D. Romero, Designing human-robot collaboration (HRC) workspaces in industrial settings: A systematic literature review, Journal of Manufacturing Systems, vol. 62, 2022, pp. 28-43.

8. E. Magid, R. Lavrenov and A. Khasianov, Modified Spline-based Path Planning for Autonomous Ground Vehicle. In *ICINCO (2)*, July 2017, pp. 132-141.

9. E. Chebotareva, E. Magid, A. Carballo and K.-H. Hsia, Basic User Interaction Features for Human-Following Cargo Robot TIAGo Base, Proceedings of 13th International Conference on Developments in eSystems Engineering (DeSE), 2020, pp. 206-211.

10. E. Chebotareva, R. Safin, K.-H. Hsia, A. Carballo and E. Magid, Person-Following Algorithm Based on Laser Range Finder and Monocular Camera Data Fusion for a Wheeled Autonomous Mobile Robot, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 12336, 2020, p. 21-33.

11. R. Lavrenov, F. Matsuno, E. Magid, Modified spline-based navigation: Guaranteed safety for obstacle avoidance, Lecture Notes in Computer Science (inc. subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 10459, 2017, pp. 123-133.

12. A. Zacharaki, I. Kostavelis, A. Gasteratos and I. Dokas, Safety bounds in human robot interaction: A survey, Safety science, 2020, pp. 127.

13. M. Rubagotti, I. Tusseyeva, S. Baltabayeva, D. Summers and A. Sandygulova, Perceived safety in physical human–robot interaction—A survey, Robotics and Autonomous Systems, 2022, pp. 151.

14. K. Dautenhahn, Methodology & themes of human-robot interaction: A growing research field, International Journal of Advanced Robotic Systems, vol. 4(1), 2007, pp. 15.

15. R. Stock-Homburg, Survey of emotions in human–robot interactions: Perspectives from robotic psychology on 20 years of research, International Journal of Social Robotics, vol. 14(2), 2022, pp. 389-411.

16. K.-H. Hsia, J.-H. Cai, S.-L. Pai and E. Magid, Control System for Maintaining Safe Following Distance while Driving, Journal of Robotics, Networking and Artificial Life, vol. 5(4), 2019, pp. 261–264.

17. M. Mustafin, E. Chebotareva, H. Li and E. Magid, Experimental Validation of an Interface for a Human-Robot Interaction Within a Collaborative Task, In International Conference on Interactive Collaborative Robotics, Cham: Springer Nature Switzerland, 2023, pp. 23-35.

Authors Introduction

Ms. Riham Salman

In 2021, she graduated from the University of Tishreen, Syria. Since 2022, she has been a master's student at the Institute of Information Technology and Intelligent Systems (ITIS), Kazan Federal University (KFU), Russia.

Dr. Shifa Sulaiman

In 2013, she received her Master's degree in Machine Design from Mahatma Gandhi University, India. In 2013-2017 worked as an Assistant Professor in various Indian Engineering Institutes. She earned her PhD from National Institute of Technology, Calicut, India, specialized in Humanoid Robotics in 2023. She is currently working as a Research Associate at LIRS, ITIS, KFU, Russia.

Ms. Renata Islamova
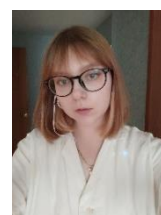
She is currently pursuing bachelor's degree program in linguistics as an English and German interpreter at the Institute of International Relations, KFU, Russia. She is also working as a laboratory assistant at the Laboratory of Intelligent Robotic Systems (LIRS) at KFU

Ms. Tatyana Tsy

In 2012 she graduated from International Area Studies Master's Degree Program at the University of Tsukuba. In 2023 completed a PhD in Computer science and information processes at ITIS, KFU, Russia.

A Comparative Analysis of Object Detection Methods for Robotic Grasping

Nikita Kolin

Intelligent Robotics Department, Kazan Federal University, 35 Kremlevskaya St., Kazan, 420008, Russian Federation

Elvira Chebotareva

Intelligent Robotics Department, Kazan Federal University, 35 Kremlevskaya St., Russian Federation
Email: nikitakolin7@gmail.com, Elvira.Chebotareva@kpfu.ru

Abstract

The objects grasping is one of the fundamental robotic problems. Accurate and efficient real-time object detection is crucial for successful grasping in robots equipped with monocular vision. Deep machine learning has made significant progress in solving problems of object detection and image segmentation. At the same time, classical computer vision methods do not lose their relevance and can also be used for these tasks. In this research, we conduct a comparative analysis of the effectiveness the YOLOv8-seg neural network model versions for solving the image segmentation problem with classical segmentation methods. The obtained results allowed us to formulate some recommendations on the choice of a particular method for object detection depending on the surrounding environment conditions.

Keywords: Robotic grasping, Robotic arm, Object detection, Object segmentation

1. Introduction

Robotic grasping is one of the fundamental tasks of robotics, relevant to all its branches [1], including industrial [2], service [3] and social robotics [4]. Grasping can be particularly important in collaborative robotics, when a robot and a person use common tools or perform joint assembly [5].

Obtaining information about the shape of the captured objects is an important stage in the implementation of grasping [6]. The shape of objects can be determined using various sensors, including cameras and lidars. At the same time, solutions that allow detect grasping objects using monocular cameras are also of practical interest. Deep machine learning has made significant progress in solving problems related to object detection [7] and image segmentation [8]. However, classical computer vision methods do not lose their relevance and can also be used for these tasks. In this paper, we present the results of our experiments aimed at comparing classical object detection methods with object detection and segmentation using the YOLOv8-seg models family. The aim of our research is to evaluate the accuracy and effectiveness of various methods for detecting object contours in robotic grasping.

2. Related Work

In the past few years, there has been a lot of research reported on robotic grasping problems. Some approaches to solving the problem of robotic grasping are presented in the review [9]. The overview of different approaches to computer vision-based robotic grasping presented in [10] and [11]. The effectiveness of deep machine learning in object detection on RGB-D images for robotic grasping is well demonstrated in the work [12]. At the same time, detecting objects in RGB images seems to be a much more difficult task, since such images do not contain depth information. Examples of works exploring the challenges of robotic grasping using monocular vision and RGB images include works [13] and [14].

Often, the effectiveness of grasping area detection methods is evaluated using one of the popular datasets, such as Cornel dataset [15] and Jacquard dataset [16]. The use of publicly available datasets allows comparing the effectiveness of different approaches. According to the data from the "Paper with Code" platform [17] the accuracy of various grasping methods based on machine learning, estimated with the help of Cornel Dataset was made from 88% in 2016 for Multi-Modal Grasp Predictor [18] to 98.2% in 2021 for Ainetter and Fraundorfer CNN-based architecture [19] and 98.9% for Efficient Grasping model by Cao et. al [20]. At the same time, the accuracy estimates on the Jacquard dataset for the Ainetter and Fraundorfer model [19] was 92.95%, and for the Efficient Grasping model by Cao et. al [20] was 95.6%.

As it can be seen, estimates of the accuracy of different methods may vary from the quality of the test dataset. In addition, although public datasets provide an overview of the effectiveness of a given method, these datasets do not always take into account the peculiarities of objects and

Nikita Kolin, Elvira Chebotareva

environments encountered in real cases. Therefore, datasets that are close to the real conditions of a particular task are always of special interest. In this research work we will make a comparative analysis of various methods of objects detection on own dataset, adapted to the real object grasping task. As an example of a model that provides acceptable accuracy along with the high inference speed and usability we use the YOLOv8-seg model for object segmentation. The YOLOv8-seg model is based on variant of the U-Net architecture [21]. The YOLO models family based on hybrid convolutional neural network architecture [22], [23] and often used in robotics.

3. Materials and Methods

3.1. *Object detection problem statement*

We model a situation in which the working cell of a robotic manipulator is equipped with a single monocular camera fixed above the working area. During the experiment, the camera does not change its position. This camera transmits image of the table area, on which three objects can be placed at random: pliers, screwdrivers and PIR sensor. We assume that the objects are homogeneous, so we can assume that the center of mass of each object coincides, or is close, with its geometric center. Therefore, the optimal gripping points must also be near the geometric center of each object.

Thus, in the context of the task of detecting objects for their robotic grasping, we need to determine the object class (pliers, screwdriver, or PIR sensor), the contours of the object, the geometric center of the object (the center of mass of the contour) and the orientation of the object. Note that in order to calculate the center and orientation of an object in this case it is enough to define its contour.

3.2 *Datasets*

For our experiments, we have prepared a set of 100 images with resolution 640×480 pixels containing three different target objects: pliers, screwdriver and PIR sensor. The choice of objects is determined by the context of collaborative tasks in which it is planned to use this robot in the future. In addition, these objects have different shapes and colors, which can help identify potential issues with specific object detection methods. The images were taken under various lighting conditions and against different backgrounds.

However, to test the hypotheses of our study, we also prepared 20 images with backgrounds that differed from those in the training dataset images. A total of 165 annotated images were prepared for model training, with 132 included in the training dataset, 17 images in the validation dataset, and 16 in the test dataset. To improve the quality of training, 236 images derived from the original 132 training dataset images were added to the

training dataset through rotation and perspective distortion.

3.3 *Object detection methods*

For our experiments, we used the YOLOv8n-seg and YOLOv8s-seg models pre-trained on the COCO dataset by Microsoft. YOLOv8n-seg and YOLOv8s-seg models trained on a set of 368 images of size 640×480.

To extract the object's contours we used two classical methods. The first method is based on the conversion to HSV space and channel thresholding. And the second method uses the Canny edge detector, as well as auxiliary erosion and dilation transformations. The classification of the extracted contours is based on a single feature – the aspect ratio of the bounding box.

3.2. *Evaluation*

Each method of object detection and segmentation was evaluated using two commonly used metrics: the Jaccard index (1) – also known as the IoU (Intersection over Union) metric, and the Dice coefficient (2) – also known as the F1-score metric. These metrics are based on the formulas (1) и (2). In formulas (1) and (2) A is the set of pixels belonging to the real object in the image, B is the set of pixels of the result of segmentation.

$$IoU = \frac{A \cap B}{A \cup B} \qquad (1)$$

$$F1 = 2\frac{|A \cap B|}{|A| + |B|} \qquad (2)$$

For each method, we also measured the inference time per frame in the test video and calculated the inference speed as the number of processed frames per second (FPS).

4. Experiment Results and Discussion

At the first stage of the experiment, we evaluated the average values of IoU and F1-score for 100 images from our dataset. Then, at the second stage of our experiment, we conducted the evaluation using 20 images with significantly different backgrounds compared to the backgrounds of the images in the training dataset. The results of these stages are presented in Table 1 and Table 2.

Table 1. The object detection methods accuracy obtained on images with conditions similar to the training dataset.

Method	IoU	F1-score
Thresholding	0.4112	0.4658
Canny edge detector	0.1720	0.2257
YOLOv8n-seg	0.8248	0.8811
YOLOv8s-seg	0.8612	0.9185

Table 2. The object detection methods accuracy obtained on images with conditions differing from the training dataset.

Method	IoU	F1-score
Thresholding	0.8047	0.8771
Canny edge detector	0.0252	0.0469
YOLOv8n-seg	0.3186	0.3477
YOLOv8s-seg	0.2299	0.2535

Thus, in the case where the images background was present in the training dataset, the best result was obtained by YOLOv8s-seg and YOLOv8n-seg models. However, the YOLOv8n-seg model showed an expectedly lower result compared to YOLOv8s-seg model. The thresholding-based method showed significantly lower results than the YOLOv8-seg models, but exceeded the Canny edge detection method. Examples of object detection and segmentation using models YOLOv8n-seg and YOLOv8s-seg are presented in Fig. 1 and Fig. 2.

Fig. 1. The example of successful recognition of all objects using the YOLOv8n-seg model.

However, in cases where the background on which the objects are located differed from the backgrounds in the training images, the method based on threshold processing of channels in the HSV color space showed better results compared to both the YOLOv8s-seg and YOLOv8n-seg models, as well as the method based on Canny edge detector. Fig. 3 presents an example of object detection and segmentation on an image with a background different from the training dataset. In this instance, the YOLOv8s-seg and YOLOv8n-seg models failed to detect object contours on the complex

Fig. 2. The example of successful recognition of all objects using the YOLOv8s-seg model.

background not included in the training dataset. The method based on Canny edge detector also did not give positive results.

Fig. 3. The example of object contour detection using a classic threshold-based method in cases where YOLOv8n-seg and YOLOv8s-seg models failed.

Table 3 presents the results of FPS evaluation on the test video for all the considered methods. The best inference speed result belongs to the method based on the Canny edge detector, followed by the threshold-based method in second place, and the YOLOv8s-seg model demonstrated the worst result.

Table 3. Results of FPS evaluation for the considered methods.

Method	FPS
Thresholding	21.8
Canny edge detector	32.6
YOLOv8n-seg	4.2
YOLOv8s-seg	2.1

5. Conclusion

The obtained results allowed us to formulate some recommendations on the choice of a particular method for object detection depending on the surrounding environment conditions. In cases where the scene is not overloaded with a large number of objects, the environmental conditions remain stable, and the objects significantly differ from the background and have relatively homogeneous texture, classical methods can be quite successfully applied to extract the contours of the target objects. Moreover, in contrast to machine learning methods, these methods do not require the preliminary collection of a large amount of training data and annotation, and they also characterized by higher inference speed. However, if the environmental conditions change dynamically, these methods become practically inapplicable. In this case, the optimal solution would be to train a previously pre-trained neural network model. However, the training dataset must be close to the real conditions in which recognition is planned, in particular, it is necessary that the background on which the target objects are located is the same as in real

Nikita Kolin, Elvira Chebotareva

conditions. Otherwise, the quality of detection and segmentation is significantly reduced.

Acknowledgements

References

1. R. Hodson, How Robots are Grasping the Art of Gripping. Nature, vol 557, 2018, pp. 23–25.
2. M. Mustafin, E. Chebotareva, E. Magid, H. Li and E. Martinez-Garcia, Features of Interaction Between a Human and a Gestures-Controlled Collaborative Robot in an Assembly Task: Pilot Experiments, The 2023 International Conference on Artificial Life and Robotics (ICAROB), Oita, Japan, 2023, pp. 158–162.
3. K. Khusnutdinov, A. Sagitov, A. Yakupov, R. Lavrenov, E. A. Martinez-Garcia, K.-H. Hsia and E. Magid, Household Objects Pick and Place Task for AR-601M Humanoid Robot. Lecture Notes in Computer Science, 11659, 2019, p. 139-149.
4. K. Khusnutdinov, A. Sagitov, A. Yakupov, R. Meshcheryakov, K.-H. Hsia, E. A. Martinez-Garcia and E. Magid, Development and Implementation of Grasp Algorithm for Humanoid Robot AR-601M. Proceedings of the 16th International Conference on Informatics in Control, Automation and Robotics, 2019, pp. 379-386.
5. M. Mustafin, E. Chebotareva, H. Li and E. Magid, Experimental Validation of an Interface for a Human-Robot Interaction Within a Collaborative Task. In International Conference on Interactive Collaborative Robotics (ICR 2023), Baku, Azerbaijan, 2023, pp. 23–25.
6. R. Abdulganeev, R. Lavrenov, R. Safin, Y. Bai and E. Magid, Door handle detection modelling for Servosila Engineer robot in Gazebo simulator. Siberian Conference on Control and Communications, (SIBCON 2022), 2022, p. 1-4.
7. V. Myrzin, T. Tsoy, Y. Bai, M. Svinin and E. Magid, Visual data processing framework for a skin-based human detection. In Interactive Collaborative Robotics: 6th International Conference, ICR 2021, Proceedings 6, September 2021, pp. 138-149.
8. Y. Yu, C. Wang, Q. Fu, R. Kou, F. Huang, B. Yang, T. Yang and M. Gao, Techniques and Challenges of Image Segmentation: A Review. Electronics 12, no. 5, 2023, pp. 1199.
9. L. Wang, Z. Zhang, J. Su and Q. Gu, Robotic Autonomous Grasping Technique: A Survey. 2021 5th Asian Conference on Artificial Intelligence Technology (ACAIT), Haikou, China, 2021, pp. 287–295.
10. K. Kleeberger, R. Bormann, W. Kraus, and M. F. Huber Kleeberger, A Survey on Learning-Based Robotic Grasping. Current Robotics Reports 1, 2020, pp. 239–249.
11. Z. Xie, X. Liang and R. Canale, Learning-Based Robotic Grasping: A Review. Frontiers in Robotics and AI 10, 2023, p. 1038658.
12. Q. Zhang, J. Zhu, X. Sun and M. Liu, HTC-Grasp: A Hybrid Transformer-CNN Architecture for Robotic Grasp Detection. Electronics 12, no. 6, 2023, p. 1505.
13. C. Veiga Almagro, R.A. Muñoz Orrego, Á. García González, E. Matheson, R.M. Prades, M. Di Castro and M.F. Pérez, (MARGOT) Monocular Camera-Based Robot Grasping Strategy for Metallic Objects. Sensors 23, no. 11, 2023, p. 5344.
14. W. Prew, T. Breckon, M. Bordewich and U. Beierholm, Improving Robotic Grasping on Monocular Images Via Multi-Task Learning and Positional Loss, 2020, https://arxiv.org/abs/2011.02888
15. I. Lenz, H. Lee and A. Saxena, Deep learning for detecting robotic grasps. The International Journal of Robotics Research 34, 2013, pp. 705–724.
16. A. Depierre, E. Dellandréa and L. Chen, Jacquard: A Large Scale Dataset for Robotic Grasp Detection. 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Madrid, Spain, 2018, pp. 3511–3516.
17. Paper with Code, https://paperswithcode.com/
18. S. Kumra and C. Kanan, Robotic grasp detection using deep convolutional neural networks. 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vancouver, BC, Canada, 2017, pp. 769–776.
19. S. Ainetter and F. Fraundorfer, End-to-end Trainable Deep Neural Network for Robotic Grasp Detection and Semantic Segmentation from RGB, 2022, https://arxiv.org/abs/2107.05287
20. H. Cao, G. Chen, Z. Li, J. Lin and A. Knoll, Lightweight Convolutional Neural Network with Gaussian-based Grasping Representation for Robotic Grasping Detection, 2021, https://arxiv.org/abs/2101.10226
21. Ultralytics YOLOv8 Tasks, Ultralytics Inc, https://docs.ultralytics.com/tasks/
22. J. Redmon, S. Divvala, R. Girshick and A. Farhadi, You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 779–788.
23. J. Terven and D. Cordova-Esparza, A Comprehensive Review of YOLO: From YOLOv1 and Beyond, 2023, https://arxiv.org/abs/2304.00501

Authors Introduction

Mr. Nikita Kolin

He received his Bachelor's degree in Kazan Federal University, Russia, in 2022. Currently he is a Master's student at the Institute of Information Technologies and Intelligent Systems, Kazan Federal University, Russia.

Assistant Professor Elvira Chebotareva
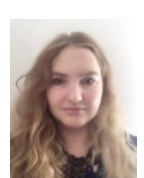
She received her PhD in physics and mathematics from Kazan Federal University. She is currently an assistant professor in Laboratory of Intelligent Robotic Systems (LIRS) at Kazan Federal University, Russia.

Vision-based autonomous navigation for medical examination using a UR3e manipulator

Bulat Abbyasov
Intelligent Robotics Department, Kazan Federal University, 420111 Kazan, Russian Federation

Aidar Zagirov
Intelligent Robotics Department, Kazan Federal University, 420111 Kazan, Russian Federation

Timur Gamberov
Intelligent Robotics Department, Kazan Federal University, 420111 Kazan, Russian Federation

Hongbing Li
Department of Instrument Science and Engineering, Shanghai Jiao Tong University, Shanghai, Minhang, 200240, China

Evgeni Magid
Intelligent Robotics Department, Kazan Federal University, 420111 Kazan, Russian Federation
E-mail: abbyasov@it.kfu.ru, ai.zag@it.kfu.ru, TRGamberov@stud.kpfu.ru, lihongbing@sjtu.edu.cn, magid@it.kfu.ru

Abstract

Medical robotics is an emerging field of robotics within a healthcare sector. This interdisciplinary field focuses on prototyping, building and developing advanced robots for numerous clinical application. Modern robotic technologies have a tremendous potential for performing medical examination procedures such as ultrasonography using robotic manipulators that act in an autonomous manner. A manipulator navigation plays a key role in a safe and efficient exploration of a human body. This paper presents a development of a vision-based autonomous navigation system for a 6-DOF UR3e robotic arm. The developed system is based on a 3D point cloud and uses MoveIt for path planning. Gazebo was used as a simulation framework to validate the navigation system.

Keywords: Medical robotics, Manipulator, Autonomous navigation, ROS, MoveIt, Gazebo

1. Introduction

The healthcare industry has been making a notable progress in recent years with an advancement of robotics technologies [1], [2]. Due to COVID-19 pandemic an automation of medical procedures employing robotics technologies became critical [3], [4]. Robots are progressively being employed for a broad spectrum of tasks that include surgery [5], rehabilitation [6], diagnostics [7], and drug delivery [8], offering precision, efficiency and remote operational capabilities [9]. Modern robotic technologies have a tremendous potential for performing medical examination procedures using robotic manipulators [10] that employ various end-effectors to provide additional functionalities within specific tasks. Nowadays, robotic arms became a major focus for medical research groups [11]. Robotic manipulators feature a lightweight reconfigurable arm design, high precision actuators and motion control systems. These robots are used in medical scenarios that require a high degree of precision and quality of a task performance, e.g., tissue suturing [12], minimally invasive surgery [13], [14], and ultrasonography [15].

An autonomy enables an unskilled human worker to easily configure a robotic system in a setup phase before performing a medical procedure. Autonomous manipulators execute tasks consistently without a variability that can be introduced by human operators [16]. This consistency is crucial in medical examination scenarios where uniformity is essential. Manipulator navigation plays a key role in a safe and efficient autonomous exploration of a human body [17]. The exploration refers to autonomously move, position, and control a trajectory of a manipulator within a designated space. Vision-based manipulator navigation is an approach of guiding and controlling a robotic manipulator's movement using visual information. A robotic system relies on cameras or other vision sensors to perceive and interpret an environment [18].

This paper presents a development of a vision-based autonomous navigation system for medical examination using a 6 degrees of freedom (DoF) UR3e robotic arm. The developed system is based on 3D point cloud data. 3D point clouds are produced by a depth camera, which measure physical dimensions of a human body. MoveIt package [19] was used for motion planning, manipulation,

and control of the robotic arm. Gazebo [20] was used as a simulation framework to validate the navigation system.

2. System setup

A UR3e manipulator (Fig. 1, left), manufactured by Universal Robots, stands as a suitable robotic system for automating low-weight processing tasks. The capability to grab and move a medical tool is facilitated by its payload capacity of 3 kg, providing a sufficient control to navigate an end-effector. Flexible motion planning is achieved using the UR3e arm's capability for 360 degrees of rotation in all its joints. For a virtual robot setup a ROS-based UR3/UR3e controller with simulation in Gazebo was used (Fig. 1, right).
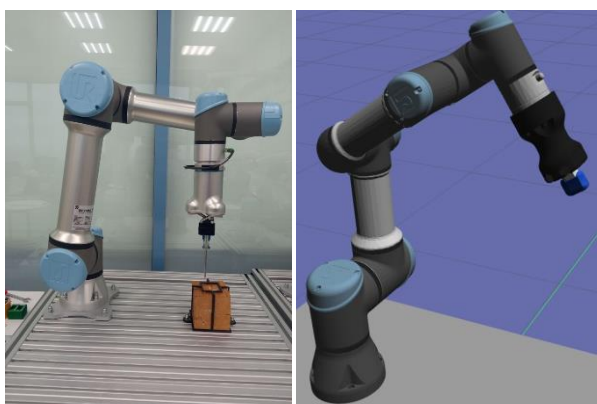
Fig. 1. UR3e robotic arm: a real robot (left) and a virtual manipulator model (right)
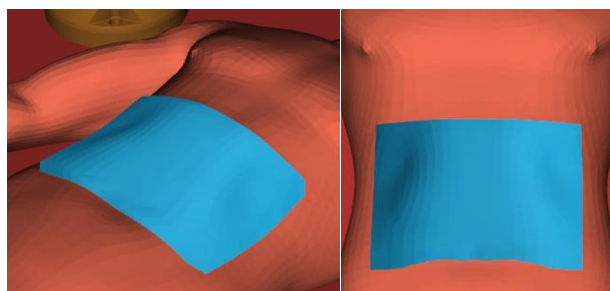
Fig. 2. Depth data: a 3D point cloud (blue) of a region of interest

A depth sensor is a type of imaging device that captures depth data along with visual data. Unlike traditional cameras that record color and intensity information, depth cameras provide additional details about a distance from a camera to an object. In a Gazebo simulation environment, integrating a Kinect camera model involves configuring a depth camera plugin. ROS/Gazebo provides a package *gazebo_ros_pkgs* that contains wrappers and tools for using ROS with Gazebo. The package includes *gazebo_ros_openni_kinect* plugin that serves as a driver for interfacing a Kinect camera with Gazebo simulation framework. The Gazebo camera plugin publishes depth data on a ROS topic of type

sensor_msgs/Image. Depth data captured by the virtual Kinect camera is used by a 3D point cloud processing module. Fig. 2 shows a 3D point cloud of an abdomen.

A simulation environment in Gazebo contains a UR3e manipulator model. UR3e is mounted on a supporting base that allows to effectively cover a table used for examinations. The Kinect camera model with a depth camera plugin captures visual and depth data of a human body. A human body model is placed on the examination table within the manipulator workspace and the Kinect camera range. UR3e is equipped with a transducer device used as an end-effector. The virtual environment is depicted in Fig. 3.
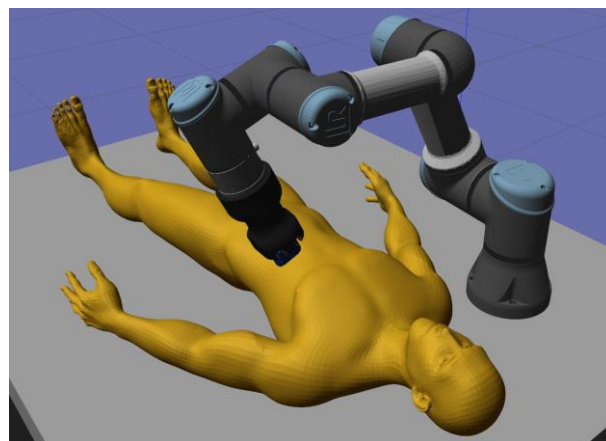
Fig. 3. Gazebo environment: a human, an examination table and UR3e manipulator, that explores a chest.

3. Navigation

A proposed navigation system consists of three modules: a 3D point cloud processing, a path planning and a motion planning. The first module serves as a foundation for a spatial perception. The second module, which is responsible for waypoint generation, takes the processed point cloud data and calculates series of feasible waypoints for navigation. The third module focuses on translating the planned paths into executable motion commands for the manipulator.

A point cloud is a collection of data points defined by their coordinates in 3D space. Each point represents a specific position in an environment and may include additional information, such as color or intensity. The module processes raw 3D point cloud data obtained from the depth camera. This step involves filtering input depth data and removing invalid points. Point Cloud Library (PCL) library [21] was utilized to process raw 3D point cloud data and create a digital representation of a human body's specific region of interest.

Waypoint generation determines a set of points that represents a desired path for the manipulator. The module is responsible for generating waypoints and utilizes the processed 3D point cloud model as an input. Normal vectors of points are employed to calculate a rotation matrix for the end-effector. The output is a sequence of feasible waypoints for motion planning (Fig. 4, bottom).

A motion planning process involves converting a trajectory into a series of motion commands. The motion planning module uses MoveIt framework to plan a motion for the robotic system, validate it and then execute the motion plan (Fig. 4, top). MoveIt uses inverse kinematics solvers for determining joint configurations and collision checking to ensure safe motion of UR3e.
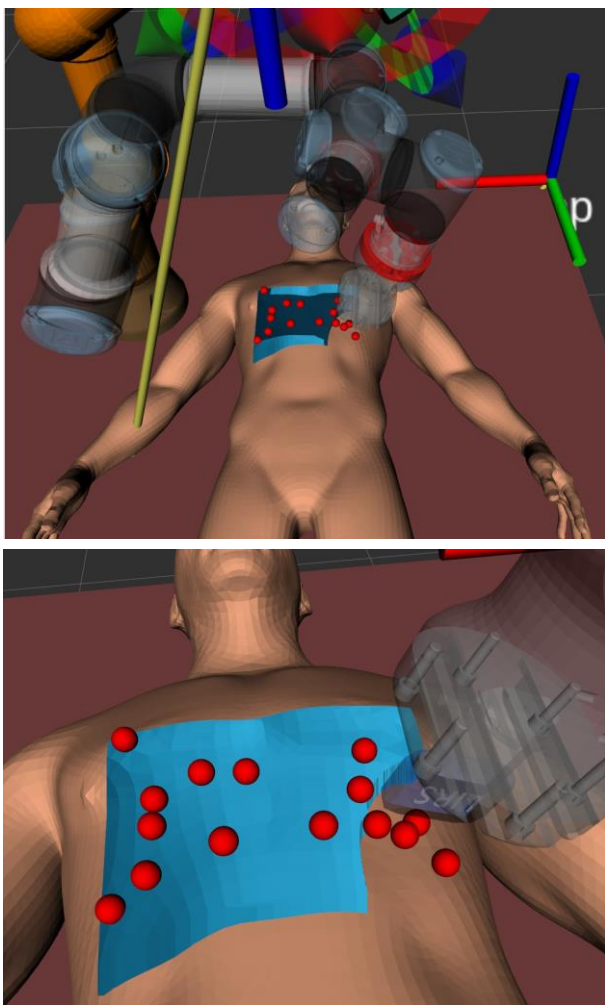
Fig. 4. Navigation in RViz: red markers denote waypoints that MoveIt uses for motion planning

RViz software [22] is commonly used together with MoveIt for visualization of motion planning tasks. A control ROS node is written to bring up the autonomous navigation system and start path planning.

4. Conclusions

This paper presented a vision-based autonomous navigation system for medical examination using a 6-DOF UR3e robotic arm. The developed system relies on depth-sensing cameras and utilizes a 3D point cloud. MoveIt motion planning framework was used for path planning and control of the robotic arm. The validation of the developed navigation system was conducted in Gazebo simulation framework. As a part of our future work, we plan to validate the proposed navigation system on a real UR3e robot. To achieve this, we will use a real Kinect depth camera and a human mannequin.

Acknowledgements

References

1. J. Holland, L. Kingston, C. McCarthy, E. Armstrong, P. O'Dwyer, F. Merz and M. McConnell, Service robots in the healthcare sector. Robotics, 10(1), 2021, p. 47.
2. M. Kyrarini, F. Lygerakis, A. Rajavenkatanarayanan, C. Sevastopoulos, H. R. Nambiappan, K. K. Chaitanya and F. Makedon, A survey of robots in healthcare. Technologies, 9(1), 2021, p. 8.
3. E. Magid, A. Zakiev, T. Tsoy, R. Lavrenov and A. Rizvanov, Automating pandemic mitigation. Advanced Robotics, vol. 35(9), 2021, p. 572-589.
4. A. Khamis, J. Meng, J. Wang, A. T. Azar, E. Prestes, A. Takacs, I. J. Rudas and T. Haidegger, Robotics and intelligent systems against a pandemic. Acta Polytechnica Hungarica, 18(5), 2021, pp.13-35.
5. H. Li, X. Nie, D. Duan, Y. Li, J. Zhang, M. Zhou and E. Magid, An admittance-controlled amplified force tracking scheme for collaborative lumbar puncture surgical robot system. The International Journal of Medical Robotics and Computer Assisted Surgery, 18(5), 2022, e2428.
6. Q. Liu, J. Zuo, C. Zhu and S. Q. Xie, Design and control of soft rehabilitation robots actuated by pneumatic muscles: State of the art. Future Generation Computer Systems, vol. 113, 2020, pp. 620-634.
7. N. Simaan, R. M. Yasin and L. Wang, Medical technologies and challenges of robot-assisted minimally invasive intervention and diagnostics. Annual Review of Control, Robotics, and Autonomous Systems, vol. 1, 2018, pp. 465-490.
8. E. B. Joyee and Y. Pan, Additive manufacturing of multi-material soft robot for on-demand drug delivery applications. Journal of Manufacturing Processes, vol. 56, 2020, pp. 1178-1184.
9. X. Kong, P. Gao, J. Wang, Y. Fang and K. C. Hwang, Advances of medical nanorobots for future cancer treatments. Journal of Hematology & Oncology, 16(1), 2023, p. 74.
10. H. Afrisal, A. D. Setiyadi, M. A. Riyadi, R. Ismail, O. Toirov and I. Setiawan, Performance Analysis of 4-DOF RPRR Robot Manipulator Actuation Strategy for Pick and Place Application in Healthcare Environment. International Journal on Advanced Science, Engineering and Information Technology, 12(6), 2022, pp. 2258-2265.

11. L. Zongxing, L. Wanxin and Z. Liping, Research development of soft manipulator: A review. Advances in Mechanical Engineering, 12(8), 2020, p. 1687814020950094.
12. A. Sagitov, T. Tsoy, H. Li and E. Magid, Automated open wound suturing: detection and planning algorithm. J. Robotics Netw. Artif. Life, 5(2), 2018, pp. 144-148.
13. J. Zhu, L. Lyu, Y. Xu, H. Liang, X. Zhang, H. Ding and Z. Wu, Intelligent soft surgical robots for next-generation minimally invasive surgery. Advanced Intelligent Systems, 3(5), 2021, p. 2100011.
14. S. Nisar, A. Hameed, N. Kamal, O. Hasan and F. Matsuno, Design and realization of a robotic manipulator for minimally invasive surgery with replaceable surgical tools. IEEE/ASME Transactions on Mechatronics, 25(6), 2020, pp. 2754-2764.
15. J. Cornejo, J. P. Perales-Villarroel, R. Sebastian and J. A. Cornejo-Aguilar, Conceptual design of space biosurgeon for robotic surgery and aerospace medicine. In 2020 IEEE ANDESCON, October 2020, pp. 1-6.
16. Y. Liu, Z. Li, H. Su, L. Jiang and C. Y. Su, Whole body control of an autonomous mobile manipulator using series elastic actuators. IEEE/ASME Transactions on Mechatronics, 2021.
17. A. Kumar, M. Vohra, R. Prakash and L. Behera, Towards deep learning assisted autonomous uavs for manipulation tasks in gps-denied environments. In 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), October 2020, pp. 1613-1620.
18. K. Asadi, V. R. Haritsa, K. Han and J. P. Ore, Automated object manipulation using vision-based mobile robotic system for construction applications. Journal of Computing in Civil Engineering, 35(1), 2021, p. 04020058.
19. M. Görner, R. Haschke, H. Ritter and J. Zhang, Moveit! task constructor for task-level motion planning. In 2019 International Conference on Robotics and Automation (ICRA), May 2019, pp. 190-196.
20. N. Koenig and A. Howard, Design and use paradigms for gazebo, an open-source multi-robot simulator. In 2004 IEEE/RSJ international conference on intelligent robots and systems (IROS) (IEEE Cat. No. 04CH37566), vol. 3, September 2004, pp. 2149-2154.
21. C. C. Wong, H. M. Feng, Y. C. Lai and H. Y. Chen, Manipulator system designs for drawing random objects through point cloud posture estimation. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 236(14), 2022, pp. 1867-1884.
22. A. Dobrokvashina, R. Lavrenov, E. Magid, Y. Bai and M. Svinin, How to Create a New Model of a Mobile Robot in ROS/Gazebo Environment: An Extended Tutorial. International Journal of Mechanical Engineering and Robotics Research, 12(4), 2023.

Authors Introduction

Mr. Bulat Abbyasov

Bulat Abbyasov received his master degree at Institute of Information Technology and Intelligent Systems (ITIS), Kazan Federal University in 2022. Currently, he is a second year PhD student in Robotics at ITIS, Kazan Federal University (KFU).

Mr. Aidar Zagirov

He obtained his B.Sc. degree from KFU in 2023. Currently, he is a first-year student in the Master's program for Intelligent Robotics at ITIS, KFU. Aidar works as a research assistant at the Laboratory of Intelligent Robotic Systems (LIRS) at KFU, Russia.

Mr. Timur Gamberov

He is a 4rd year student at the ITIS, KFU, Russia. He works as a research assistant at LIRS, KFU.

Associate Professor Hongbing Li

He received his Ph.D. degree in Mechano-Micro Engineering at the Tokyo Institute of Technology, Japan. Currently he works at the Shanghai Jiaotong University as an Associate Professor. His research interests include surgical robots, surgical instrument design, robot force control and haptic perception for minimally invasive surgery.

Professor Evgeni Magid

He is a Head of Intelligent Robotics Department, a founder, and a Head of Laboratory of Intelligent Robotic Systems (LIRS) at Kazan Federal University, Russia. Senior member of IEEE. He obtained Ph.D. degree in Robotics Engineering from the University of Tsukuba, Japan.

Robot-Assisted Language Learning: Scientific Data Analysis

Karina Sadyikova
Institute of Information Technology and Intelligent Systems, Kazan Federal University, Kazan, Russia

Valeriya Zhukova
Laboratory of Intelligent Robotics Systems (LIRS), Institute of Information Technology and Intelligent Systems, Kazan Federal University, Kazan, Russia

Roman Lavrenov
Laboratory of Intelligent Robotics Systems (LIRS), Institute of Information Technology and Intelligent Systems, Kazan Federal University, Kazan, Russia
Email: sadyikovakarina@mail.ru, v.zhukova@it.kfu.ru, lavrenov@it.kfu.ru

Abstract

Robot-assisted language learning (RALL) is a direction in education that uses robots in the process of studying foreign languages. In this paper, we present the results of our research work devoted to studying trends in the quantity of publications in the field of using social robots for learning foreign languages. Additionally, we have analyzed some published paper in the field of RALL, which has allowed us to identify several popular robot models used in practical research in recent years. Among these models were the NAO, Pepper and DARWIN OP-2 robots. We have found that these robots are actively used in the educational process for learning foreign languages. In particular, these models are used to improve language and communication skills during implicit learning and interactive games.

Keywords: Robot-assisted language learning, Foreign language learning, Robot teaching assistant, Social robotics

1. Introduction

The last decades have seen significant progress in the field of Artificial Intelligence (AI), which has become an integral part of human life. This development has resulted in the emergence of many systems with the ability to replace certain aspects of human activity. Modern AI systems are innovative technological solutions capable of functioning effectively in various fields of activity, including manufacturing, medicine, finance, transportation, and others. Some studies speculate about the prospects of symbiosis of human and cybernetic systems [1]. Robotics is closely related to artificial intelligence and also applies the results of progress in this area [2]. Thanks to artificial intelligence, it is possible to generate a three-dimensional space for robot testing in simulation [3], generate a maze [4] and create obstacles to further calibrate algorithms for complex tasks [5]. Robots, thanks to machine learning algorithms, are able to navigate in the Gazebo simulation environment [6], [7], in real-world environments [8], and can also follow humans [9] and independently plan a route to move [10].

The significant achievement of robotics is the direct assistance to humans, which makes this field in demand nowadays. Modern robots perform a huge range of tasks: they can provide the necessary assistance to the elderly [11], they can be used for medical procedures and examinations [12], [13] and find application as personnel in hospitals [14], and they are also used in areas of high risk for humans, in rescue operations [15].

Currently, robotics is widely used in education [16]. There it finds application in the creation of programs and exercises for students [17], as well as can be an assistant in learning foreign languages [18]. The use of social robotics in education has great potential and is of significant interest to researchers in both the humanities and technical fields of science. Robot assisted language learning (RALL) is a direction in education that uses robots in the process of studying foreign languages. Education using robots is aimed primarily at increasing student motivation and the degree of interactivity of the learning process itself, as well as adapting the educational process to the individual needs of students.

The aim of this paper is to study the dynamics of the number of publications devoted to the use of social robots for learning foreign languages. Additionally, we will examine the robot models used by researchers in the field of RALL.

Karina Sadyikova, Valeriya Zhukova, Roman Lavrenov

2. Analysis of Publication Activity in the Field of RALL

In order to assess the interest of researchers in the field of RALL, we counted the number of published papers for the query "robot-assisted language learning" in the Google Scholar search engine. In total, based on the results of this search query, we received more than 900 scientific papers. Among the collected papers, we have selected those published no earlier than 2010 (totaling 889 papers) and counted the number of publications in each year from 2010 to 2023. The resulting diagram of the number of published papers in each year is presented in Fig. 1. Based on the diagram presented in Fig. 1 we can conclude that interest in social robotics and, in particular, in robot-assisted language learning has increased over the last decade.
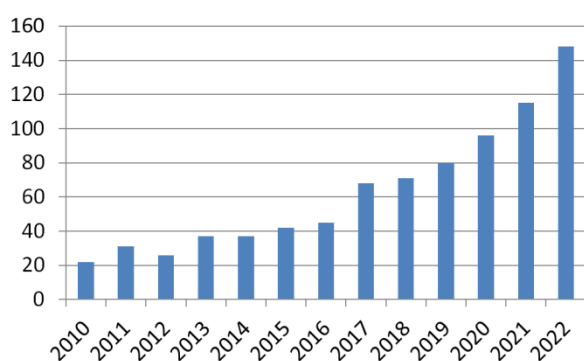
Fig.1 A number of published papers (X-axis) on robot-assisted language learning in the Google Scholar search engine per year (Y-axis)

According to Fig. 1, starting in 2019, interest in robots assisting with foreign language learning began to grow noticeably. Thus, among the papers presented in the search results there were 80 papers published in 2019, and 148 papers published in 2022. This increase in interest may be attributed, among other factors, to the impact that the COVID-19 pandemic has had on the methods of communication in society in general, and within educational systems in particular. Due to restrictions on traveling and face-to-face meetings, many people have turned to online education and digital solutions to learn foreign languages. Robots capable of teaching languages provide a convenient and affordable way to learn virtually. This has further influenced the growth of technology in this field. These results confirm that social robotics and robot-assisted language learning have great potential and will continue to attract the attention of researchers and society at large.

3. Robot Models Used in RALL

To identify the robots currently utilized in robot-assisted language learning research, we conducted an analysis of around 100 papers derived from the search query "robot-assisted language learning" in the Google

Scholar search engine. Among the most popular models of robots used in practical research were the NAO robot and the Pepper robot from SoftBank Robotics, as well as the DARWIN-OP2 robot from ROBOTIS.

The NAO robot model was referenced in 3 research papers [19], [20], [21]. The study [19] examined the impact of the NAO robot functioning as a teacher's assistant on university students' English vocabulary improvement. The research [20] implemented a system of implicit learning utilizing two NAO robot models. In study [21], the NAO robot was utilized in teaching a second language to 5-year-old children. Two research papers [22], [23] referenced the Pepper robot model. In [22], the Pepper robot was employed for interactive game-based learning in single-user and team versions of a quiz game. In study [23], a project was presented which involved the Pepper robot being used as the basis for developing for robotic teaching of the endangered Ainu language. The DARWIN-OP2 robot was also mentioned in two research papers [18], [24]. The study [18] presented the experimental verification of English language lesson scenarios with the DARWIN-OP2 robot in the Gazebo simulator. The study [24] introduced a pilot experiment involving the interaction of preschool children aged 5-6 years with the DARWIN-OP2 robot, serving as a teacher's assistant.

4. Discussion

In this section, we provide the general characteristics of the NAO, Pepper, and DARWIN-OP2 robots, and also review key characteristics of the studies presented in the previous section.

The robots NAO, Pepper, and DARWIN-OP2 share characteristics that make them attractive for research in the field of RALL. The robot NAO has small dimensions - about 58 cm in height, and a humanoid form with a head, arms, and legs. The DARWIN-OP2 robot has more compact dimensions - about 45 cm in height, and a simpler design. Pepper is a taller robot - about 120 cm in height. Like NAO and DARWIN-OP2, Pepper has a head and two arms. However, unlike NAO and DARWIN-OP2, Pepper does not have legs, and overall, Pepper has a more monumental appearance. Overall, the visual design of all three robots is conducive to social interaction.

From the point of view of using these robot models in an educational environment when learning foreign languages, all three robot models also have a number of common attractive characteristics. First of all, these models are programmable models and can be adapted to various educational contexts. From a technical point of view, robots NAO, Pepper and DARWIN-OP2 provide ample opportunities for speech synthesis and recognition, integration of computer vision and machine learning. All this allows researchers to achieve a high

level of human-robot interaction, including the interaction with children.

In Table 1 we provide a comparative analysis of some of the characteristic features of the studies we indicated in previous section, in which the NAO, Pepper and DARWIN-OP2 robot models were used.

Table 1. Key characteristics of studies using the NAO, Pepper and DARWIN-OP2 robots.

The application area	
NAO	Assistant English teacher for university students improving their vocabulary. Implicit learning through conversation.
Pepper	Studying English grammar in universities. Learning an endangered language.
DARWIN-OP2	English lessons with preschool children.
Objectives of the experiments	
NAO	Studying the effect of a robot on improving students' vocabulary.
Pepper	The effectiveness of game-based learning.
DARWIN-OP2	Interaction with preschool children in English lessons.
Distinctive features of the studies	
NAO	The use of implicit learning as a primary learning style.
Pepper	Using a robot as a play partner with an educational interactive quiz.
DARWIN-OP2	The optimization of the English language learning process with the help of the Gazebo simulator and global human-robot interaction scenarios

From the presented in Table 1 training scenarios, it can be concluded that the NAO, Pepper and DARWIN-OP2 robots demonstrate a wide range of capabilities in the context of educational environments. They have been successfully used in university classrooms, ranging from helping students learn foreign language and improving students' vocabulary to using them in game-based learning. In addition, these robots successfully interact with preschool children, optimizing the language learning process using advanced simulators and global human-robot interaction scenarios. Thus, learning cases using these robots reflect their applicability in various educational contexts and their ability to effectively support and improve students' language skills.

5. Conclusions

Our analysis of the number of published papers related to "robot-assisted language learning" demonstrates a steadily growing interest among researchers in the RALL field. Data obtained from a Google Scholar search show a noticeable increase in the number of

publications since 2019, indicating a growing interest in robots helping with foreign language learning. This increase in interest can be partially attributed to the impact of the COVID-19 pandemic, which has prompted a shift to online education and digital language learning solutions due to travel restrictions and social distancing measures. Among the most frequently used robot models in RALL research were the NAO, Pepper, and DARWIN-OP2 robots. These robots are used to improve language skills and communication in a variety of educational scenarios, including implicit learning and interactive games.

Acknowledgements

References

1. M. Talanov, J. Vallverdú, A. Adamatzky et al., Neuropunk Revolution. Hacking Cognitive Systems towards Cyborgs 3.0, International Journal of Unconventional Computing, Vol. 18 (2-3), 2023, pp. 145−201.
2. E. Magid, A. Pashkin, N. Simakov et al., Artificial Intelligence Based Framework for Robotic Search and Rescue Operations Conducted Jointly by International Teams, Proceedings of 14th International Conference on Electromechanics and Robotics "Zavalishin's Readings" ER (ZR), Kursk, Russia, 2019. pp. 15−26.
3. E. Denisov, A. Sagitov, R. Lavrenov et al., DCEGen: Dense Clutter Environment Generation Tool for Autonomous 3D Exploration and Coverage Algorithms Testing, in Interactive Collaborative Robotics. ICR 2019. Lecture Notes in Computer Science, eds. A. Ronzhin, G. Rigoll, R. Meshcheryakov, Springer, Cham, vol 11659, 2019, pp. 215−225.
4. A. Iskhakova, B. Abbyasov, T. Mironchuk et al., LIRS-MazeGen: An Easy-to-Use Blender Extension for Modeling Maze-Like Environments for Gazebo Simulator, in Frontiers in Robotics and Electromechanics. Smart Innovation, Systems and Technologies, eds. A. Ronzhin, V. Pshikhopov, Springer, Singapore, vol 329, 2023, pp. 147−161.
5. R. Gabdrahmanov, T. Tsoy, Y. Bai, M. Svinin and E. Magid, Automatic Generation of Random Step Environment Models for Gazebo Simulator, in Robotics for Sustainable Future. CLAWAR 2021. Lecture Notes in Networks and Systems, eds. D. Chugo, M.O. Tokhi, M.F. Silva, T. Nakamura, K. Goher, Springer, Cham, vol 324, 2022, pp. 408–420.
6. R. Safin, R. Lavrenov and E.A. Martínez-García, Evaluation of Visual SLAM Methods in USAR Applications Using ROS/Gazebo Simulation, Proceedings of 15th International Conference on Electromechanics and Robotics "Zavalishin's Readings", Ufa, Russia, 2021, pp. 371−382.
7. E. Mingachev, R. Lavrenov, T. Tsoy, F. Matsuno, M. Svinin, J. Suthakorn and E. Magid, Comparison of ROS-based monocular visual SLAM methods: DSO, LDSO,

ORB-SLAM2 & DynaSLAM, in 5th International Conference on Interactive Collaborative Robotics (ICR 2020), St. Petersburg, Russia, 2020, pp. 222–233.

8. E. Chebotareva, T. Tsoy, B. Abbyasov, J. Mustafina, E.A. Martinez-Garcia, Y. Bai and M. Svinin, On the Problems of SLAM Simulation for Mobile Robots in the Arctic Conditions, in 5th International Conference on Interactive Collaborative Robotics (ICR 2020), St. Petersburg, Russia, 2020, pp. 33–44.

9. E. Chebotareva, R. Safin, K.-H. Hsia, A. Carballo and E. Magid, Person-Following Algorithm Based on Laser Range Finder and Monocular Camera Data Fusion for a Wheeled Autonomous Mobile Robot, in 5th International Conference on Interactive Collaborative Robotics (ICR 2020), St. Petersburg, Russia, 2020, pp. 21–33.

10. Y. Bai, M. Svinin, E. Magid and Y. Wang, On Motion Planning and Control for Partially Differentially Flat Systems, Robotica, Vol. 39, no. 4, 2021, pp. 718−734.

11. K. Khusnutdinov, A. Sagitov, A. Yakupov, R. Lavrenov, E. A. Martinez-Garcia, K.-H. Hsia, and E. Magid, Household Objects Pick and Place Task for AR-601M Humanoid Robot, in 4th International Conference on Interactive Collaborative Robotics (ICR 2019), Istanbul, Turkey, 2019. pp. 139−149.

12. A. Sagitov, T. Tsoy, H. Li and E. Magid, Automated open wound suturing: detection and planning algorithm. Journal of Robotics, Networking and Artificial Life, Vol. 5, no. 2, 2018, pp. 144−148.

13. A. Shafikov, T. Tsoy, R. Lavrenov et al, Medical palpation autonomous robotic system modeling and simulation in ROS/Gazebo, in 13th International Conference on Developments in eSystems Engineering (DeSE), Liverpool, United Kingdom, 2020, pp. 200−205.

14. R. Safin, R. Lavrenov, T. Tsoy, E. Magid, M. Svinin, S. D. Roy and S. K. Saha, Prioritizing tasks within a robotic transportation system for a smart hospital environment, in 6th International Conference (ICR 2021), St. Petersburg, Russia, 2021, pp. 182−193.

15. V. Myrzin, T. Tsoy, Y. Bai, M. Svinin and E. Magid, Visual data processing framework for a skin-based human detection, 6th International Conference (ICR 2021), St. Petersburg, Russia, 2021, pp. 138−149.

16. E. Chebotareva and M. Mustafin, Android Based Educational Mobile Robot Design and Pilot Evaluation, in 2023 International Conference on Artificial Life and Robotics (ICAROB2023), Oita, Japan, 2023, pp. 145−148.

17. S. Bulatov, E. Kharisova, R. Lavrenov, B. Abbyasov and E. Magid, Structural Model of a Training Computer Program for Improving Professional Skills of a Student in a Role of a District Polyclinic Physician, Journal of Robotics, Networking and Artificial Life, Vol. 8, no. 2, 2021, pp. 122−126.

18. L. Gavrilova, A. Kotik, T. Tsoy et al., Facilitating a preparatory stage of real-world experiments in a humanoid robot assisted English language teaching using Gazebo simulator, in 2020 13th International Conference on Developments in eSystems Engineering, Liverpool, United Kingdom, 2020, pp. 222−227.

19. H. Banaeian and I. Gilanlioglu, Influence of the NAO robot as a teaching assistant on university students' vocabulary learning and attitudes, Australasian Journal of Educational Technology, 37(3), 2021, pp. 71−87.

20. A. Khalifa, T. Kato and S. Yamamoto, Effect of Implicit Learning in Joining-in-type Robot-assisted Language Learning System, International Journal of Emerging Technologies in Learning (iJET), 14(02), 2019, pp. 105–123.

21. M. de Haas, P. Vogt and E. Krahmer, The Effects of Feedback on Children's Engagement and Learning Outcomes in Robot-Assisted Second Language Learning, Front. Robot. AI, 7:101, August (2020).

22. I. Buchem, A. Mc Elroy and R. Tutul, Designing and Programming Game-Based Learning with Humanoid Robots: a Case Study of the Multimodal "Make or Do" English Grammar Game with the Pepper Robot, ICERI2022 Proceedings, Seville, Spain, 2022, pp. 1537-1545.

23. K. Nowakowski, M. Ptaszynski and F. Masui, Spicing up the Game for Underresourced Language Learning: Preliminary Experiments with Ainu Language-speaking Pepper Robot, in 6st Workshop on Linguistic and Cognitive Approaches to Dialog Agents (LaCATODA 2020), 2020.

24. L. Gavrilova, V. Petrov, A. Kotik et al., Pilot Study of Teaching English Language for Preschool Children with a Small-Size Humanoid Robot Assistant, in 2019 12th International Conference on Developments in eSystems Engineering (DeSE), Kazan, Russia, 2019, pp. 253−260.

Authors Introduction

Ms. Karina Sadyikova

She received her Bachelor's degree in Accounting Analysis and Audit from the Faculty of Economics, Kazan Federal University, Russia, in 2022. She is currently a Master's student at the Institute of Information Technologies and Intelligent Systems, Kazan Federal University, Russia.

Ms. Valeriya Zhukova

In 2022 she graduated from Kazan Federal University with a degree in linguistics as an English and Korean interpreter at the Institute of International Relations. Currently is a Master's student in the foreign language teaching program and works as a Research Assistant at the Institute of Information Technologies and Intelligent Systems, Kazan Federal University, Russia.

Assistant Professor Roman Lavrenov

Currently he works as an Assistant Professor at Kazan Federal University. He obtained his Ph.D. at Kazan Federal University in 2020. His main interests are path planning, robot operating system, and autonomous mobile robots.

Monitoring Beehive Sound Levels with Arduino-based System

Ramis Kulmukhametov
Intelligent Robotics Department, Kazan Federal University, 420008, Kazan, Russian Federation

Ramil Safin
Intelligent Robotics Department, Kazan Federal University, 420008, Kazan, Russian Federation

Tatyana Tsoy
Intelligent Robotics Department, Kazan Federal University, 420008, Kazan, Russian Federation

Kuo-Hsien Hsia
National Yunlin University of Science and Technology, Taiwan

Evgeni Magid
Intelligent Robotics Department, Kazan Federal University, 420008, Kazan, Russian Federation
Email: RNKulmukhametov@stud.kpfu.ru, tt@it.kfu.ru, khhsia@yuntech.edu.tw, magid@it.kfy.ru

Abstract

An automated beekeeping is a promising approach to addressing various issues associated with a beekeeping. Among primary problems, a swarming procedure stands out as a major concern. An uncontrolled swarming can lead to significant financial losses. During a swarming period a potential for losing a bee swarm is high, therefore a noise monitoring at this period gains a significant importance. Our long term research aims at a development of an intelligent monitoring system for beehive conditions, based on a hive-generated noises analysis. This paper presents an experiment that collected data about acoustic characteristics of bee states using an Arduino microcontroller and a MAX9814 microphone module. The obtained data analysis is discussed.

Keywords: beehive condition monitoring, acoustic level analysis, Arduino microcontroller, microphone, anthropogenic disturbance

1. Introduction

Currently, beekeepers are increasingly facing an issue of a rising mortality among honeybees, attributed to various factors [1]. This problem carries significant ecological and economic consequences that may cause a heavy impact on a biodiversity of wild plants [2] and agricultural production [3], [4]. One of the most destructive phenomena in this field is Colony Collapse Disorder, in which honeybees abandon their hives [5], [6], [7]. This colony collapse or demise is a result of multiple negative factors that can act independently, combine, or amplify one another [8]. A mortality of honeybees can be attributed to various factors, which can be classified into two main groups: a natural factor and a anthropogenic factor. Natural factors include bee infestations by Varroa mites and a swarming. The former are major parasites of honeybees, leading to a reduced hive viability and a decreased honey production [9]. A swarming, on the other hand, is a natural process of dividing a bee colony into two parts [10]. During the swarming, some worker bees and one or several queens leave a current nest to establish a new colony elsewhere. However, this process can pose a challenge for beekeepers as it involves a loss of some bees and, consequently, a decline in honey production.

To deal with Varroa mite a number of methods were proposed including computer vision approaches [11], [12], [13] and real-time visualization systems [14] that allow monitoring of honeybee activity by counting their entry and exit from a beehive [15], [16]. Sensor fusion [17], [18] and fiducial markers [19] can also be incorporated into the monitoring process to provide more robust and accurate measurements. Researchers identified a correlation between the swarming and a hive weight, a sound, a temperature, and a humidity inside a hive [20]. Particularly noteworthy is an analysis of sounds produced by honeybees, which has proven to be one of the most effective methods for detecting the swarming [21]. Scientific studies in this field demonstrated a correlation between specific sound frequencies and the swarming process [22], [23], making an acoustic analysis of bee sounds a promising avenue for monitoring and studying a condition of bee colonies.

This article presents results of a pilot experiment that analyzed a hive noise under normal conditions and after

Ramis Kulmukhametov, Ramil Safin, Tatyana Tsoy, Kuo-Hsien Hsia, Evgeni Magid

bee disturbances. It is a part of our long term goal of constructing an IoT-based monitoring system that could improve a productivity of honeybees by a timely detection of various events inside beehives and reporting results of AI-based analysis to beekeepers.

2. Experimental Setup

The Arduino UNO microcontroller acts as a main processor, performing key monitoring functions of the experimental setup. The microcontroller receives data from MAX9814 microphone module. The microcontroller and the microphone were placed together into a cardboard box with a power and data wire connected to a notebook computer. The box was placed under a hive cover, above hive frames. The setup and the device placement within a beehive are depicted in Fig. 1.

a   b
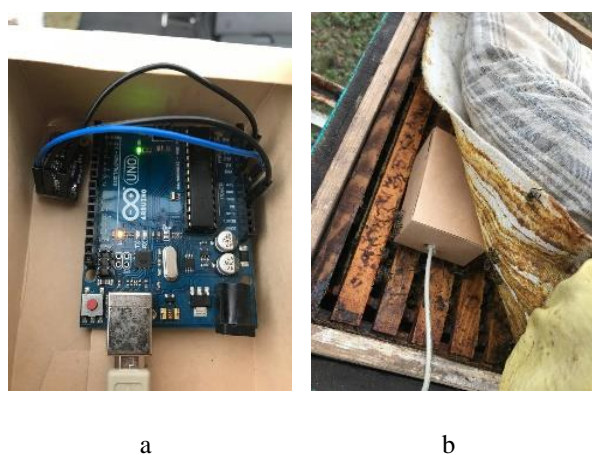
Fig. 1. Experimental setup: (a) the device; (b) the container placement inside a beehive

The device was programmed using the Arduino open-source software development platform. Arduino IDE was used for programming the microcontroller and reading data from the microphone. Upon powering up the microcontroller, the program initializes variables and opens a serial port at a baud rate of 115200. Data are transmitted through UART communication, and obtained by the microphone values are converted to decibels. A flow of the program is depicted in Fig. 2.

3. Results

A pilot experiment targeted to examine a state and a behavior of bees in terms of acoustic changes under artificial human-made disturbances. The pilot experiment was conducted on October 8, 2023, at 14:30:00 of the local time. By this time of the year, bees completed their main honey gathering season and started preparations for the upcoming winter season.
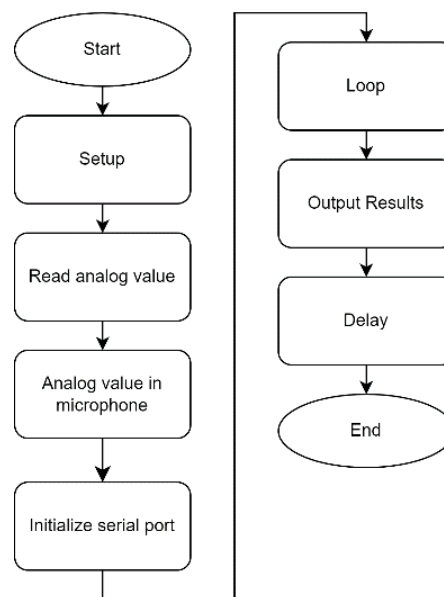
Fig. 2. Acoustic monitoring flowchart

The device was placed into a beehive and the microphone collected acoustic data from bees. The experiment was divided into two phases, 30 seconds each. The first phase corresponded to normal conditions of the beehive. In the second phase an anthropogenic disturbance was generated by knocking the hive for 3 seconds in order to provoke a response from the bees. The phase aimed to investigate changes in the bees' behavior.
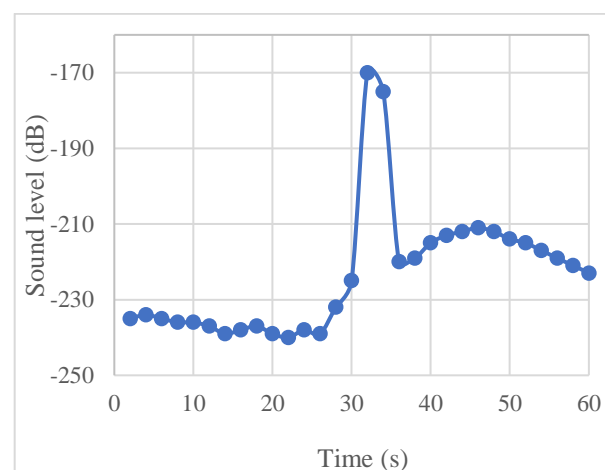
Fig. 3. An acoustic level at a human disturbance event

Fig. 3 presents experimental results of acoustic levels that were measured within the beehive for one minute. It is worth noting that the measurements were taken during a cool weather, which could have affected the acoustic level, making it lower compared to the main honey harvesting period. Initially, in a state of a rest, the measured acoustic level was -240 dB. The highest recorded acoustic level was -170 dB, which occurred at the beginning of the second phase (14:30:31) when the

beehive was subjected to a 3-seconds knocking. At this moment, the bees transitioned to an alarmed state, that was reflected in the acoustic level of -210 dB, which gradually decreased over several seconds. Interestingly, when the device was removed from the beehive, there was a clustering of bees around its location. This behavior could be interpreted as an attempt of the bees to protect their colony. They formed a small cluster around the device, and as their stings proved ineffective against the potential threat, they attempted surrounding the "*intruder*" with their bodies, creating a defensive barrier. This behavior impacted the acoustic level, which increased by several decibels in the period from 14:30:27 to 14:30:29, prior to the anthropogenic disturbance.

Based on the experimental observations, it was concluded that while the device could collect valuable acoustic data, a safer placement of a microphone and the processing block should be selected (which could be done using a virtual environment [24]), e.g., on an outer surface of a back wall of a beehive that is opposite a bees' entrance. This would protect the device from overheating and other factors generated by the bees themselves and ensure a more reliable data acquisition.

4. Conclusions

In this study, a part of the bee monitoring system designed to collect data on an acoustic level within a beehive was successfully developed and tested. The experiment aimed to study a bees' behavior under conditions of an anthropogenic disturbance. This experiment represents a step towards a better understanding of bees behavior and their response to interventions. Based on the experimental observations, it was also concluded that a proper selection of the monitoring device and sensors' location is crucial.

Acknowledgement

References

1. A.-M. Klein and B. E. Vaissière, Importance of pollinators in changing landscapes for world crops, Proceedings of the royal society B: biological sciences, Vol. 274, 2007, pp. 303-313.

2. D. A. Stanley, S. M. Msweli and S. D. Johnson, Native honeybees as flower visitors and pollinators in wild plant communities in a biodiversity hotspot, Ecosphere, Vol. 11(2), 2020, e02957.

3. D. L. Gazzoni and J. V. G. R. P Barateiro, Soybean yield is increased through complementary pollination by honey bees, Journal of Apicultural Research, 2023, pp. 1-12.

4. G. MacInnis and J. R. K. Forrest, Pollination by wild bees yields larger strawberries than pollination by honey bees, Journal of Applied Ecology, Vol. 56(4), 2019, pp. 824-832.

5. H. M. Al-Solami, N. A. Alkenani, A. G. Alghamdi, M. M. M. Ahmed, K. Javeed and A. A. Dar, Influence and Management of Colony Collapse Disorder (CCD) Damaging European Honeybee Apis mellifera, Specialusis Ugdymas, Vol. 2(43), 2022, pp. 3117-3130.

6. Nikita, A. Grover, P. Kalia, R. Sinha and P. Garg, Colony collapse disorder: A peril to apiculture, Journal of Applied and Natural Science, Vol. 14(3), 2022, pp. 729-739.

7. P. A. Woods, Illustrative Document of the Colony Collapse Disorder, 2022.

8. F. Requier, L. Garnery, P. L. Kohl, H. K. Njovu, C. W. W. Pirk, R. M. Crewe and I. Steffan-Dewenter, The conservation of native honey bees is crucial, Trends in ecology & evolution, Vol. 34(9), 2019, pp. 789-798.

9. K. S. Traynor, F. Mondet, J. R. de Miranda, M. Techer, V. Kowallik, M. A. Y. Oddie. P. Chantawannakul and A. McAfee, Varroa destructor: A complex parasite, crippling honey bees worldwide, Trends in parasitology, Vol. 36(7), 2020, pp. 592-606.

10. R. A. Ellis, T. Weis, S. Suryanarayanan and K. Beilin, From a free gift of nature to a precarious commodity: Bees, pollination services, and industrial agriculture, Journal of Agrarian Change, Vol. 20(3), 2020, p. 437-459.

11. S. Sevin, H. Tutun and S. Mutlu, Detection of varroa mites from honey bee hives by smart technology var-gor: a hivemonitoring and image processing device, Turkish Journal of Veterinary & Animal Sciences, Vol. 45(3), 2021, pp. 487-491.

12. G. J. Tu, M. K. Hansen, P. Kryger and P. Ahrendt, Automatic behaviour analysis system for honeybees using computer vision, Computers and Electronics in AgriculR12re, Vol. 122, 2016, pp. 10-18.

13. M. Roy, R. A. Boby, S. Chaudhary, S. Chaudhury, S. D. Roy, S. D. and S. K. Saha, Pose estimation of texture-less cylindrical objects in bin picking using sensor fusion. In 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), October 2016, pp. 2279-2284.

14. E. Magid, R. Lavrenov, T. Tsoy, M. Svinin and R. Safin, Real-time Video Server Implementation for a Mobile Robot. Proceedings of 11th International Conference on Developments in eSystems Engineering (DeSE), 2018, p. 180-185.

15. C. Chen, E.-C. Yang, J.-A. Jiang and T.-T. Lin, An imaging system for monitoring the in-and-out activity of honey bees, Computers and electronics in agriculR12re, Vol. 89, 2012, pp. 100-109.

16. T. N. Ngo, K.-C. Wu, E.-C. Yang and T.-T. Lin, A real-time imaging system for multiple honey bee tracking and activity monitoring, Computers and Electronics in Agriculture, Vol. 163, 2019, p. 104841

17. R. Safin, T. Tsoy, R. Lavrenov, I. Afanasyev and E. Magid, Modern Methods of Map Construction Using Optical Sensors Fusion. International Conference on Artificial Life and Robotics (ICAROB 2023), 2023, pp. 167-170.

18. E. A. Martínez-García, Cyber-Physical Robotics: Real-Time Sensing, Processing and Actuating. Cyber-Physical Systems for Industrial Transformation. CRC Press, 2023, pp. 57-74.

19. T. Tsoy, R. Safin, R. Sultanov, S. K. Saha and E. Magid, Recommended criteria for qualitative comparison of fiducial markers performance. Siberian Conference on Control and Communications, (SIBCON 2022), 2022, p. 1-5.

20. W. Hong, B. Xu, X. Chi, X. Cui, Y. Yan and T. Li, Long-term and extensive monitoring for bee colonies based on

internet of things, IEEE Internet of Things Journal, Vol. 7(8), 2020, pp. 7148-7155.

21. A. Terenzi, S. Cecchi and S. Spinsante, On the importance of the sound emitted by honey bee hives, Veterinary Sciences, Vol. 7(4), 2020, p. 168.

22. K. I. Dimitrios, C. V. Bellos, K. A. Stefanou, G. S. Stergios, I. Andrikos, T. Katsantas and S. Kontogiannis, Performance Evaluation of Classification Algorithms to Detect Bee Swarming Events Using Sound, Signals, Vol. 3(4), 2022, pp. 807-822.

23. G. Voudiotis, S. Kontogiannis and C. Pikridas, Proposed smart monitoring system for the detection of bee swarming, Inventions, Vol. 6(4), 2021. p. 87.

24. A. Zagirov, A. Apurin, E. Chebotareva, Modeling of Human Actions in a Collaborative Robotic Space Using AR601M Humanoid Robot: Pilot Experiments in the Gazebo Simulator. International Conference on Artificial Life and Robotics (ICAROB 2023), 2023, pp. 163-166.

Authors Introduction

Mr. Kulmukhametov Ramis

He received a BSc degree from Ufa State Aviation Technical University in 2022. Currently, he is a first-year student of Master degree program in Intelligent Robotics at the Institute of Information Technology and Intelligent Systems at the Kazan Federal University. He is also a professional beekeeper.

Mr. Ramil Safin

Ramil Safin is a Ph.D. student and a research assistant at the Laboratory of Intelligent Robotic Systems, Institute of Information Technology and Intelligent Systems, Kazan Federal University, Russia. He teaches the following courses: «Fundamentals of technical vision», «Robotic systems sensors», «Algorithms and data structures» His main interest is computer vision for robotic applications.

Ms. Tatyana Tsoy

In 2012 she graduated from International Area Studies Master's Degree Program at the University of Tsukuba. In 2023 completed a PhD in Computer science and information processes at the Institute of Information Technology and Intelligent Systems (ITIS) of Kazan Federal University.

Associate Professor Kuo-Hsien Hsia

He received a Ph.D. degree in Electrical Engineering at the National Sun Yat-Sen University, Kaohsiung, Taiwan. He works at the National Yunlin University of Science and Technology.

Professor Evgeni Magid

He is a Head of Intelligent Robotics Department, a founder, and a Head of Laboratory of Intelligent Robotic Systems (LIRS) at Kazan Federal University, Russia. Senior member of IEEE. He obtained Ph.D. degree in Robotics Engineering from the University of Tsukuba, Japan.

Restoration of Guqin Music by Deep Learning Methods

Takashi Kuremoto[1], Kazuma Fujino[1], Hirokazu Takahashi[1], Shun Kuremoto[2, 3],
Mamiko Koshiba[2], Hiroo Hieda[3], Shingo Mabu[2]

[1]Nippon Institute of Technology, Gakuendai 4-1, Miyashiro-Machi, Saitama, 345-8501, JAPAN
[2]Yamaguchi University, Tokiwadai 2-16-1, Ube, Yamaguchi, 755-8611, JAPAN
[3]Institute of Future Engineering, Fukagawa 2-6-11, Koto-ku, Tokyo, 135-8473 JAPAN
E-mail: kuremoto.takashi@nit.ac.jp

Abstract

Guqin （古琴）music played an important role in the history of Asia cultures. The notation of Guqin ancient music remained more than 600, however, only about 100 of them are played in nowadays. The reason is that the handwritten Guqin notations named "Jianzi Pu" is hard to be understood, however, we challenge to restore the Guqin music by deep learning methods and few Jianzi Pu images. VGG16 and YOLOv5 were adopted in the recognition experiment for Guqin music restoration. For a well-known Guqin music "Sen-O-So"（仙翁操）, 55 kinds of single characters of Jianzi Pu and 4,951 images of them were collected from 23 kinds of Sen-O-So version found by the Internet or obtained by image processing such as rotation, enlarge (zoom-in), reduce (zoom-out), filtering, etc. The average accuracies of VGG16 and YOLOv5 were 87.50% and 88.47% for the test data, respectively. Additionally, it was realized an online output of Guqin music as its output of audio or video forms by YOLOv5 in this study.

Keywords: deep learning, VGG16, YOLOv5, SVM, Guqin, Jianzi Pu, AI music

1. Introduction

Many kinds of ancient cultures and civilizations have disappeared in the past. A Chinese stringed instrument named "Guqin" （古琴）(or "Qin"（琴）) (Figs. 1 and 2) has a history more than 3,000 years, however, few players in the world exists in the last century. Recently, more Chinese youngers are learning this ancient musical instrument, however, it is very difficult to use its notation, named "Jianzi Pu", which is a kind of handwritten character using Chinese character and its original characters (see Figs. 3 and 4). There are more than 600 Guqin musical pieces in the history, meanwhile, the performed pieces in nowadays only about 100 [1], [2]. These Guqin music are translated from Jianzi Pu to modern staff notation by experts of Guqin players, and this translation process is called "Dapu" [3].

 Jianzi Pu recognition by computer has been studied since last century in China [4], [5], [6], however, the practical results are not obtained.
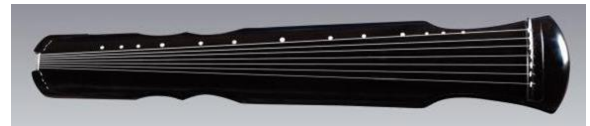
Fig. 1 A musical instrument: Guqin (7 strings Qin).

Fig. 2 Playing Guqin.

Takashi Kuremoto, Kazuma Fujino, Hirokazu Takahashi, Shun Kuremoto, Mamiko Koshiba, Hiroo Hieda, Shingo Mabu
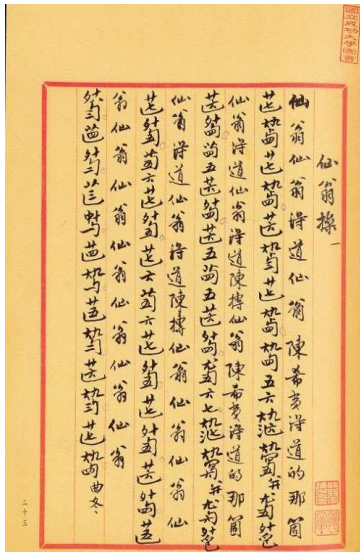
Fig. 3 A sample of Guqin musical notation "Sen-O-So".
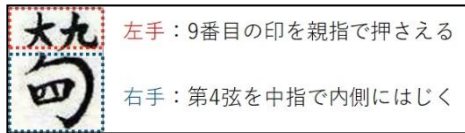
Fig. 4 A sample of a single Jianzi Pu and its meaning in Japanese.

In our previous work [7], [8], machine learning methods such as support vector machine (SVM), deep learning models VGG16, ResNet50, and the hybrid models VGG16 with SVM, ResNet50 with SVM were adopted to be classifiers of single character of Jianzi Pu. In the comparison experiments using a dataset of a well-known Guqin music "Sen-O-So"（仙翁操）, VGG16 with SVM had the highest average accuracies to 15 kinds of single characters data which were the first and the second line of Sen-O-So Jianzi Pu in 1,500 images, i.e., 99.11% in training and 88.33% for unknown data.

In this study, we built a bigger dataset for total single characters of Sen-O-So Guqin notation. 55 kinds of images of single characters were collected from 23 kinds of original Jianzi Pu from the Internet, and 4,951 single characters were obtained including data augmentation by image processing such as rotation, reversion, size enlargement (zoom in), size reduction (zoom out), and filtering. To realize online recognition, deep learning model YOLOv5 [9], [10] was adopted into the system as a classifier. The validation accuracy of YOLOv5 after 300 training arrived at 88.47% which was higher than the case of VGG16 87.50%. Additionally, audio/video data of single characters of Sen-O-So were created by Japan Society for the promotion of Guqin [2], and they were matched to the recognition results by YOLOv5.
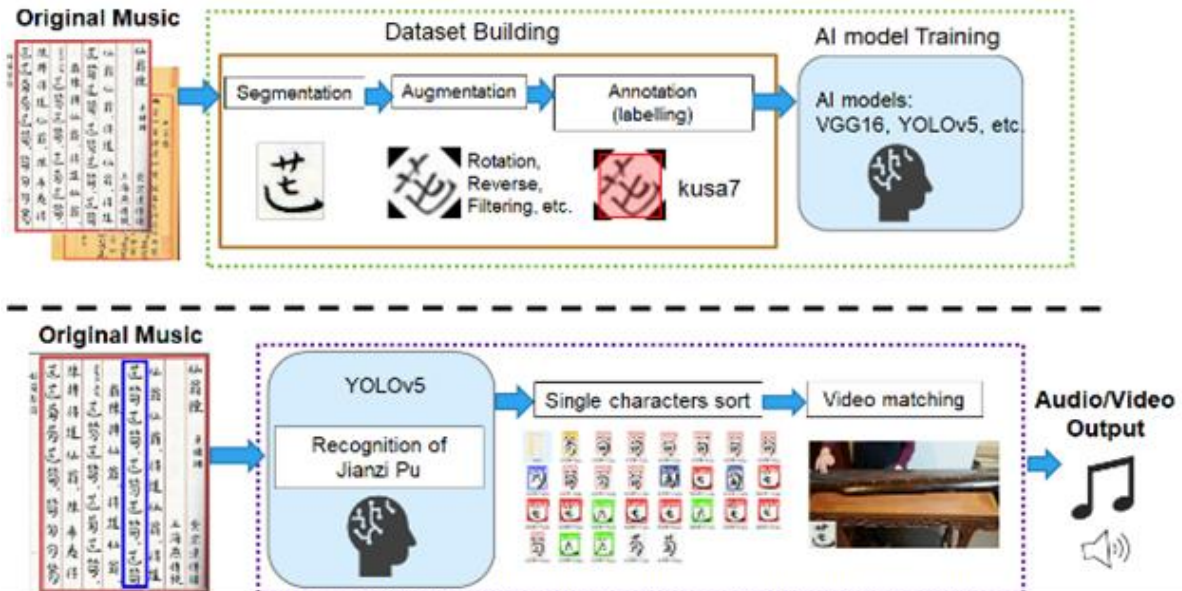
Fig. 5 A restoration system of Guqin music.

2. A Guqin Music Restoration System

A Guqin music restoration system, as shown in Fig. 5, is proposed in this study. It is aimed to realize that when a Jianzi Pu image is presented to the system, Guqin music in audio/video forms are output. The system is constructed by a fine-tuned deep learning model YOLOv5 [9], [10] and a matching process of the output of YOLOv5 and a single character database in audio and video forms.

Fig. 6 A sample of 55 single characters of Sen-O-So.

2.1. *A Database of Jianzi Pu*

To train the deep learning models, a database of single characters of a well-known Guqin music "Sen-O-So" in Jianzi Pu form is built at first. 23 versions of Sen-O-So Jianzi Pu images were collected by the Internet. 55 kinds of single characters were segmented from the Sen-O-So notation, and totally 343 images of single characters were obtained. Data augmentation is adopted by image processing such as rotation, reflection, parallel translation, enlarge (zoom-in), reduce (zoom-out), filtering to the original single character images. As the result, 4,951 images of single characters were obtained as a database of Jianzi Pu. A sample of the original single character images and a sample of data augmentation were shown in Fig. 6 and Fig. 7, respectively.

ファイル名	画像	画像処理の詳細
7 (14).jpg		元の画像
7 (14).jpg_**30**.jpg		高さと幅を30pxに変更
7 (14).jpg_**100**.jpg		高さと幅を100pxに変更
7(14).jpg_**200**.jpg		高さと幅を200pxに変更
7 (14).jpg_**blur**.jpg		カーネルサイズ（7，7）標準偏差1.5で平滑化
7 (14).jpg_**brighter**.jpg		色の値を1.5倍
7 (14).jpg_**darker**.jpg		色の値を0.9倍
7 (14).jpg_**noise**.jpg		ランダムな場所を対象にRGBランダムに選んでその値をランダムな少数分±する
7 (14).jpg_**r45**.jpg		反時計回りに45度回転
7 (14).jpg_**r90**.jpg		反時計回りに90度回転
7 (14).jpg_**r135**.jpg		反時計回りに135度回転
7 (14).jpg_**r180**.jpg		反時計回りに180度回転
7 (14).jpg_**r225**.jpg		反時計回りに225度回転
7 (14).jpg_**r270**.jpg		反時計回りに270度回転
7 (14).jpg_**r315**.jpg		反時計回りに315度回転

Fig. 7 A sample of data augmentation.

2.2. *Deep Learning Models*

A well-known deep learning model VGG16 [7], [8] and an online object recognition model YOLOv5 [9], [10] were investigated their recognition accuracies to single characters of Jianzi Pu. The state of art of YOLO is YOLOv8, and we are investigating its performance for Jianzi Pu recognition recently.

3. Experiments and Results

Using the database of Sen-O-So described in Section 2, VGG16 and YOLOv5 were fine-tuned with 300 epochs. The average recognition accuracies of the single characters were 87.50% and 88.47% respectively.

The change of accuracies in training of two models is shown in Fig. 8. It can be confirmed that the accuracies (training accuracy and validation accuracy) of VGG16 were converged from 200 epoch (Fig. 8a), meanwhile, it converged from 100 epoch in the case of YOLOv5 (Fig. 8b).

Takashi Kuremoto, Kazuma Fujino, Hirokazu Takahashi, Shun Kuremoto, Mamiko Koshiba, Hiroo Hieda, Shingo Mabu
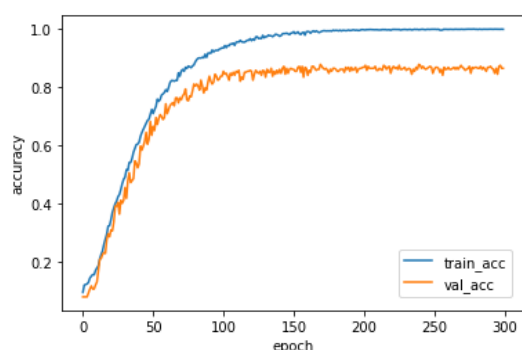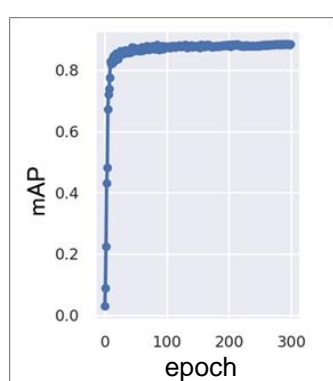
a. VGG16

b. YOLOv5

Fig. 8 Learning curves of VGG16 and YOLOv5.

Fig.9 A sample of audio/video data matched to a single character of Jianzi Pu.
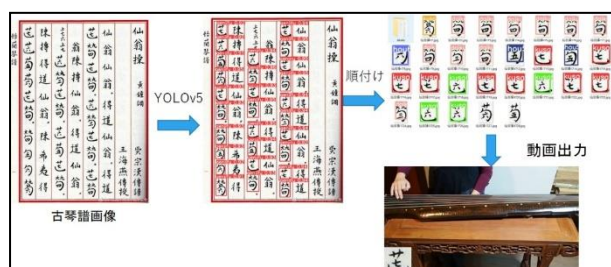
Fig.10 A sample of the processing result of the proposed system.

Single characters of Jianzi Pu were played by a professional Guqin player, Mrs. Yuki Takei [2] and recorded in audio/video forms as shown in Fig. 9. The order of single characters in a Guqin musical notion was given by their position detected by YOLOv5. By matching the recognized single characters to audio/video data, the system output music and videos as shown in Fig. 10.

4. Conclusion

To restore Guqin （古琴）music written in Jianzi Pu notation, a deep learning system constructed by YOLOv5 was developed in this study. The main contribution includes the making of dataset of a well-known Guqin music Sen-O-So (仙翁操) and the restoration system developed with artificial intelligence (AI) technology. The experiment results showed the effectiveness of the proposed system, and the future work of this study is to build more single character datasets of Jianzi Pu using more Guqin musical notations.

Acknowledgements

References

1. China Script, "The difference between Qin and So", http://chugokugo-script.net/chugoku-bunka/kin.html (2016) (in Japanese)
2. Japan Society for Promotion of Guqin: https://www.guqin.jp/about
3. R. Takao, "A survey and analysis of the Dapu process in Qin music", Journal of the Musicological Society of Ochanomizu University, No. 10, 2008 pp.40-45 (in Japanese)
4. Z.X. Pan, C. L. Zhou: "Text Segmentation and Extraction from Images of Guqin Jianzi Pu", Mind an Computation, Vol.1, No.2, pp. 286-295 (2007) (in Chinese)
5. Chen Shi: "Guqin Notation and Music Style Recognition", Computer Science (2016)
6. 王利，孫洋，罗兆麟，張輝：「AI自动翻译"减字谱"——以《流水》和《不染》为例」，Art Education, Vol.3 (2019) (in Chinese)
7. B. Yang, M. Sato, S. Kuremoto, M. Koshiba, S. Mabu, H. Hieda, T. Kuremoto, "Recognition of Guqin music by deep learning methods", Proceedings of the Electronics, Information and Systems Conference Electronics, Information and Systems Society, I.E.E. of Japan，pp.512-515 (2022) (in Japanese)
8. Bowen Yang, Shun Kuremoto, Mamiko Koshiba, Shingo Mabu, Hiroo Hieda, Takashi Kuremoto, "Recognition of Guqin Notation using Deep Learning", in Proceedings of Innovative Application Research and Education (ICIARE2022), pp.61-64 (2022)
9. J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You Only Look Once: Unified, Real-Time Object Detection, aiZiv:1506.02640v5 (2016)
10. G. Jocher, https://github.com/ultralytics/yolov5/ (2020)

Authors Introduction

Prof. Dr. Takashi Kuremoto

He received his B.E. degree from the University of Shanghai for Science and Technology, China, in 1986, M.E. and Ph.D. degrees from Yamaguchi University, Japan, in 1996 and 2014, respectively. He worked as a system engineer at the Research Institute of Automatic Machine of Beijing from 1986 to 1992 and until 2021, he was an Assistant Professor at Yamaguchi University. He was an Academic Visitor of the School of Computer Science, the University of Manchester, U.K., in 2008. Currently, he is a Full Professor with the Department of Information Technology and Media Design, Nippon Institute of Technology, Japan. His research interests include artificial neural networks, fuzzy, evolutionary computation, time series forecasting, and intelligent systems

Mr. Kazuma Fujino

He received his B.E. degree from Nippon Institute of Technology, Japan, in 2023. He works at SET Software Co., Ltd., Tokyo, Japan, currently.

Mr. Hirokazu Takahashi

He received his B.E. degree from Nippon Institute of Technology, Japan, in 2023.

Mr. Shun Kuremoto

He received his B. F. degree from Shanghai University of Finance and Economics, China, in 1989, and M. A. in Commerce degree from Waseda University, Japan, in 1997. Currently, he is a guest researcher of the Institute for Future Engineering, Japan, as well as Director of Japan Society of Promotions of Guqin, and a doctoral program student in Yamaguchi University, Japan.

Dr. Mamiko Koshiba

1. She received her B.E., M.E. and Ph.D. degrees from Tokyo University of Agriculture and Technology, Japan, in 2001, 2002 and 2005, respectively. She worked as a postdoctoral neuroscientist at National Center of Neurology and Psychiatry 2005 to 2007 and until 2013, she was a Visiting Associate Professor at Tokyo University of Agriculture and Technology and from 2013 to 2016 at Saitama Medical University, Currently, she was an Associate Professor of Yamaguchi University, Japan. Her research interests include application of artificial neural networks for neuronal system evolution and development.

Mr. Hiroo Hieda

He is a Fellow of the Institute for Future Engineering. Longtime, he did comprehensive research on science and art. Currently, he researches on the history of Qin (七絃琴) in early modern Japan, especially in the Edo period. He published a book named "A study to the history of Qin in early modern Japan" (近世琴学史攷) in 2018.

Prof. Dr. Shingo Mabu

He received his B.E and M.E. degrees from Kyushu University, Japan, in 2001 and 2003, respectively, and his Ph.D. degree from Waseda University, Japan, in 2006. From 2006 to 2007, he was a visiting lecturer at Waseda University, Japan, and from 2007 to 2012, an assistant professor at the same university. From 2012 to 2017, he was an assistant professor at Yamaguchi University, Japan, and from 2017 to 2020, an associate professor at the same university. He is currently a professor at the Graduate School of Sciences and Technology for Innovation at the same university.

Constructive Nurse Scheduling Using Reinforcement Learning Considering Variations in Nurse Work Patterns

Masato Nagayoshi
Niigata College of Nursing, 240 Shinnan-cho Joetsu, Niigata 943-0147, Japan
E-mail: nagayosi@niigata-cn.ac.jp

Hisashi Tamaki
Kobe University, 1-1 Rokkodai-cho, Nada-ku, Kobe, Hyogo 657-8501, Japan
E-mail: tamaki@al.cs.kobe-u.ac.jp

Abstract

It is very difficult to create a work schedule that satisfies all the different requirements in nurse scheduling. For this reason, numerous studies have been conducted on the nurse scheduling problem. However, the created shift schedule is often not practical as it is, because adjustments including various constraints and evaluation criteria are required. Therefore, we have proposed a work revision method using reinforcement learning in a constructive nurse scheduling system. In this paper, we investigate the feasibility of creating a practical work schedule even when there are nurses with different work pattern evaluations, i.e., the feasibility of creating a practical work schedule that takes each nurse's life stage into consideration.

Keywords: nurse scheduling, reinforcement learning, constructive search, work patterns, life stages

1. Introduction

Various studies have been conducted on the nurse scheduling problem [1], which is the creation of a shift schedule for nurses. However, for practical use, adjustments including various constraints and evaluation values are required, and the created shift schedule is often not practical as it is, so many head nurses still feel burdened by creating shift schedules [2]. Therefore, we have proposed a work revision method [3] using reinforcement learning [4] on a constructive nurse scheduling system [5].

In this paper, we investigate the feasibility of creating a practical work schedule even when there are nurses with different work pattern evaluations, i.e., the feasibility of creating a practical work schedule that takes each nurse's life stage into consideration. In addition, the work pattern evaluations of the obtained modified work schedules are visualized. Furthermore, none of the previous studies on nurse scheduling considered the evaluation of different work patterns.

2. Constructive nurse scheduling system

2.1. *Features*

The features of the constructive nurse scheduling system [5] are as follows.
1. The system creates a schedule for each day, starting from the first day.
2. The priority calculation can be extended to consider detailed conditions.
3. It does not consider the evaluation value for the entire shift schedule for a month.

2.2. *Work Revisions*

The constructive scheduling system considers only the basic constraints that would be required in a hospital with many nurses, and the possibility exists that a feasible solution that does not satisfy the head nurse is obtained. For this reason, Kurashige et al. [5] describe the following two procedures for the actual modification.
(1) A work shift of the nurse in the case that does not satisfy the head nurse is manually exchanged with a work shift of another nurse.
(2) A work shift of the nurse in the case that do not satisfy the head nurse is exchanged to other work shift as designated work shift, and the rescheduling is done.

Next, we introduce our proposed system [3] that learns this exchange procedure using reinforcement learning.

3. Work Revision Method Using Reinforcement Learning

3.1. *Problem Setting for Reinforcement Learning*

The shift schedule created by the constructive nurse scheduling system, which is created in order from the first day, satisfies the shift constraints (such as the number of nurses required for each day). On the other hand, the shift schedule for the entire scheduling period (e.g., one month) is checked, there may be several cases in which the nurse constraints (such as the limited number of workdays) are not satisfied for each nurse.

Therefore, the number of violations *Vnw* of work shift *w* is calculated as the number of days exceeding *UTnw*, the upper limit of the number of assignments of work shift *w* to each nurse *n*, from the work schedule, and a revision is repeated according to the following:

$$min \ \sum_n \sum_v V_{nw} \qquad (1)$$

The following procedure is to be used for one revision.
(1) Select a work shift w_0 that is the source of the exchange (usually the one with the most violations).
(2) Determine the nurse n_0 with the highest number of violations in the shift w_0.
(3) If the shift w_0 is the night shift, the shift w_0 with the highest number of violations, whether it is evening or late-night shift, is designated as w_0 for the nurse n_0.
(4) If there is a work shift that is below the lower limit of the number of assignments for the nurse n_0, that work shift w_1 is designated as a destination of the exchange shift. If not, day shift without the upper and lower limits of the number of assignments is used as the exchange.
(5) Determine the day d_0 with the highest priority among the days when the shift w_0 is exchanged to w_1 for nurse n_0.
(6) Deduce the group $g(j_0)$ in which the nurse n_0 is in charge of a job j_0, which is assigned as the shift w_0.
(7) Determine a nurse n_1 who belongs to group $g(j_0)$ and whose shift on the day d_0 is w_1. If there is more than one applicable nurse, determine the nurse n_1 with the highest priority among the nurses when the shift w_1 is exchanged to w_0 on d_0.
(8) The nurses n_0 and n_1 are exchanged their shifts on d_0.

In case there is no corresponding nurses in any of the procedures, the exchange is not valid. In addition, it is also not valid to undo a previous exchange.

Here, minimizing the number of violations is a very difficult problem, because the number of possible modifications depends on which work shift is being exchanged.

3.2. *RL Agent*

Q-learning [6] is applied to the proposed method to learn an appropriate exchange procedure. The state space of the RL agent consists of 4 dimensions: the previous exchange days (1 to 30), the total number of violations by all nurses for evening, late-night shift, and holiday: V_{nw} (w=1,2,3), to be a Markov decision process. The number of possible actions is 4, which is the exchange of evening, late-night, holiday, and night shift.

1 step is defined as 1 exchange including unsuccessful cases, 1 episode is defined as the time when the shift schedule reaches the target state, or 100 steps passed. Here, the target state is defined as the sum of violations for all nurses and shifts, $\sum_n \sum_v V_{nw} = 0$ (excluding over-holiday violations). The positive reinforcement signal $r_t = 10$ (reward) is given only when the target state is reached and the reinforcement signal $r_t = 0$ is given at any other steps. At the start of each episode, the shift schedule will be in its initial state.

4. Computational Experiments

Table 1. Evaluations of work pattern for 2 days.

shift on previous day	shift on the day			
	day	evening	late-night	holiday
day	15	1	13	11
evening	0	5	0	12
late-night	0	8	5	4
holiday	23	3	0	17

Table 2. Evaluations of work pattern for 2 days allowing transitions from the evening shift to the late-night shift.

shift on previous day	shift on the day			
	day	evening	late-night	holiday
day	15	1	13	11
evening	0	2	8	7
late-night	0	8	5	4
holiday	23	3	0	17

Table 3. Parameters for experiments

Parameter	Value
α_Q	0.1
γ	0.9
τ	0.1

4.1. *Nurse Scheduling Problem*

The extended method is applied to a nurse scheduling problem like that of Kurashige et al. [5]. First, a three-

shift system (day, evening, and late-night shift) is adopted, and the number of nurses is 23, including the head nurse. Furthermore, the number of positions is classified as 3 (head nurse, assistant head nurse, and general), the number of teams is 2 (A and B), and the skill level is 3 (experienced, mid-career, and new). The other constraints are outlined below.

- Restrictions on the number of nurses for each shift:
  (1) Required number of day shift on weekdays is greater than or equal to 10.
  (2) Required number of day shift for weekends and holidays is 5.
  (3) Required number of late-night shift is 5.
  (4) Required number of evening shift is 5.

Next, Table 1 shows the evaluation of work patterns for 2 days with M = 2. The evaluation of work patterns can be set using the current work schedule as a reference.

Here, we investigate the feasibility of creating a feasible modified work schedule in the situation where all nurses have the work pattern evaluations shown in Table 1 (Situation 1) and only Staff 6 and Staff 17, the experienced nurses of Team A and Team B, respectively, have the work pattern evaluations allowing transitions from the evening shift to the late-night shift shown in Table 2 (Situation 2).

4.2. *RL Agen
t*

In the state space of the RL agent, the total number of violations is assumed to be [0, 2] and can take 3 states.

The computational experiments have been done with parameters as shown in Table 3. In addition, all initial Q-values are set at 5.0 as the optimistic initial values.

Fig 1. Required steps for accomplishing the task.

Fig. 2. Total number of violations when accomplishing the task.

4.3. *Results*

The average of the numbers of steps required for accomplishing the task and the average of the total number of violations when accomplishing the task in Situations 1 and 2 were observed during learning over 20 simulations, as described in Fig. 1 and Fig. 2, respectively.

It can be seen from Fig. 1 and Fig. 2 that (1) in Situation 1, the target state was reached in about 5 steps, and the number of violations was reached at 3, (2) in Situation 2, the target state was reached in about 2 steps, and the number of violations was reached at 2. Thus, we confirmed that a feasible modified work schedule could be created not only in Situation 1, but also in Situation 2, where nurses have different work pattern evaluations.

The average of each nurse's work pattern evaluations for each day (the average of the previous day's and the day's work pattern evaluation, and the day's and the next day's work pattern evaluation) and the total average of the work pattern evaluations on the modified work schedule for Situations 1 and 2, respectively, are shown in Figs. 3 and 4. In addition, the 24, 25, and 26 days of Staff 6, indicated by the red boxes in Fig. 4, represent the average of the pattern evaluation of the late-night shift, evening shift, and late-night shift, respectively, and the 7, 8, and 9 days of Staff 17 represent the average of the work pattern evaluation of the holiday, evening shift, and late-night shift, respectively.

It can be seen from Fig. 3 and Fig. 4 that (1) the total average of the work pattern evaluations is the same for Situation 1 and Situation 2.

Although additional experiments are needed, setting the work pattern evaluations to values that each nurse him/herself desires will enable him/her to realize a work style (work pattern) that suits him/her life stage and, in addition, will help him/her to understand the work schedule creation process.

Fig. 3 Average evaluations of work pattern in the revised work schedule in Situation 1.

Masato Nagayoshi, Hisashi Tamaki

Fig. 4 Average evaluations of work pattern in the revised work schedule in Situation 2.

5. Conclusion

In this paper, we confirmed that the proposed method can create a feasible modified work schedule in a situation where there are nurses with different work pattern evaluations, specifically, where there are two nurses who can work from evening shift to late-night shift. In addition, the work pattern evaluations of the obtained modified work schedules were visualized and the total average of the work pattern evaluations in the situation was compared with the total average of the work pattern evaluations in the normal situation where evening shift to late-night shift is not possible, and we confirmed that the evaluation was equivalent to that in the normal situation.

Our future projects include to confirm the effectiveness of the proposed method in situations where nurses evaluate even more diverse work patterns, etc.

Acknowledgements

References

1. A. Ikegami, A Model for the Nurse Scheduling Problem, IPSJ SIG Notes, **5**, 1-6, 1996. (in Japanese)
2. H. Adachi, S. Nakamura, M. Nagayoshi and N. Okamura, A Survey on The Present Status about Required Time and Recognition of Supports for Managers to Prepare a Work Timetable in The Medical Treatment and Supervision Act Ward, Journal of Japan Academy of Psychiatric and Mental Health Nursing, **30** (1), 59-65, 2021. (in Japanese)
3. M. Nagayoshi and H. Tamaki, An Approach of Exchanging Work Shifts Using Reinforcement Learning on a Constructive Nurse Scheduling System, Journal of Robotics, Networking and Artificial Life, **9**(2), 154-158, 2022.
4. R. S. Sutton and A. G. Barto, Reinforcement Learning, A Bradford Book, MIT Press, Cambridge, 2018.
5. K. Kurashige, T. Hashimoto and Y. Kameyama, Nurse Scheduling System in Consideration of Versatility, Journal of Japan Industrial Management Association, **56**(2), 109-120, 2005.
6. C. J. C. H. Watkins and P. Dayan, Technical note: Q-Learning, *Machine Learning* **8**, 279-292, 1992.

Authors Introduction

Dr. Masato Nagayoshi

He is an Associate Professor of Niigata College of Nursing. He graduated from Kobe University in 2002 and received Master of Engineering from Kobe University in 2004 and Doctor of Engineering from Kobe University in 2007. IEEJ, SICE, ISCIE member.

Dr. Hisashi Tamaki

He is a Professor, Graduate School of Engineering, Kobe University. He graduated from Kyoto university in 1985 and received Master of Engineering from Kyoto University in 1987 and Doctor of Engineering from Kyoto University in 1993. ISCIE, IEEJ, SICE, ISIJ member.

A Basic Study on Indicator of Transfer Learning for Reinforcement Learning

Satoshi Sugikawa
Osaka Institute of Technology, 1-79-1 Kitayama, Hirakata City, Osaka, 573-0196 Japan
Email: satoshi.sugikawa@oit.ac.jp

Kenta Takeoka
Osaka Institute of Technology, 1-79-1 Kitayama, Hirakata City, Osaka, 573-0196 Japan

Naoki Kotani
Osaka Institute of Technology, 1-79-1 Kitayama, Hirakata City, Osaka, 573-0196 Japan

Abstract

Reinforcement learning requires a lot of learning time for the agent to learn. Transfer learning is a method to reduce this learning time, but it has the problem that the user does not know which knowledge is effective in which environment until it is learned. Therefore, it is necessary for the user to consider the relationship between the source and destination when transferring knowledge. Therefore, this study proposes Indicator of adaptation criteria that can determine this relationship in advance. In simulations, we demonstrate the usefulness of the proposed method by using some example problems.

Keywords: Reinforcement learning, Transfer learning, Maze problems

1. Introduction

In recent years, machine learning has achieved rapid growth with results in various fields such as natural language processing and image processing. Reinforcement learning [1], [2] which is capable of self-learning, is expected to further develop in all fields in the future. Reinforcement learning has the problem that it requires a lot of learning time because the agent learns from scratch in a new environment by trial and error.

Many studies have been conducted to reduce learning time, one of which is transfer learning. Transfer learning is a method of adapting and reusing previously learned material from similar tasks for a new task. Since there is no need to relearn, learning time can be reduced. However, the effectiveness of the transferred material is not known until it is transferred and learned. Therefore, when transfer learning is performed, the user needs to consider the relationship between the transferee and the transferee source, but even then, there is a possibility that the learning will not be successful and negative transfers will occur [3]. Therefore, this study proposes several indicators of adaptive criteria that can discriminate in advance the validity of knowledge.

2. Reinforcement learning

Reinforcement learning is a branch of machine learning in which an agent learns through interaction with its environment. Reinforcement learning is characterized by the fact that the agent chooses an action to achieve a certain goal and receives a reward signal for that action as it learns. Optimal behaviour rules are learnt by repeating steps 1 to 4 below.

1. Agent observes the state
2. Decides on a course of action based on cues from measures
3. Memorizes experience of which states, which actions and which rewards
4. Seeks measures based on experience

Reinforcement learning is also formulated as a Markov decision process, expressed as M=(S,A,P,R) as follows.

- S: Set of states
- A: Set of actions
- P: State transition $p(s_t = s'|s_t = s, a_t = a)$
- R: Behaviour rules $\pi(s, a) = R(s_t = s, a_t = a)$

Reinforcement learning aims to acquire behaviour rules that maximize the expected reward.

Satoshi Sugikawa, Kenta Takeoka, Naoki Kotani.

2.1. *Q-learning*

In this study, Q-learning [3] was used as a reinforcement learning algorithm learning updates the Q-values that the agent associates with a combination of states and actions, and finds an action rule that selects the action with the maximum Q-value in each state. The formula for updating the Q-values is as follows.

$$Q(s_t, a_t) = Q(s_t, a_t) + \eta \\ * (R_{t+1} + \gamma \max Q(s_{t+1}, a) \\ - Q(s_t, a_t))$$

In this study, the ϵ-greedy method was also used as the agent's action selection method. In the ϵ-greedy method, the agent acts randomly with a probability of ϵ and selects an action with a probability of 1- ϵ with a maximum Q value.

3. Transfer learning

In transfer learning [4], [5], the goal is to reduce learning time by transferring knowledge learned at the transfer source as prior knowledge at the transfer destination with similar tasks. However, if there is no similarity between the source and destination, a negative transfer may occur. Therefore, the user must consider the similarity between the source and destination in advance. Therefore, it is necessary to determine the similarity between the source and destination to determine which knowledge is transferred to which destination.

In this study, we aim to formalize the adaptation criterion by determining the similarity between mazes using the maze problem.

3.1. *learning model*

The algorithm used for learning in reinforcement learning is Q-learning. There are five actions that the agent can take: up, down, left, right, left and right, and no action. The reward design is +1 if the agent reaches the goal, and -0.01 if it collides with a wall or a step elapse.

3.2. *transition method*

As the transfer method uses Q-learning, the transfer learning adopts the value function transfer type.
The Q-table obtained by the learning of the transfer source is transferred to the transfer destination agent.
The degree of re-use of the behavioural value function is adjusted using the transfer rate τ. The transfer rate τ is adjusted in the range of $0 < \tau < 1$. In this study, τ was set at 0.5.

$$Q^c(s, a) = Q^t(s, a) + \tau Q^s(s, a)$$

4. Similarity Indicators

The following five indices are used to compute the similarity.

4.1. *Similarity by pixel value*

In this calculation example, the similarity by pixel value is determined. As the mazes used in this study have the same size, the percentage of pixel values matched between the images is determined.

4.2. *Cosine similarity between maze sequences*

In this example calculation, the maze is transformed into arrays and the cosine similarity between the source and destination arrays is calculated. The transformed array is the same as the array used for training, with the maze walls set to 1 and the paths set to 0.
- s is a vectorization of the base maze
- t is a vectorization of the destination maze.

$$cos(s, t) = \frac{\sum_i s_i t_i}{\sqrt{\sum_i s_i^2} \sqrt{\sum_i t_i^2}}$$

4.3. *Similarity using maxQ as weights*

In this calculation example, maxQ is used to determine the similarity. The total of the differences between the mazes is calculated using maxQ of maze A, which is the source of the transition, as the weight. The sum of the differences between the mazes is calculated using maxQ of maze A as the weight, and divided by the total value of maxQ.

$$p = 1 - \frac{\sum_i (s_i - t_i) \max_a Q_{s_{i,a}}}{\sum_i \max_a Q_{s_{i,a}}}$$

4.4. *Cosine similarity using maxQ*

Finally, in this calculation example, the Q values obtained after learning are used to obtain the cosine similarity. The Q values obtained after normal learning in each maze $(12 \times 12 \times 5)$ are changed to an array of 12×12 by finding the maximum value (maxQ) in each square. However, this is not known before application as it is an outcome. It is calculated as part of the evaluation for measuring similarity.

$$cos(Max_a Q_s, Max_a Q_t) = \frac{\sum_i Max_a Q_{s_{i,a}} Max_a Q_{t_{i,a}}}{\sqrt{\sum_i Max_a Q_{s_{i,a}}^2} \sqrt{\sum_i Max_a Q_{t_{i,a}}^2}}$$

5. Simulation

Five mazes were prepared to test the usefulness of the index. Maze A is the original maze and was applied to mazes B, C, D, and E, respectively. Fig. 1, Fig. 2, Fig. 3, Fig. 4, and Fig. 5 show the mazes.
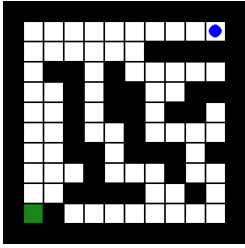
Fig.1 Maze A
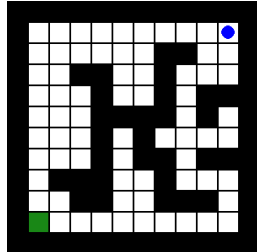
Fig.2 Maze B
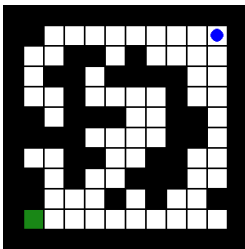
Fig.3. Maze C
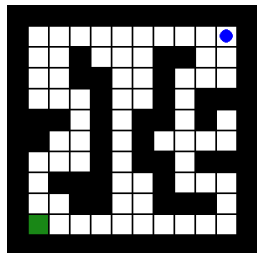
Fig.4 Maze D
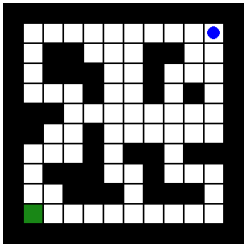
Fig.5 Maze E

The similarity of the mazes to each indicator is shown in Table 1.

Table 1 Similarity Results

	Maze A to Maze B	Maze A to Maze C	Maze A to Maze D	Maze A to Maze E
Similarity by pixel value	82.95	77.08	80.02	79.44
Similarity between maze sequences	80.29	75.96	78.71	77.66
Similarity using Qmax	98.86	60.96	82.49	77.37
Cosine similarity using Qmax	72.97	5.89	41.17	28.91

S The Similarity results show that maze B, maze D, and maze C have the highest similarity in that order. Similarity by pixel value, similarity between maze sequences, and similarity using Qmax, in that order. The use of Q-values allows us to focus on differences only in important locations. Fig. 7, Fig. 8, and Fig. 9 show results of the transfer learning. Transition from maze A to maze C is not shown in the figure because the goal was not achieved. Reinforcement learning is represented by the red line and transfer learning by the blue line.
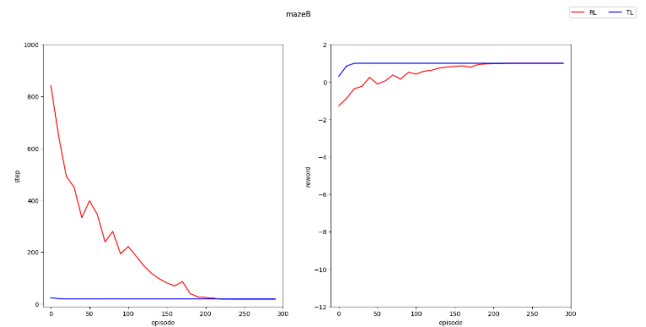
Fig.7. Transfer learning from maze A to maze B, left panel shows number of steps to goal and episodes, right panel shows total reward and number of episodes.
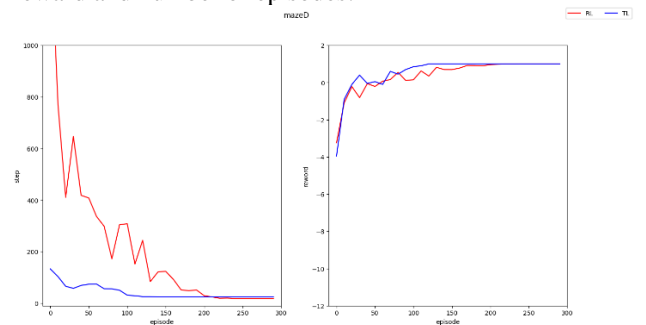
Fig.8. Transfer learning from maze A to maze D, left panel shows number of steps to goal and episodes, right panel shows total reward and number of episodes
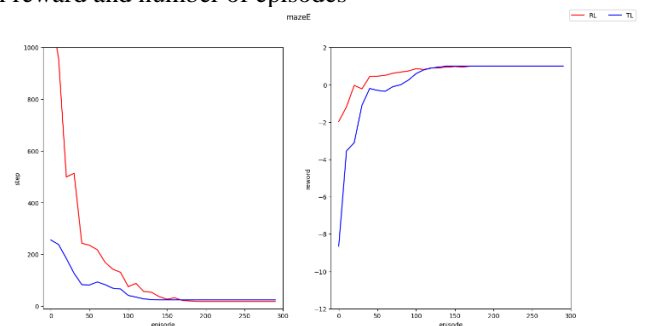
Fig.9. Transfer learning from maze A to maze E, left panel shows number of steps to goal and episodes, right panel shows total reward and number of episodes

From these figures, the goal is reached faster with transfer learning than with reinforcement learning. This indicates that transition learning is effective in these mazes. Based on the results in the figures and the cosine similarity using Qmax, it seems reasonable to focus on the Q value.

Satoshi Sugikawa, Kenta Takeoka, Naoki Kotani.

Fig. 10 shows the results of applying reinforcement learning to the mazes. Fig. 11, Fig. 12, Fig. 13 and Fig. 14 show the results of applying transition learning to each maze. Red indicates locations with high rewards. The reward of maze A was applied to each maze. Therefore, the maze was routed from the top towards the goal. This demonstrates the validity of this simulation.
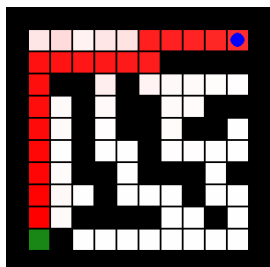
Fig.10 Result of Reinforcement Learning on maze A
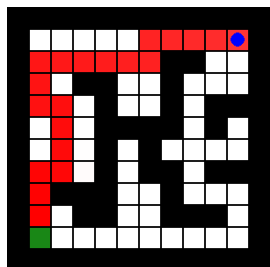
Fig.11 Result of Transfer learning on maze B
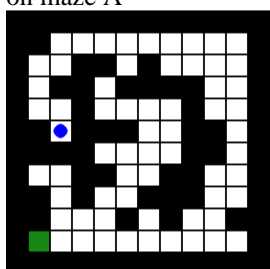
Fig.12 Result of Transfer learning on maze C
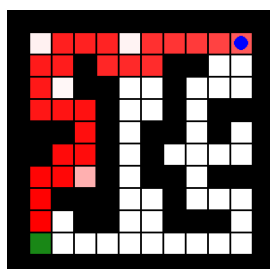
Fig.13 Result of Transfer learning on maze D
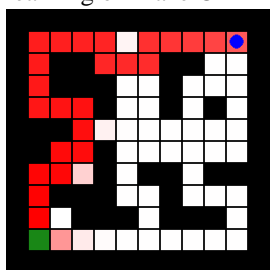
Fig.14 Result of Transfer learning on maze E

6. Conclusion

In this study, we conducted basic research on measures focusing on environmental similarity in reinforcement learning. Simulation results confirmed the validity of focusing on maxQ. The validity of the indicator was also confirmed to some extent. In the future, the creation of more accurate indicators is required.

References

1. Sutton, R, S., Barto, A, G: Reinforcement Learning: An Introduction. The MIT Pres（2007）

2. Leslie Pack Kaelbling. Michael L. Littman and Andrew W. Moore: Reinforcement Learning-A Survey. Journal of Artificial Intelligence Research.vol.4.237/285（1996）

3. C. J. C. H. Watkins: Learning from Delayed Rewards, PhD thesis, King's College, Cambridge, UK（1989）

4. Haitham B. Ammar, Karl Tuyls, Matthew E. Taylor, Kurt Driessens, Gerhard Weiss: Reinforcement Learning Transfer via Sparse Coding, Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems（2012）

5. Matthew E. Taylor, Peter Stone: Transfer Learning for Reinforcement Learning Domains, A Survey ,Journal of Machine Learning Research, vol.10,1633/1685（2009）

Authors Introduction

Dr. Satoshi Sugikawa

He received his Dr. Eng. degrees from Kobe University, Japan, in 2011. In 2013, he joined Osaka Institute of Technology, where he is currently an Assistant Professor. His research interests include Mathematical Optimization. He is a member of IEEJ

Mr. Kenta Takeoka

He received his Bachelor's degree in Engineering in 2022 from the Faculty of Information Science and Technology, Osaka Institute of technology in Japan. He is currently a master student in Osaka Institute of Technology, Japan

Dr. Naoki Kotani
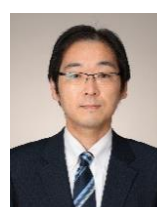
He received a Ph.D. degree from Osaka University, Osaka, Japan, in 2011. He is an assistant professor of Information Science and Technology at Osaka Institute of Technology. His research interests machine learning and robotics. Kotani is a member of the ISCIE, the SICE, the RSJ, the JSAI and hte IEEE.

Machine Learning Approach
to Predict Cooling Load for Existing Buildings

Makoto Ohara

Department of Information Technology, International Professional University of Technology in Osaka, 3-3-1 Umeda, Kita-ku, Osaka, 530-0001, Japan

Hideo Isozaki

Carbon Neutral Promotion Headquarters, Kobe University, 1-1, Rokkodai-cho, Nada-ku, Kobe 657-8501, Japan

E-mail: ohara.makoto@iput.ac.jp, isozaki@person.kobe-u.ac.jp
www.iput.ac.jp/osaka/

Abstract

The objective of this study is to predict air conditioning loads for existing buildings using operational data, weather forecasts and visitor forecasts. The proposed prediction method is based on a neural network approach. However, it is important to note that the proposed method does not learn the entire loads. Loads are divided into factors which can be predicted by traditional thermodynamics and factors which are subject to machine learning. The proposed method has been applied to an example instance using operational data from an underground mall in Kobe, and its validity has been confirmed.

Keywords: Cooling Load Prediction, Air Conditioning System, Existing Building Data, Machine Learning

1. Introduction

If air-conditioning load are accurately predicted, it is expected which the system is optimally operated, e.g. setting of changes in chilled water supply temperature, selection of heat source models, setting of operation times, for saving energy [1].

The purpose of this study is to predict future load of an air conditioning systems on the next day for which exists, using operational data of existing buildings, weather forecasts and visitor volume forecasts. The prediction algorithm is combining mathematical calculation and neural network (hereafter NN) methods. The proposed method has applied to an example based on operational data of an underground shopping mall in Kobe and compared with the actual measured cooling load.

Air conditioning load prediction for existing buildings differ from load calculations for new buildings in the followings.

1. Therefore historical operating data and corresponding load data exist, they can be used for forecasting by machine learning. The large amount of training data can be used for the load prediction without having to establish physical causal relationships between load calculation conditions and load values.
2. The actual situation regarding how the building is used is known, not the assumptions made at the time of design.
3. The process of changing the parameters and improving the accuracy of the predictions is expected, as results are immediately available on the correctness of the predictions.

The scope of this study is as follows.

1. The proposed approach predicts the hourly heat source load for the next day in the existing building.

2. The target of forecasting is limited to cooling loads, where more energy savings are expected, and does not deal with heating loads.

3. If the predicted cooling load is spatially subdivided, the approach is more valuable for air-conditioning operations, although, load performance monitoring points must also be taken at the same granularity. The detailed measurement has become realistic in recent years due to advances in IT technology, however, this research aims the cooling load forecast for the existing building, considering the suppression of renovation work costs.

2. Air-Conditioning Load

With reference to the dynamic load calculation method based on the response factor method [2], [3], the air-conditioning loads were classified into two categories: those estimated by mathematical calculations and those estimated by machine learning, assuming that each element has its own explanatory variables. The explanatory variables for each load element were selected on the basis of whether they could be measured in a typical building.

In this section, the term "current" indicates the point in time at which the forecast is made.

2.1. *Load Factors Subject to Machine Learning*

1. Load of exterior wall through-flow heat due to temperature difference between interior and exterior including through-flow load from glass. The explanatory variables for this load are: current outdoor air temperature, past outdoor air temperature (going back about 50 hours), past daily average outdoor air temperature (going back about 3 months), current horizontal solar radiation, and past horizontal solar radiation (going back about 50 hours).

2. Load of interior wall penetration heat from adjacent room. The explanatory variables for this load are the current outdoor air temperature and the past outdoor air temperature (going back about 50 hours).

3. Load due to solar radiation transmitted through glass. The explanatory variables for this load are current daylight hours, horizontal solar radiation, and historical horizontal solar radiation (going back about 10 hours).

4. Load due to drafts. The explanatory variables for this load are current outdoor air temperature, current outdoor air humidity, current external wind direction and speed, and building opening condition.

5. Load due to heat storage of intermittent air conditioning operation. The explanatory variable

for this load is the past room temperature of the room concerned (going back about 20 hours).

6. System losses, .e.g. fan reheat, duct losses and mixing losses.

2.2. *Load Factors Subject to Actuarial Calculations*

7. Load due to human metabolism. The explanatory variables for this load are current manpower, past manpower (going back about 10 hours), and average work intensity.

8. Load due to heat dissipation of lighting fixtures. The explanatory variables for this load are current power consumption and past power consumption (going back about 10 hours).

9. Load due to heat dissipation from other indoor equipment. The explanatory variables for this load are current heat generation, and past heat generation (going back about 10 hours).

10. Outdoor air load. The explanatory variables for this load are outdoor air intake, outdoor air temperature, and outdoor air enthalpy.

3. Prediction Method for Cooling Load

A prediction method which combines mathematical computation and machine learning approach is build.

3.1. *Learning Phase*

When training with measurement data, the data set is composed for NN [3] as follows.

- Teaching data are the differences between the overall loads and the loads of Section 2.2.
- Input data are the explanatory variables for loads of Section 2.1.
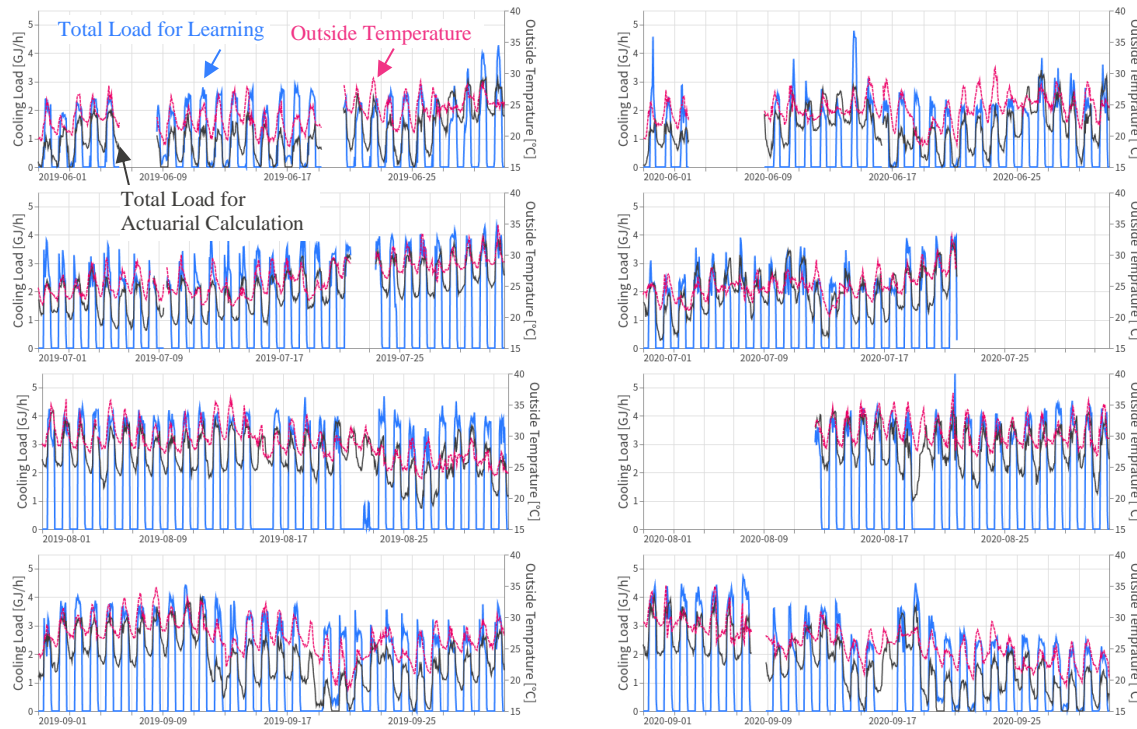
Fig. 1 Monthly Cooling Load of the Example Instance in 2018 and 2019.

4. Computational Experiment

4.1. *Example*

An example of air conditioning load forecasting is consists of two years of measured data for the overall air conditioning load and the explanatory variables for each air conditioning load for the underground shopping mall in Kobe, see Table 1. Fig. 1 shows the cooling loads of 2 years.

Table 1. Outline of Underground Mall

Name	Santica
Year of construction	1963 (1st phase)
	1966 (2nd phase)
Management Entity	Kobe Chikagai Co., Lt
Total floor area	19,109m2
Retail area	10,145m2
Business Hours	10:00-20:00 (merchandising)
	11:00-21:00 (food and
	beverage)
	6:00 - 24:00 (passageway)

4.2. *Learning setting*

Since the subject is an underground mall, the effect of solar radiation was ignored among the air conditioning load elements. See [4] for the measurement method of the current people flow. The settings of the NN method are shown in Table 2.

Table 2. Setting of NN method.

Input Layer	Input Size	28
Intermediate Layer	Number of Layers	2
	Type	Fully-Connected Layer
	Activation Function	ReLU
	Number of Units	20 (first layer)
		10 (second layer)
Output Layer	Type	Fully-Connected Layer
	Output Size	1
Others	Batch Size	10
	Epoch Size	30
	Trial	10

Makoto Ohara, Hideo Isozaki

4.3. *Prediction setting*

In this manuscript, calculations are performed on historical data sets to verify the performance of the NN method. Where weather forecasts or other forecasting data should be used in actual use, measurement values are used. The people-flow prediction method is described in [5].

4.4. *Result*

Fig. 2 shows the difference between normalized predicted and actual measurements by time period. The errors are particularly large between 8:00 and 10:00 and at 20:00. The prediction accuracy is expected to improve when the conditions at the start and end of cooling are considered.
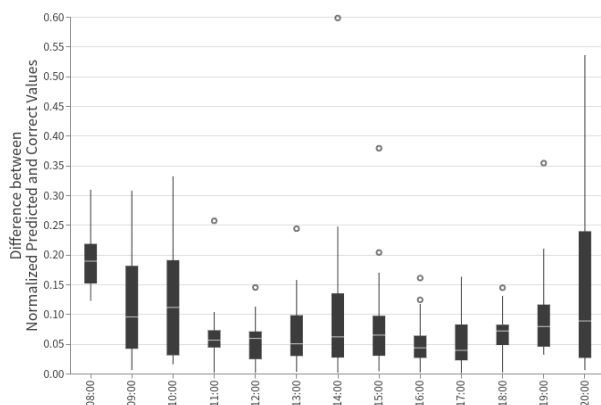
Fig. 2 Result of Prediction.

5. Conclusion

In this study, a method for predicting air-conditioning loads is investigated with the goal of efficient air-conditioning operation in underground malls with open areas. The proposed method has been applied the example based on actual data from the underground mall in Kobe. Future issues include evaluation of predictions based on actual use and examination of accuracy in comparison with existing studies.

References

1. Y. Matsuo, K. Yokoyama, H. Ishino, S. Kawamoto, "Introduction to Dynamic Heat Load Calculations for Air Conditioning Systems", Japanese Association of Building Mechanical and Electrical Engineers, 1980. (In Japanese)
2. H. Oda, M. Nakao, S. Todo, H. Isozaki, "Study on an Identification Model of Cooling Coil Characteristics of Air Conditioners (Report 1) Setting Up a Gray-Box Model and Examining the Identification Accuracy", Proceedings of Annual Conference of the Society of Heating, Air-Conditioning and Sanitary Engineers of Japan, E-6, 2020. (In Japanese)
3. T. Okatani, "Deep Learning", Kodan-sha, 2015. (In Japanese)
4. M. Ohara, T. Matsumoto, Y. Suzuki, H. Nagahiro, H. Tamaki, "An Estimation Method for People Flow with Laser Rangefinders on an Underground Walkway", Proceedings of Annual Conference of the Institute of Systems, Control and Information Engineers, 2018. (In Japanese)
5. M. Ohara, T. Matsumoto, K. Sakakibara, Y. Suzuki, H. Nagahiro, H. Tamaki, "Pedestrian Density Prediction Method for Air Conditioning System on Underground Mall with Exterior Openings", Proceedings of SICE SSI, 2020. (In Japanese)

Authors Introduction

Dr. Makoto Ohara

He received Doctor of Philosophy in Engineering from Kobe University in 2012. He works as an assistant professor in International Professional University of Technology in Osaka now. His research interests optimization for social systems. In recent years he has also research into machine learning approaches.

Mr. Hideo Isozaki

He received his Master in Engineering from Waseda University in Japan. He is a visiting professor Carbon Neutral Promotion Headquarters, Kobe University in Japan.

Intelligent Logistics Handling Robot: Design, Control, and Recognition

Yanchao Bi

Tianjin University of Science and Technology, 1038 Dagu South Road, Hexi District, Tianjin, 300222, China

Jiale Cheng

Tianjin University of Science and Technology, 1038 Dagu South Road, Hexi District, Tianjin, 300222, China

Limei Wang

Tianjin University of Science and Technology, 1038 Dagu South Road, Hexi District, Tianjin, 300222, China

Yizhun Peng

Tianjin University of Science and Technology, 1038 Dagu South Road, Hexi District, Tianjin, 300222, China

Email: pengyizhun@126.com

Abstract

This study aims to investigate various aspects of intelligent logistics handling robots, including mechanical design, automatic control, and image recognition. With the continuous development of the logistics industry and advancements in automation technology, intelligent logistics handling robots play a crucial role in improving logistics efficiency and reducing costs. Leveraging existing technologies, we have designed and developed an omnidirectional mobile intelligent logistics handling robot equipped with a SCARA-type robotic arm. The robot integrates functions such as task acquisition, global positioning, material detection, warehouse identification, material handling, and stacking, achieving a fully automated and streamlined logistics handling process.

Keywords: Logistics handling robots, Automatic control, Image recognition, Global positioning

1. Introduction

In modern industrial production, logistics handling is a time-consuming and labor-intensive task. The modern industry faces rapidly changing market demands and a complex and dynamic production environment, making traditional manual logistics handling methods relatively outdated. In this situation, there is an urgent need for a more flexible, efficient, and adaptive logistics handling approach to meet the requirements of modern industry. This new type of handling method should be able to adapt to changes in different factory environments and production processes, completing a large number of handling tasks in a short time to improve production efficiency. This is precisely the significant advantage that intelligent logistics handling robots can bring to modern industry.

The intelligent logistics handling robot studied in this project comprehensively designs and analyzes aspects such as mechanical structure, robotic arm control and coordinate calculation, chassis control and global positioning, image recognition, and processing. A multifunctional intelligent logistics handling robot (AGV) is designed based on the omnidirectional wheel structure chassis and planar articulated (SCARA) robotic arm [1], [2], [3], [5]. It integrates functions such as positioning, environmental perception, path planning, material recognition and grasping, material placement, and stacking. This integration significantly reduces the number of operators, enhances the speed and accuracy of logistics operations, and greatly improves the efficiency of the logistics system.

2. Intelligent Logistics Handling Robot Overall Program Design
2.1. Operational site

In this thesis, the robot's operation site is based on the 9th China University Students Engineering Practice and Innovation Competition "Intelligent +" track logistics handling robot competition. The robot drives on the gray lane, and the rest of the area is matte white or yellow. In the competition venue, the start-stop zone, raw material zone, roughing zone and temporary storage zone are set up. The start-stop zone is blue in color and is used for the robot to go back and forth. Fig. 1 shows the exact location of the given raw material area, roughing area and staging area within the competition venue.

Yanchao Bi, Jiale Cheng, Limei Wang, Yizhun Peng
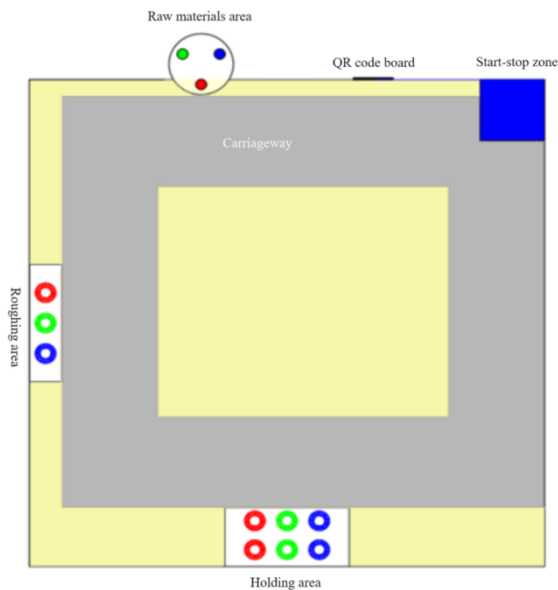
Fig. 1 Schematic diagram of the robotics field

The raw material area adopts a round electric turntable to arrange materials, and materials are identified by color. The top surface of roughing area, staging area, finishing area and finished product area are equipped with color rings and or circles for measuring the accuracy of material placement.

2.2. Mandate requirements

The robot follows a "one-key" start-up procedure, moving from the start-stop zone to the QR code panel within a specified time to read the QR code. It obtains two handling tasks (the sequence for handling materials of three different colors). According to the task requirements, the robot moves between the material area, rough processing area, and temporary storage area, carrying out placement and stacking tasks. After completing the tasks, it returns to the start-stop zone.

2.3. *Design proposal*

According to the task requirements, the overall design of the intelligent logistics handling robot is illustrated in Fig. 2. In terms of mechanical structure, the chassis adopts a Mecanum wheel [1], [5] double-layer shock-absorbing design, controlled by DJI brushless motors to achieve closed-loop control. The robotic arm employs a planar articulated (SCARA) design, driven by stepper motors and closed-loop stepper motors, allowing real-time reading of the angles of each joint for feedback control.

The electronic control system is mainly divided into three parts: chassis motion control system, robotic arm motion control system, and visual recognition system. The chassis motion control system, centered around the

chassis control board, includes external devices such as a QR code recognition module, gyroscope, touchscreen, and brushless motor driver. The robotic arm motion control system integrates servos and utilizes closed-loop stepper motors for precise control of the robotic arm. The visual module, developed based on OpenCV, includes functions for material color recognition and target ring center recognition. The chassis control board, robotic arm control board, and visual module are all powered simultaneously by the power module.
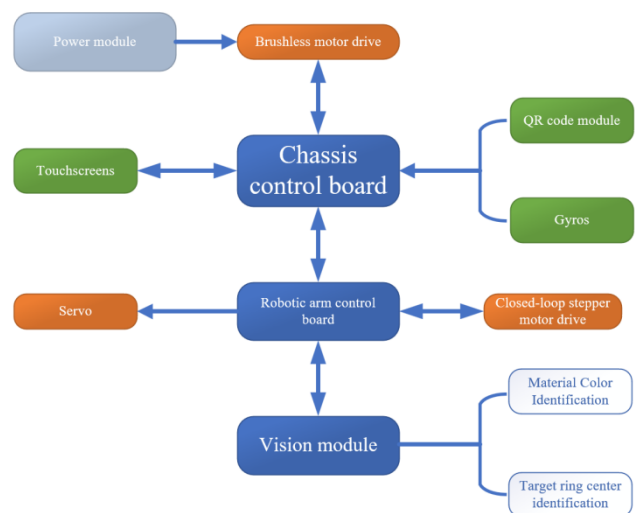
Fig. 2 Overall design schematic

3. Design of Mechanical Structural Systems

3.1. *Intelligent logistics handling robot overall mechanical structure design*

To ensure the flexibility and efficiency of the robot, the chassis adopts a Mecanum wheel chassis, and the mechanical arm structure adopts a standard three-axis SCARA structure.

a. Robot 3D drawing
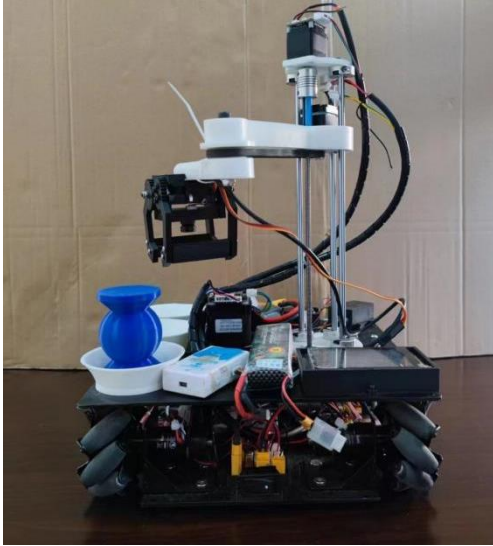
b.   Robot Physical Picture
Fig. 3 Image of Intelligent Logistics Handling Robot

3.2. *Chassis mechanical structure design*

The chassis adopts a double-layer design with Mecanum wheels, made of aluminum alloy with electroplating process. The lower layer of the chassis houses the electronic control system, while the upper layer accommodates the mechanical arm, material storage, and human-machine interaction module, among others. The overall structure is well-defined.

The motion system of the chassis requires a high level of grounding for the Mecanum wheels [5]. To ensure stable movement of the robot, the lower layer of the chassis adopts a double-plate coaxial structure, ensuring that all four wheels touch the ground simultaneously. This design increases shock absorption and solves the problem of Mecanum wheels slipping when not in contact with the ground [1].

a.   Top view

b.   Bottom view
Fig. 4 Chassis structure

3.3. *Mechanical structure design of mechanical arm*

This robotic arm is of the three-axis SCARA type [2], made of PLA material and manufactured using 3D printing technology. It is lightweight in quality.
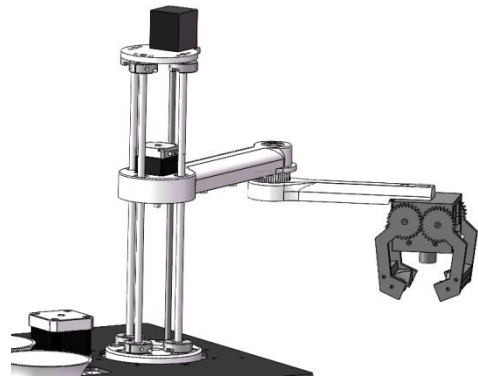
Fig. 5 3D model of the robot arm

The upper arm of the robotic arm is connected to the arm axis through a linear bearing, while the lower arm is equipped with a thrust ball bearing to withstand axial forces and reduce friction between the upper and lower arms. The robotic arm can achieve three degrees of freedom through the motion of stepper motors, including vertical movement, rotation of the upper arm, and rotation of the lower arm.

3.4. *Mechanical Structure design of mechanical claw*

The mechanical claw is one of the key components of an intelligent logistics handling system. It is manufactured using 3D printing technology and made of PLA material.

Yanchao Bi, Jiale Cheng, Limei Wang, Yizhun Peng
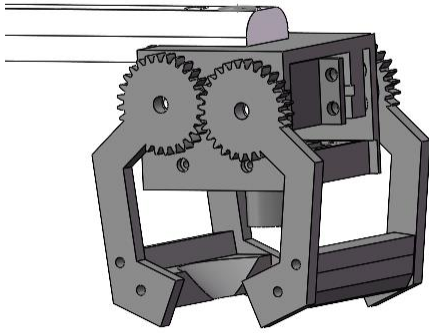
a.   3D drawing
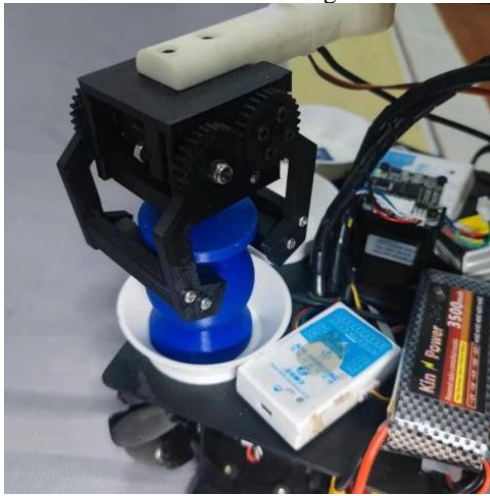
b.   Concrete drawing
Fig. 6 Robotic Gripper Structure

The mechanical claw is used for grasping and identifying materials. When grasping materials, the mechanical claw needs to firmly grip the materials. The mechanical claw is divided into active and passive parts, with the active part being directly driven by a servo motor to rotate, and the active part drives the passive part to rotate through gears. The identification of materials involves the use of a camera and a visual recognition system. The camera is installed in the center of the mechanical claw.

4. Three-axis SCARA Robot Kinematic Model

4.1. *Kinematic modeling of a three-axis SCARA robotic arm*

SCARA robotic arms are usually characterized by high accuracy, high speed and repeatability [2], [3]. In this project, the SCARA robotic arm is designed as a three-degree-of-freedom, containing one rotational degree of freedom in the XY plane, one lifting degree of freedom in the Z-axis, and one moving degree of freedom in the

Z-axis. The process of solving the operational science model of the three-axis SCARA robotic arm is as follows:

Determination of joints, connecting rods and their parameters. The robot arm uses a rotating chassis at the bottom, lifting and lowering using a stepper motor screw Z-axis, and a rotating arm in the XY plane, for a total of three degrees of freedom, and the analysis process is simplified by dropping the end-effector [3].

Simplify the model to set the coordinate system. With the vertical upward as the +Z-axis and the center of the Z-axis as the coordinate origin from the top view, establish a plane right-angle coordinate system with the body as the reference system.
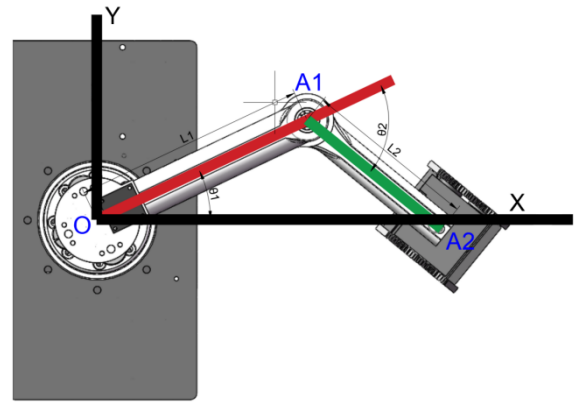
Fig. 7 Simplified coordinate system model

As shown in Fig. 7, for the big arm of the robot arm, for the small arm of the robot arm. For the rotation angle of the big arm, controlled by the bottom disk; for the rotation angle of the small arm, controlled by the stepping motor through the synchronous belt.

Inverse kinematic modeling calculations. We set the point coordinates to (x, y, z), the length of the big arm to be L1, and the length of the small arm to be L2, then the length $L3 = \sqrt{L1^2 + L2^2}$. The angle of rotation of the large arm θ_1 and the angle of rotation of the small arm θ_2 need to be solved for. This leads to Eq. (1). The rotation angle of the big arm θ_1 and the rotation angle of the small arm θ_2 can be found (Eq. (2)).

$$\begin{cases} \cos(\theta_1) = \dfrac{L2^2 + L3^2 - L1^2}{2 * L2 * L3} \\ \cos(\theta_2) = -\dfrac{L1^2 + L2^2 - L3^2}{2 * L1 * L2} \end{cases} \quad (1)$$

$$\begin{cases} \theta_1 = arccos(\dfrac{L2^2 + L3^2 - L1^2}{2 * L2 * L3}) \\ \theta_2 = arccos(-\dfrac{L1^2 + L2^2 - L3^2}{2 * L1 * L2}) \end{cases} \quad (2)$$

4.2. *Robotic arm electronic control system design*

(1) Robotic arm control board circuit design

The robotic arm control board uses STM32F405RGT6 as the main control chip [4], with an externally integrated power module, stepper motor drive module and communication module. The power supply module part, the input voltage is 12V, using TPS5430 decompression to generate 5V voltage, and then through the AMS1117 on the 5V voltage regulator to generate 3.3V, thus generating all the voltages required for system operation. The control board integrates the TMC2209 stepper motor driver, which is powered by 12V and has the advantages of silence, high speed and low failure, and is used for open-loop control of the stepper motor. The closed-loop stepper motor driver is controlled using a UART interface. In order to realize a UART interface and multiple closed-loop stepper motor driver communication problems, we use a diode plus pull resistance way to enhance the serial port receiving ability, shielding the interference signal.

(2) PCB Overall Layout

The control board of the robotic arm is a four-layer board, and the shielding layer is set up to shield the power supply and signals, which improves the signal stability of the system. As Fig. 8 are the front and back of the chassis control board.
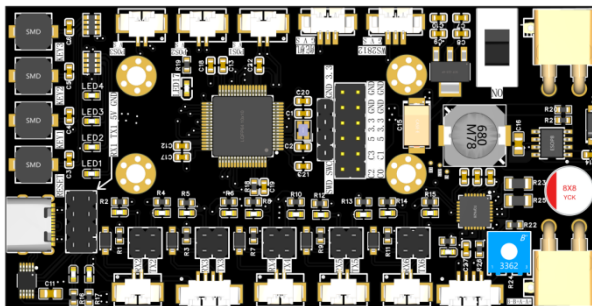
a.   Top view
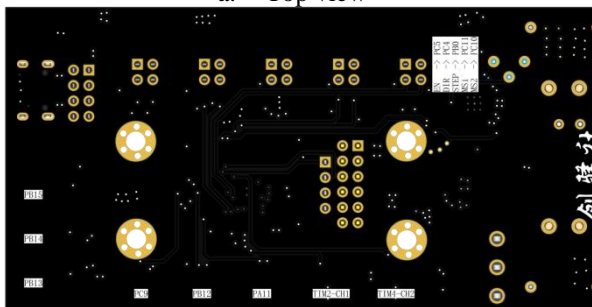
b.   Bottom view

Fig. 8 Robotic arm control board

(3) Robotic arm software design

The robotic arm performs numerous actions, and the use of sequential execution of robotic arm actions will occupy a large amount of program storage space, and the code logic is confusing and redundant. This system uses structure variables to save the spatial coordinates of each step of the robotic arm action and information such as gripping operation, and establishes structure arrays for storing the information of the robotic arm action during the whole operation process. After the robotic arm is powered on, it initializes the peripherals and stepper motors, servos and so on. After the initialization is completed, turn on the serial port to receive interrupts and wait for the chassis controller to send task execution instructions. When the task execution instruction is received, the robotic arm starts to move, reads the action information in the structure array, and executes it step by step. The vision system will send the material color, target ring coordinates and other information to the robotic arm, and the robotic arm will calibrate and correct the position through this information. When all the actions are completed, the robotic arm task is completed and returns to the initial state. The workflow of the robotic arm motion control system is shown in Fig. 9.
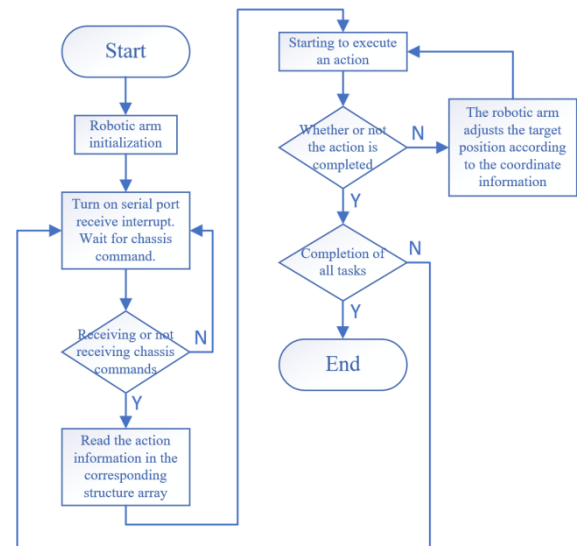
Fig. 9 Flow chart of robotic arm motion control system

5. Chassis Motion Control System Design

5.1. *Mecanum wheel chassis kinematics model*

The Mecanum wheel platform kinematic model is to establish the relationship between the rotational speed of the four Mecanum wheels and the speed of the geometric center [4]. The positive kinematics model is to know the rotational speed of the four Mecanum wheels and calculate the speed of the Mecanum wheel platform center point CENTER; the inverse kinematics model is to know the speed of the Mecanum wheel platform center point CENTER and calculate the speed of the four Mecanum wheels.

As an example, the inverse kinematics model of the Mecanum wheel is used to establish a coordinate

system with the geometric center point CENTER of the top view of the chassis as the origin [5]. The distance between the CENTER point and the point of contact between the Mecanum wheel and the ground is r. r_x and r_y denote the projected distances on the x-axis and y-axis, respectively, of the CENTER coordinate system (both positive). The velocity of the CENTER point is known to be $[v_c \ w_c]^T$. v_{cx} and v_{cy} are the components of a on the x and y axes, respectively.
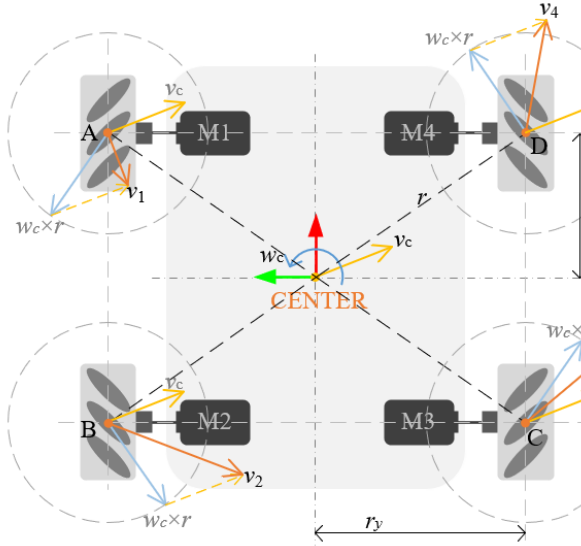
Fig.10 Decomposition of Mecanum Wheel Platform Motion

After calculation, the inverse kinematic equation of Mecanum's wheel can be obtained.

$$
\begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix} = \frac{1}{R} \begin{bmatrix} 1 & -1 & -(r_x + r_y) \\ 1 & +1 & -(r_x + r_y) \\ 1 & -1 & +(r_x + r_y) \\ 1 & +1 & +(r_x + r_y) \end{bmatrix} \begin{bmatrix} v_{cx} \\ v_{cy} \\ w_c \end{bmatrix} \quad (3)
$$

The positive kinematics equations are based on the rotational speeds of the four Mecanum wheels to calculate the velocity of the point CENTER, and the theoretical analysis is similar to the inverse kinematics, and the simplest way to do this is to invert Eq. (3), which can be expressed as Eq. (4).

$$
\begin{bmatrix} v_{cx} \\ v_{cy} \\ w_c \end{bmatrix} = \frac{R}{4} \begin{bmatrix} +1 & +1 & +1 & +1 \\ -1 & +1 & -1 & +1 \\ -1 & -1 & 1 & 1 \\ r_x + r_y & r_x + r_y & r_x + r_y & r_x + r_y \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix} \quad (4)
$$

5.2. Full-field positioning method

The full-field localization method studied in this project does not require the addition of the rest of the extra sensors or encoders and uses the robot's own

Mecanum wheel as the encoding wheel. When the robot starts to move, the Mecanum wheel starts to shift rotation. According to the positive kinematics model of the Mecanum wheel, the velocity of the robot at the current moment can be obtained from the rotational speeds of the four wheels with respect to the current coordinate system of the robot itself [5]. Assuming that the angle between the current robot's own coordinate system and the world coordinate system is θ (Fig. 11).
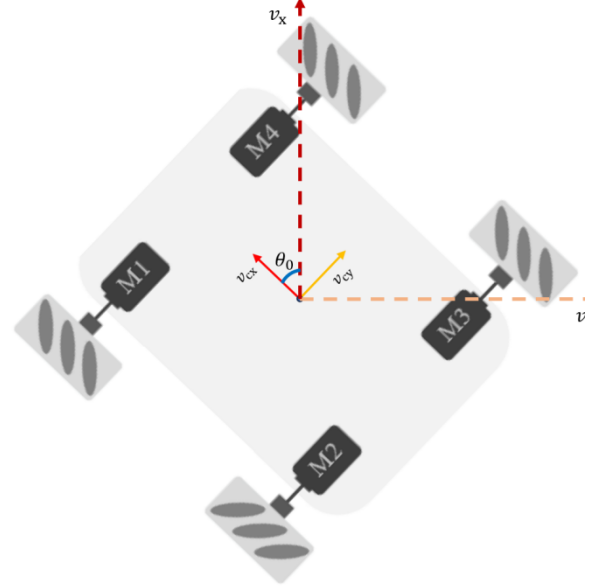
Fig.11 World Coordinate System and Own Coordinate System

The expression for converting the robot's velocity in its own coordinate system to the velocity in the world coordinate system is Eq. (5).

$$
\begin{bmatrix} v_x \\ v_y \\ w_z \end{bmatrix} = \begin{bmatrix} cos\theta_0 v_{cx} + sin\theta_0 v_{cy} \\ -sin\theta_0 v_{cx} + cos\theta_0 v_{cy} \\ w_c \end{bmatrix} \quad (5)
$$

With the same time units, the conversion to a displacement expression is Eq. (6).

$$
\begin{bmatrix} x_x \\ x_y \\ \theta_z \end{bmatrix} = \begin{bmatrix} cos\theta_0 x_{cx} + sin\theta_0 x_{cy} \\ -sin\theta_0 x_{cx} + cos\theta_0 x_{cy} \\ \theta_c \end{bmatrix} \quad (6)
$$

In this case, the calculated displacement is the relative displacement in the world coordinate system for an instantaneous moment. Since the angle between the robot's own coordinate system and the world coordinate system is always changing, therefore, to obtain the displacement of the robot in the world coordinate system over a period of time, it is necessary to integrate the displacement value calculated at any moment (discrete data summation) [6], assuming that over a period of time, a total of N times the robot's displacement relative to the

world coordinate system at a given moment, the cumulative displacement expression is Eq. (7).

$$\begin{bmatrix} \Delta x_x \\ \Delta x_y \\ \Delta \theta_z \end{bmatrix} = \begin{bmatrix} \sum_{i=0}^{N} cos\theta_i x_{cxi} + sin\theta_i x_{cyi} \\ \sum_{i=0}^{N} -sin\theta_i x_{cxi} + cos\theta_i x_{cyi} \\ \Delta \theta_c \end{bmatrix} \quad (7)$$

In Eq. (7), i represents the number of operations.

5.3. *Chassis electronic control system design*

(1) Chassis control board circuit design

The main control and power scheme of the chassis control board is the same as that of the robotic arm control board, in addition to the integrated communication module. The chassis motion control system adopts a variety of communication interfaces to communicate with various peripherals, mainly including communication with gyroscope, QR code scanning module through UART serial port, communication with touch screen through SPI and IIC, and communication with brushless motor driver through CAN [7].

(2) PCB Overall Layout

The PCB is fabricated using a two-layer board design with careful planning and design of the overall layout. As Fig. 12 show the front and back view of the chassis control board.
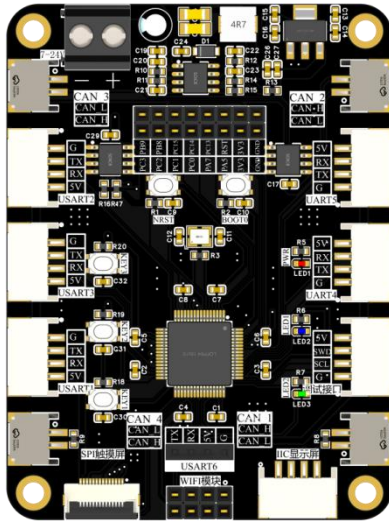
a.   Top view

b.   Bottom view

Fig.12 Chassis control board

5.4. *Chassis software design*

Prepare a chassis motion control system flowchart according to the task requirements. After the system is powered on, the program starts running and the peripherals are initialized. After the initialization is completed, the robot drives out of the start-stop area, runs to the QR code board, and starts the QR code scanning module to start scanning the QR code to receive the task code. The robot enters the next running task and proceeds to the raw material zone. When the robot runs to the raw material area, the vision system is activated to identify the color of the material according to the task code, and the robotic arm starts to grasp and place it into the bin on the robot. After completing the grasping of all materials, the robot continues to move forward. When the robot travels to the vicinity of the roughing area, the vision system starts to recognize the coordinates on the target ring in the roughing area. The chassis control board adjusts the robot position according to the coordinate data so that the robot is accurately close to the rough machining area. At this time, the vision system starts to identify the center of the target ring, and the robotic arm will place the corresponding color materials in the target ring according to the position of the target ring. After all materials have been placed, the robot will pick up the materials again in the order in which they were placed and place them in the bin on the robot. When the above operation is complete, the robot will move on to the staging area and resume the material placement operation. This operation is the same as the robot placing material in the roughing area. When all tasks are completed, the robot returns to the start-stop area.
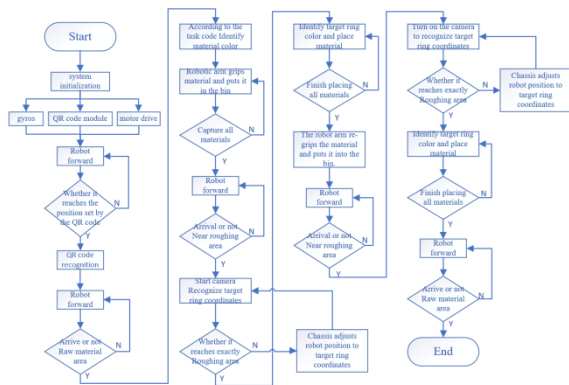
Fig.13 Chassis Motion Control System Flowchart

6. Visual identity system

6.1. *Material Color Identification*

The task requires the robot to be able to recognize the color of the material and grab the material in the order of the task code. This project uses OpenCV library for color recognition and detects the center of gravity of the material in the image. The idea of specific color recognition is to confirm the HSV color space threshold of the color and binarize the image with this HSV threshold to get a binary image [8]. At this point, the image is expanded so that the outline of the material in the image tends to be more like a circle. Based on this outline, it can be determined to circle the material and determine the coordinates of the center of gravity of the material.

As in Fig. 14, the robot gripper is directly above the blue material, at this time, the visual recognition system is turned on to identify the material color. The outline of the material will be wrapped by the black outline line, and the center of the material will be marked (center).

a. Identification process

b. Identification results

Fig.14 Visual identification of material colors

6.2. *Target ring center identification*

The task requires the robot to be able to recognize the target ring coordinates and place the material accurately into the target ring coordinates. This project uses the OpenCV library to identify the center of the target ring, which is a circle with color. When dealing with the target ring detection problem, this system does not directly detect the target ring directly by means of Hough's circle transformation, but first binarizes the image. After that, the contours in this image are searched and the contours are filtered to find the contour of the target ring. The contour of the target ring is drawn as the smallest enclosing circle, which is very close to the position and size of the target ring itself, and the circle is the center of the target ring.

As shown in Fig. 15, the robot gripper is directly above the blue target ring, and at this time, the visual recognition system is turned on to detect the center of the target ring. The center of the blue target ring will be marked by a blue dot.
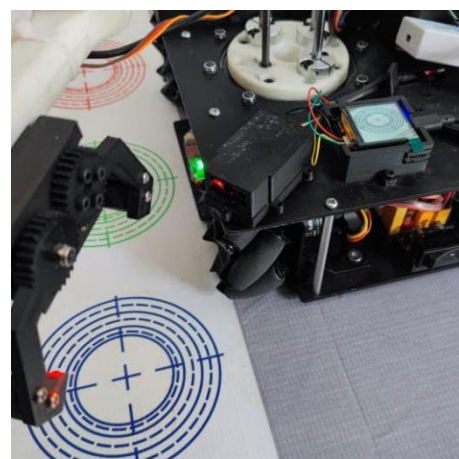
a. Identification process
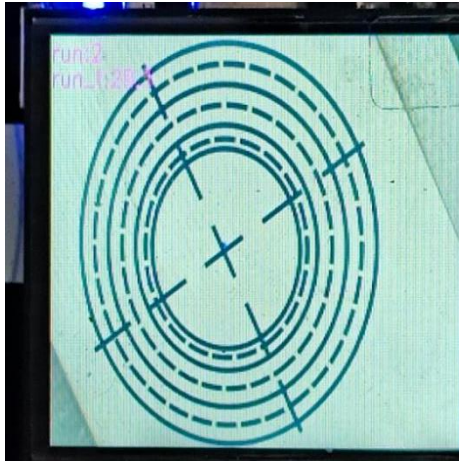
b.   Identification results
Fig.15 Visual recognition target ring center

7. Conclusion

This project designs an intelligent logistics handling robot (AGV) for industrial logistics scenarios, which is equipped with a chassis motion control system based on an omnidirectional moving chassis with Mecanum wheels, a robotic arm motion control system based on a planar articulated (SCARA) robotic arm, and a material recognition system based on machine vision. The chassis motion control system is used to control the robot motion, pick up the task code, and perform full-field localization. The robotic arm motion control system realizes material gripping and placing operations through the movement of the robotic arm. The visual recognition system then identifies the color of the material and the center of the target ring. The chassis motion control system sends commands to the chassis motion control system to control the robotic arm to perform tasks. When the robotic arm performs the task, it needs to receive the data returned by the visual recognition system to correct the coordinate position. Each system cooperates with each other to accomplish the task.

The design simulates the handling and palletizing tasks under industrial operations, which will greatly improve the rapidity and accuracy of logistics operations and greatly improve the efficiency of the logistics system.

References

1. Huimin Xu, Gaohong Yu, Yimiao Wang, Xiong Zhao, Yijin Chen Jiangang Liu. (2023). Path Planning of Mecanum Wheel Chassis Based on Improved A* Algorithm.Electronics(8).
2. Zhao Nan Ibaraki Soichi. (2022). Novel kinematic model of a SCARA-type robot with bi-directional angular positioning deviation of rotary axes.The International Journal of Advanced Manufacturing Technology(7-8), 4901-4915.
3. Rigatos G., Abbaszadeh M., Busawon K. Pomares J..(2023).Nonlinear optimal control for a 4-DOF SCARA robotic manipulator. Robotica(8), 2397-2450.
4. Long Tao. (2023).   Design of Sweeping Robot Based on STM32 Single Chip Microcomputer.Journal of Physics: Conference Series(1).
5. Dong Nguyen Minh, Hiep Do Quang, Nam Dao Phuong, Tien Ngo Manh Duy Nam Bui. (2023). An Adaptive Fuzzy Dynamic Surface Control Tracking Algorithm for Mecanum Wheeled Mobile Robot.International Journal of Mechanical Engineering and Robotics Research(6).
6. Pizá Ricardo, Carbonell Rafael, Casanova Vicente, Cuenca Ángel Salt Llobregat Julián J..(2022).Nonuniform Dual-Rate Extended Kalman-Filter-Based Sensor Fusion for Path-Following Control of a Holonomic Mobile Robot with Four Mecanum Wheels.Applied Sciences(7),
7. Jing Yan, Xuyang Tian, Xiaoyuan Luo Xinping Guan. (2019).Design of an Embedded Communication System for Underwater Asynchronous Localization..Embedded Systems Letters(3),97-100.
8. CorbalanFuertes Montserrat, Millan Maria S. Gilabert Daniel. (2001). Computer vision system for color discrimination.Univ. Politecnica de Catalunya (Spain)592-595.

Authors Introduction

Mr. Yanchao Bi

He was admitted to Tianjin University of Science and Technology, China in 2021, majoring in Intelligent Science and Advanced Manufacturing, School of Electronic Information and Automation. He is currently an undergraduate student of Tianjin University of Science and Technology, China.

Mr. Jiale Cheng

He was admitted to Tianjin University of Science and Technology, China in 2021, majoring in Intelligent Science and Advanced Manufacturing, School of Electronic Information and Automation. He is currently an undergraduate student of Tianjin University of Science and Technology, China.

Ms. Limei Wang

She was admitted to Tianjin University of Science and Technology, China in 2021, majoring in Intelligent Science and Advanced Manufacturing in the School of Electronic Information and Automation. He is currently an undergraduate student of Tianjin University of Science and Technology, China.

Dr. Yizhun Peng

He is an Associate Professor in Tianjin University of Science &Technology. He received a doctor's degree in control theory and control engineering from the Institute of Automation, Chinese Academy of Science in 2006.His research field is intelligent robot and intelligent control.

Greenhouse Design Using Visual Recognition and IoT Technology

Yuntian Xia,

College of Artificial Intelligence ,Tianjin University of Science and Technology, 300457，China；
Yizhun Peng*

College of Electronic Information and Automation, Tianjin University of Science and Technology,
300222，China
E-mail: *pengyizhun@tust.edu.cn
www.tust.edu.cn

Abstract

This device is to solve the traditional pesticide spraying method on the human body has a greater impact and other issues, through the STM32-based visualisation of the intelligent greenhouse to achieve automatic spraying of pesticides and remote monitoring and other functions, the establishment of a visualisation of the intelligent greenhouse monitoring platform. This equipment through the MQTT protocol, not only through the Internet of Things platform real-time monitoring of crop growth status in the greenhouse, but also through the platform to determine whether to spray pesticides, data transmission, so as to use the cross slide to control the position of the nozzle, and then through the visual recognition algorithms to improve the accuracy of the visual recognition part of the accuracy of the spraying of plants affected by insect pests, the realization of the digital intelligent greenhouse.

Keywords: STM32, Monitoring Platform, MQTT, Visual Recognition

1. Introduction

This project is dedicated to solving problems such as the potential impact of traditional pesticide spraying methods on the human body. Through the STM32-based [1] visualised intelligent greenhouse, automatic pesticide spraying and remote monitoring functions are realised, and a complete visualised intelligent greenhouse monitoring platform is constructed. We adopted the MQTT [2] protocol, which not only monitors the growth status of crops in the greenhouse in real time, but also replenishes different colours of light in a targeted manner according to the light conditions inside the greenhouse to improve the efficiency of photosynthesis. This is not only more energy efficient than traditional white light supplementation, but also significantly increases crop yields and helps to make the system carbon neutral.

In addition, our system also features indoor temperature monitoring and control, which promotes the accumulation of organic matter in crops by adjusting the temperature difference between day and night. For a more precise response to pests, we utilise a cross slide to control the position of the nozzles and incorporate visual recognition algorithms to improve the accuracy of spraying on pest-affected plants [3]. For a sustainable energy supply, we added solar panels to the exterior of the greenhouse to provide a constant source of energy for the system, while improving long-term human health,

reducing air pollution and greenhouse gas emissions, and positively impacting human health [4].

For the electrical control part of the project, we carried out systematic modelling and simulation through the Proteus simulation platform to determine the optimal operating conditions. Subsequently, we used Altium Designer to create the design, programming and debugging, and finally succeeded in building an efficient intelligent greenhouse monitoring platform [5].

The rest of this paper is organised as follows. The second part describes the hardware used and the theoretical calculations. The third part describes the method of controlling the nozzle and the visual recognition algorithm. Part IV describes how the smart shed works and shows some of the operational results. Part V summarises the main points of the paper.

2. Design of Hardware Structure

2.1. Selection of the main control board

This work needs to deal with a large amount of data, has a more complex circuit, and needs a lot of memory, so it needs a microcontroller that can deal with a lot of song peripheral devices at the same time, so the STM32F4 became the team's first choice. the physical diagram of the STM32F4 is shown in Fig. 1.

Fig. 1 STM32F4 Physical

2.2. Relay Modules

This work is controlled by mobile phone APP, through the Internet of Things platform, the corresponding signal will be sent to the STM32 core board, and it will control each module. In order to save power, they will use a relay to control the switch of some modules, so that some non-working modules are in the off state when they are not working, this work uses the JQC-3FF-S-Z relay, the physical picture is shown in Fig. 2.

Fig. 2 Relay Physical Diagram

2.3. Driver Modules

This work needs to be as precise as possible when it comes to pesticide spraying so that the pesticide can be sprayed where we need it, so we chose to use stepper motors and stepper motor drives. The physical drawing of the stepper motor and stepper motor drive is shown in Fig. 3.
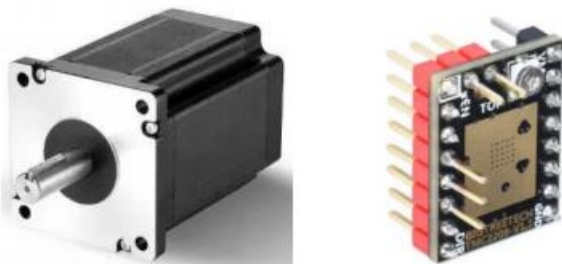
Fig. 3 Physical drawings of stepper motors and stepper motor drives

2.4. Mechanical components

The mechanical part is designed as shown in Fig. 7, its motion system consists of two axes, X-axis and Y-axis, to control the position of the nozzle and the pump.

The nozzle structure is to control different spraying methods and the pump structure is used to spray different concentrations of pesticides. The mechanical structure is modelled as shown in Fig. 4.

Fig. 4 Mechanical structure modelling drawings

The main issues considered in the design: Since the crops inside the greenhouse are scattered all over the place, a mechanical structure that can bring the nozzles and pumps to every part of the greenhouse is needed, they refer to the mechanical structure that controls the movement of the nozzles in the 3D printers, and they designed and chose the above solution.

2.5. Sensor selection

In order to obtain more information about the greenhouse, this work chooses a variety of sensors. The temperature and humidity sensor can always sense the temperature and humidity inside the greenhouse, which is convenient for the adjustment of the internal temperature of the greenhouse [6]; the light intensity sensor monitors the change of the light intensity inside the greenhouse, so as to judge what colour and intensity of light needs to be reinforced; the visual sensor can intuitively observe the growth status of the plants inside the greenhouse and transmit the information to the user's mobile phone APP, so as to achieve real-time monitoring. The digital temperature and humidity sensor, light intensity sensor and OpenMV are shown in Fig. 5.

Fig. 5 The digital temperature and humidity sensor, light intensity sensor and OpenMV

2.6. Motion control of nozzles and pumps

Motion control of spray nozzles and pumps means that the mobile phone APP or various sensors send signals to the main board, so as to control the spray nozzles and pumps to move according to the predetermined paths and speeds in the two-dimensional space through the cross-axis structure. They use stepper motors to drive the nozzles and pumps along the X-axis and Y-axis, and the rotation angle and speed of the stepper motors are controlled by pulse signals sent from the main board, and the frequency and number of pulse signals are determined by the concentration of pesticides that need to be configured [7].

The principle of this mechanism is to control the movement of X-axis and Y-axis by two motors at the same time, when the left and right motors are in the same direction, they move towards X-axis, and when the two motors are in the opposite direction, they move towards Y-axis. The simultaneous action of the two motors provides a more stable force than a single motor controlling one axis, and reduces the weight of one motor on the XY platform. The formula is shown in Fig. 6.
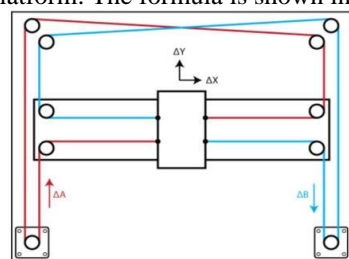
Fig. 6 Formula

2.7. Pesticide concentration configuration

They will configure pesticides in different concentrations according to the needs of different crops

There are several formulas for pesticide formulation:

Conversion between per cent concentration and parts per million concentration: parts per million (ppm) = $10,000 \times$ per cent concentration

Conversion between the multiplier method and the percentage concentration: Percent concentration (%) = (original concentration/dilution times) $\times$ 100

Calculate the amount of diluent according to the active ingredient:

Dilution of 100 times or less: Diluent dosage = Weight of the original agent $\times$ (concentration of the original agent - concentration of the formulated agent) / Formulation Concentration of formulated agent

Dilution more than 100 times: Diluent dosage = weight of the original agent $\times$ concentration of the original agent / concentration of the formulated agent

Calculate the amount of diluent by the multiplier method:

Dilution less than 100 times: Diluent dosage = Original agent weight $\times$ Dilution times - Original agent weight

Dilution more than 100 times: Diluent dosage = original agent weight $\times$ dilution times

Mass percent concentration: Mass percent concentration (%) = solute mass/solution mass 100%

Molar concentration: molar concentration (mol) = solute moles / solution volume (litres) [8].

Equivalent concentration: Equivalent concentration = gram equivalent of solute / volume of solution (litres)

Mass-volume concentration: Mass - volume concentration = mass number of solute (grams or milligrams) / volume of solution (cubic metres or litres)

2.8. Solar panel steering design

A common method in solar panel steering design is to use a microcontroller and a photoresistor module to collect light intensity data in four directions, and then convert them into digital signals through ADC, and then drive the servo or stepper motor to adjust the angle of the solar panel through comparison and control algorithms.

The specific calculation process is as follows:

Collect light intensity data in four directions through ADC, respectively x0 , x1 , y0 , y1, and can take the average value several times to improve the accuracy. Calculate the difference between the two directions of light intensity, respectively, dx = x0 - x1 , dy = y0 - y1, these two values reflect the solar panel and the degree of deviation from the sun. Based on the positive, negative and magnitude of dx and dy, the direction and angle to be adjusted are determined. For example, if dx > 0 and dy > 0, the solar panel needs to be rotated to the upper left; if dx < 0 and dy < 0, the solar panel needs to be rotated to the lower right; and if dx = 0 and dy = 0, the solar panel has been aligned with the sunlight.

According to the direction and angle to be adjusted, calculate the corresponding PWM signals or pulse numbers, and then generate PWM signals or pulse signals to drive the servo or stepper motor to rotate the corresponding angle by timer interrupt or other means.

Repeat the above steps until dx and dy are close to zero, i.e. the solar panel is perpendicular to the sunlight.

3. Software design

3.1. Pesticide spraying design

When the vision system detects that a plant is infested, it automatically feeds back a signal to the stepper motor system. The stepper motor system responds quickly and accurately drives the pesticide nozzle to position itself near the infected plant. The stepper motors then co-ordinate the spraying of the pesticide to ensure efficient and precise spraying of the pesticide in the affected area. This intelligent system not only responds to pest problems in a timely manner, but also provides a reliable solution for plant healthcare by maximising spraying efficiency through the collaborative operation of the vision system and stepper motors [3], [7]. The stepper motor working process is shown in Fig. 7.
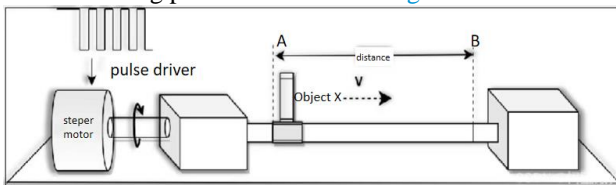
Fig. 7 The stepper motor working process

3.2. Visual Inspection

(1) Acquisition of plant leaf datasets

They acquired a detailed and rich batch of plant leaf datasets and leaf disease datasets, a valuable resource that will fully support in-depth research. The data covers multiple dimensions, including plant growth, dynamics of physiological parameters, and the potential impact of environmental factors on tobacco quality. By carefully analysing this dataset, they will be able to gain insight into all aspects of plant growth. The plant leaf dataset is shown in Fig. 8, and the diseased plant dataset is shown in Fig. 9.
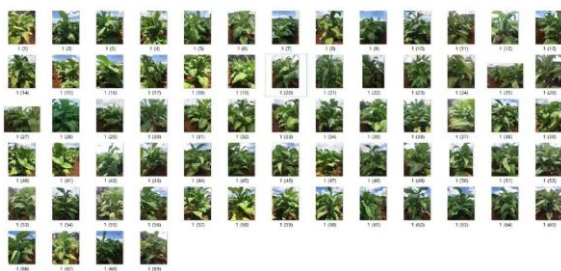
Fig. 8 The plant leaf dataset

Fig. 9 The diseased plant dataset

(2) Annotate

He carefully and systematically annotated the leaf dataset and leaf disease dataset of this batch of plants. Through the annotation scheme designed by him, he captured and recorded the characteristics of the plants at different growth stages, from roots to leaves, from physiological parameters to environmental adaptations. And he captured the characteristics of plants suffering from pests, diseases and so on. This kind of annotation not only enriches the connotation of plant data, but also lays a solid foundation for subsequent data analysis and model construction. The process diagram of leaf annotation is shown in Fig. 10, and the process diagram of disease point annotation is shown Fig. 11.
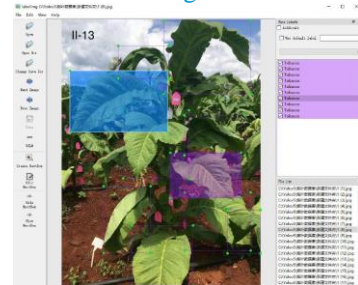
Fig. 10 The process diagram of leaf annotation
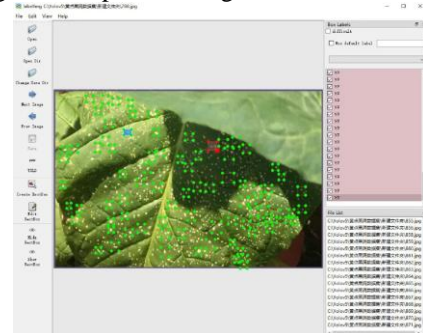
Fig. 11 The process diagram of disease point annotation

(2) Identification

Through CNN neural networks, he successfully implemented the functions of leaf recognition and pest identification [9], and cleverly combined them to form an efficient and comprehensive system. Through deep learning and computer vision techniques, he accurately recognised different types of tobacco leaves in the leaf images, and based on this, the system was able to keenly detect possible insect pests through careful algorithm design. The leaf recognition and pest recognition diagram is shown in Fig. 12. The final recognition result is shown in Fig. 13.

Fig. 12 The leaf recognition and pest recognition diagram
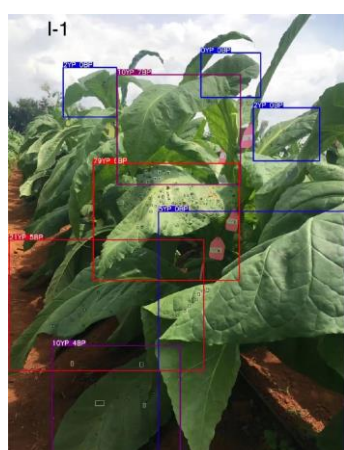
Fig. 13 The final recognition result

3.3. Optimisation

Since the disease target is very small, the sensory field of the original smallest detection head is still larger than the size of the target, so it is necessary to improve the original network, starting from the optimisation algorithm, firstly, the activation function is improved, they use leaky relu activation function in the middle layer, and the S-type activation function is selected for the last detection layer, and at the same time, the mish activation function is used; and the loss function is improved , using the CIOU loss as the loss of the bounding box, which can

bring faster convergence speed at the same time, have better performance [10]. The activation function is shown in Fig. 14.
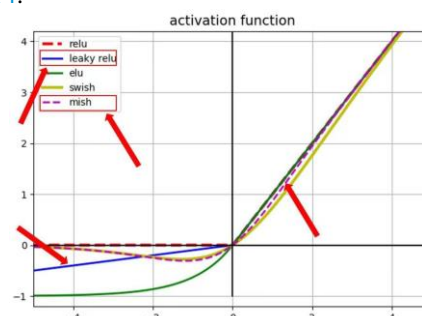
Fig. 14 activation function

4. Experiment

4.1. Pesticide spraying programmes

In the smart greenhouse system, they have introduced an advanced disease monitoring and prevention mechanism. Once the system recognises the presence of a plant disease, the automated control system is instantly activated, transferring the information to the stepper motor system. The stepper motors precisely manoeuvre the pesticide nozzles to move to the exact location of the infected plant. The intelligent system then implements precise and efficient spraying according to a pre-set pesticide spraying programme to maximise the suppression of the spread of the disease.

Future research directions include further optimising disease detection algorithms, improving the system's ability to accurately identify different disease types, and incorporating advanced pesticide selection and use techniques to achieve more precise and sustainable disease prevention and control strategies. Such innovations will lead to more sustainable and efficient plant health management solutions in the field of smart agriculture.

4.2. Mobile Application Programming

By running a well-designed mobile application, he successfully implemented a real-time greenhouse monitoring system on his mobile phone. The application provides an intuitive and detailed user interface that enables users to easily monitor the status of the greenhouse remotely and access key environmental indicators in real time. Users can easily access data such as temperature, humidity, light and soil moisture, as well

as plant growth conditions in each area of the greenhouse, anytime, anywhere, using their smartphones [11].

Through the mobile app, users can set customised alarm thresholds, and the system will immediately send notifications to alert users of any abnormal changes in environmental conditions. This real-time monitoring and response mechanism provides users with timely decision support, enabling them to quickly take the necessary measures to safeguard plant health and growth.

In addition, the mobile app supports remote control of the greenhouse system. Users can adjust parameters such as temperature control, humidity regulation, light management, and irrigation system through the mobile phone to achieve intelligent remote control of the greenhouse environment. Such flexibility enables users to make adjustments according to real-time changing needs, improving the adaptability and operability of the greenhouse system. The running diagram of the programme is shown in Fig.15.

Fig. 15 The running diagram of the programme

5. Conclusion

Nowadays, with the rapid development of artificial intelligence and Internet of Things technology, the visualised smart greenhouse designed in this paper is precisely in line with the trend of technological development, combining the knowledge of microcontroller, machine vision, artificial intelligence, and many other things with each other to form a complete system, and it is believed that in the future, this design will be more widely used in the agricultural market. This design combines the knowledge of multiple disciplines,

improves the human-computer interaction experience, and improves efficiency while protecting people's safety, and I believe that the demand for this type of product in the agricultural market will grow.

References

1. Bufeng Zhang, Guoli Li, Puryu Liu. Design of remote monitoring system for agricultural greenhouses based on STM32. *Science and Technology Innovation*,2023,(27):205-208.
2. Fengshuo Wang,Li Zeng. Design and implementation of internet of things platform based on MQTT protocol. Computer Knowledge and Technology,2023,19(25):73-76.
3. Chen Peng, Jin-Lian Mo, Feng Wang. Research and design of machine vision in environmental monitoring system of agricultural vegetable greenhouse. *Journal of Hunan College of Arts and Sciences (Natural Science Edition)*,2021,33(02):68-72.
4. Meigui Ya. Safe production of agricultural products and the use of biopesticides. *Rural Practical Technology*, 2023, (10):93-94.
5. Jun Jia, Jingyu Li, Fangjuan DONG. Discussion on the application of internet of things technology in the design of monitoring system for smart agricultural greenhouses. Intelligent Agriculture Guide,2023,3(13):9-12.
6. Baoling Hu, Jun Ma, Lei ZHOU et al. Design of intelligent control system for air temperature and humidity in agricultural greenhouses. Journal of Lanzhou College of Arts and Sciences (Natural Science Edition),2023,37(05):68-74.
7. Renming Tan, Renjie Zhang, Tao Jiang. Design of closed-loop stepper motor position control system based on STM32. *Journal of Heilongjiang Institute of Technology (General Edition)*,2023,23(01):58-62.
8. Feineng Zheng. Representation of pesticide concentration and calculation of dilution. *Plant Protection Technology and Extension*,1994, (06):19-20.
9. Tong Zhao,Chaofeng Sha. Revisiting test sample selection for CNN-oriented models:considering model calibration. *Computer Science* 1-13.
10. Guowei Dai, Zhimin Tian, Jingchao Fan et al. A neural network structure search oriented enhanced recognition method for plant leaf diseases. *Journal of Northwest Forestry College*,2023,38(05):153-161+193.
11. Xiaoyuan Bo,Jiaqi Wu,Zhiyi Hong et al. Design and research on the application programme of Huinong Huimin. *Modern Agriculture*, 2020, (11):17-19.

Authors Introduction

Mr. Yuntian Xia

He is currently pursuing his undergraduate degree at the School of Artificial Intelligence, Tianjin University of Science and Technology. His research area is neural networks.

Dr. Yizhun Peng

He is an Associate Professor in Tianjin University of Science &Technology. He received a doctor's degree in control theory and control engineering from the Institute of Automation, Chinese Academy of Science in 2006.His research field is intelligent robot and intelligent control.

A Design of Modular Photovoltaic Environmentally Friendly Portable Stroller

Suqing Duan, Yizhun Peng*

College of Electronic Information and Automation, Tianjin University of Science and Technology, China
*E-mail: *pengyizhun@tust.edu.cn*
www.tust.edu.cn

Abstract

The work is centered on ROS and integrates speech recognition and natural language processing modules to enhance the environmental awareness of the stroller and monitor the baby's condition to ensure a timely and appropriate response. The Raspberry Pi is the main control unit of the stroller and connects to a cloud-based IoT platform via the MQTT protocol. The platform facilitates seamless communication between the Raspberry Pi and cloud data for efficient data visualization on mobile devices. This innovative solution solves the challenge of balancing parental responsibilities with career advancement, promoting healthier and happier babies while allowing parents to maintain a harmonious work-life balance.

Keywords: Natural Language Processing, MQTT Protocol, Data Visualization, Cloud Computing, Raspberry Pi

1. Introduction

In order to reduce greenhouse gas emissions and at the same time enhance the safety and interactivity of the stroller, this work designs a modular photovoltaic intelligent portable stroller system with active safety, micro-environmental monitoring and remote monitoring of mobile terminals [1], and carries out a scientific and intelligent design of the software and hardware of the photovoltaic environmentally friendly portable stroller movement system and monitoring system [2].

The work uses polycrystalline solar panels to charge lithium-ion batteries, sets three power modes, and adopts aluminum alloy and other carbon-saving materials, which helps to reduce air pollution and greenhouse gas emissions, and responds to the national call for promoting green and low-carbon development [3]. The work uses STM32 development board as the bottom control center, and DC motor and drive circuit to build the bottom drive system hardware; adopts ROS as the core, and applies IoT technology with voice recognition and natural language processing module to realize the stroller's perception of the surrounding environment and the baby's state and make appropriate responses; adopts Raspberry Pi as the upper host computer of the stroller, and builds a cloud IoT platform and a cloud IoT platform based on MQTT protocol, and uses Raspberry Pi as the upper host computer. Raspberry Pi is used as the host controller of the stroller, and a cloud-based IoT platform based on MQTT protocol is built and the interaction scheme between Raspberry Pi and cloud data is designed to realize the convenient interaction function of data visualization in the mobile terminal [4].

The rest of the paper is organized as follows. The second part introduces the system program design. The third part introduces the working principle of the stroller, which includes the design of automatic following and safe obstacle avoidance function, automatic cradle function design, microenvironment sensing system design, baby crying sound sensing and recognition design and baby expression recognition design. The fourth part summarizes the main content of this paper.

2. System Programming

The team carried out the overall design and technical analysis of the system for the modular photovoltaic smart eco-friendly baby stroller. The function realization part of the stroller is a set of integrated motion device, which contains modules such as automatic cradle, autonomous following, etc.; and the monitoring part is a complete set of intelligent IoT solutions, which contains modules such as intelligent interaction, active safety, and visual surveillance [5]. The multi-angle view of the smart eco-friendly stroller is shown in Fig.1.
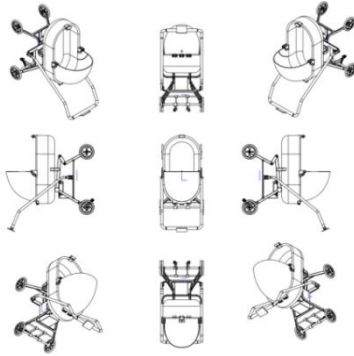
Fig. 1 The multi-angle view of the smart eco-friendly stroller

The modular photovoltaic intelligent eco-friendly stroller realizes the corresponding functions by controlling a variety of intelligent modules such as motor drive module, autonomous motion module, automatic brake module, automatic bed shaking module, deep learning module, visual monitoring module, intelligent interaction module and so on through the embedded system, and the intelligent system based on the Internet of Things (IoT) can make the stroller connected to the mobile terminal and provide real-time feedback for the parents on all kinds of information of the baby in the stroller. The system block diagram is shown in Fig. 2.
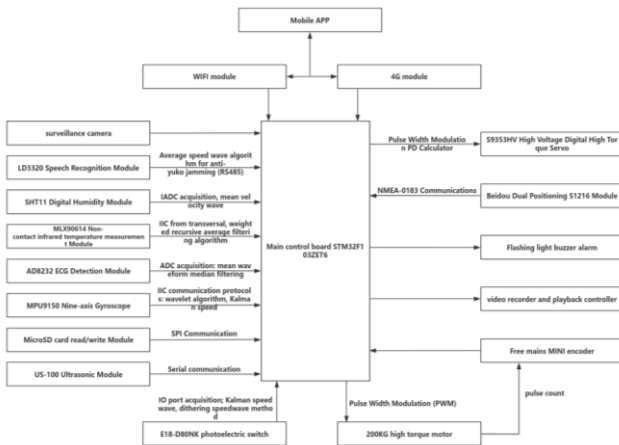
Fig. 2 The system block diagram

The sensor data collected by the controller is susceptible to external interference and is not highly accurate. In this work, different sensor data types are processed by mean filtering, weighted recursive average filtering, median filtering, Kalman filtering, dithering filtering, first-order hysteresis filtering, etc., and the data accuracy is improved. The schematic diagram of data processing is shown in Fig. 3.
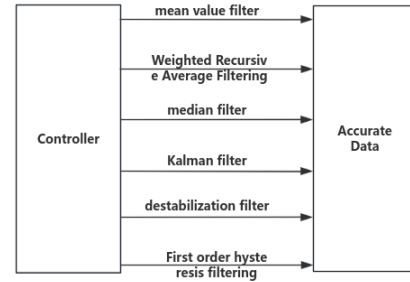
Fig.3 The schematic diagram of data processing

3. How Strollers Work

3.1. Automatic Following and Safe Obstacle Avoidance Function Design

The automatic following function is calculated by the ultrasonic distance measuring module through the acoustic wave reflection time length, to get the distance between the stroller and the person being followed, and through the PID algorithm to control the motor drive module to make the distance between the stroller and the person being followed stay within the set range [6]; the safety obstacle avoidance function also uses the ultrasonic distance measuring module to measure the distance, when the distance between the stroller and the obstacle in front of the stroller is less than the set value, the stroller brakes and stops. When the distance between the stroller and the obstacle in front is measured to be less than the set value, the stroller brakes and stops. In signal processing, Kalman filtering algorithm will be used to eliminate other interferences, to ensure that the smart stroller follows accurately, so that parents can have a more secure experience.

3.1.1. Distance Measuring Sensor

The ultrasonic distance module was selected for the distance sensor. Use low power consumption, low price and easy to connect HC-SR04 ultrasonic distance sensor. The module uses a transceiver split probe to range objects within a distance of 1cm-6m, to achieve the function of avoiding obstacles and following people. Ultrasonic distance measurement structure is shown in Fig. 4.
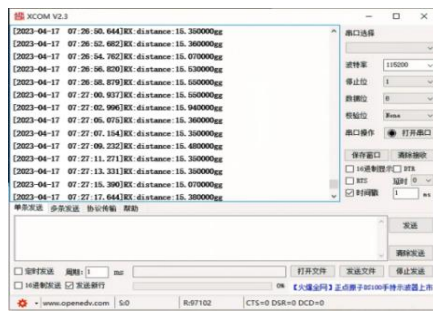
Fig. 4 Ultrasonic distance measurement structure

The team added temperature and humidity sensors to the work, thus using different speeds of sound for distance measurements by monitoring different temperatures, increasing the accuracy of the distance measurements and further securing the safety of the stroller. The temperature-speed of sound relationship graph is shown in Fig. 5.
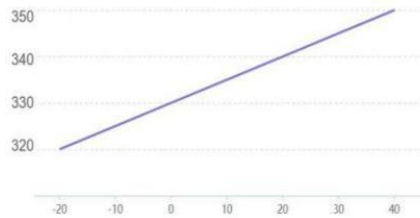
Fig. 5 The temperature-speed of sound relationship graph

3.1.2. Bluetooth Module

The system uses HC-05 Bluetooth module to realize two-way wireless function, through the micro-controller control device and voice control system communication, to realize the remote control cradle car front and back, left and right in all directions, or cell phone APP connects to Bluetooth to realize the cell phone APP control movement [7]. The cell phone Bluetooth app interface is shown in Fig. 6.

Fig.6 The cell phone Bluetooth app interface

3.2. Automatic Cradle Function Design

The design adopts the temperature and humidity compound sensor DHT11. It has humidity measurement elements to measure humidity and NTC temperature measurement elements to measure temperature, and its correction factor is stored in OTP memory by program.

3.2.1. Cradle Control Functional Design

The function of this module is divided into semi-automatic bionic rocking and fully automatic bionic rocking. When the semi-automatic cradle is turned on, parents can control it by connecting the cell phone APP through the Bluetooth module, and the stroller will simulate the parents' shaking once for every click on the turn on automatic shaking button; while when the fully automatic cradle is turned on, the automatic cradle module will work together with the cry detection module and the camera to automatically turn on the automatic cradle function when the baby's cry is detected and the function will decide the direction and strength of shaking according to the baby's posture fed back by the camera. This function will decide the direction and strength of rocking according to the baby's posture feedback from the camera, so that the baby is in the most comfortable cradle environment.

3.2.2. Attitude Sensing Module

The stroller is equipped with posture sensing module MPU6050, which sets the balance state data limit, autonomously monitors the swinging condition of the cradle without voice control [8], and the Raspberry Pi accepts the data and transmits it to the control board, which ensures that the cradle is always kept in a horizontal state without human intervention, providing safety for infants and toddlers to move around in the cradle. The test results of the attitude sensing module are shown in Fig. 7.

Fig.7 The test results of the attitude sensing module

3.2.3. Voice Interaction Module

The stroller is also equipped with voice interaction module LD3320, which can recognize the user's commands and make corresponding interactive commands. Considering that some babies are accustomed to listening to their parents' voices to sleep, when the stroller is told to "play mommy's words" or other commands, it will play the audio recorded by the parents in advance to coax the child to sleep, which reduces the pressure on parents' childcare. The test data of the voice interaction module is shown in Fig. 8.

Fig.8 The test data of the voice interaction module

3.3. Design of Microenvironment Sensing Systems

Sensing system design, temperature and humidity sensors can monitor the temperature and humidity in the stroller, when the temperature is too high, it will remind the baby moms and dads through the cell phone to ventilate in time, and when the humidity is too low, it will prompt the baby moms and dads to turn on the humidification device; the light intensity sensor automatically senses the strength of the light around the sun, and when it meets with the strong light environment, the sunshade will automatically extend and unfold, and play the effect of the sun, and if the baby moms and dads want to let their children sunbathe in the sun, they can turn it off the automatic sun-shading function through the cell phone app and turn the sunshade on.

3.3.1. Light Intensity Sensor

BH1750 internal by the photodiode, operational amplifier, ADC acquisition, crystal, etc. PD diode through the photovoltaic effect will be converted into an input light signal into an electrical signal, amplified by the operational amplifier circuit, the voltage collected by the ADC, and then through the logic circuit is converted into a 16-bit binary number stored in the internal registers, in short, is that the stronger the light, the greater the photocurrent, the greater the voltage is. larger, and using

its high resolution can detect a larger range of light intensity changes.

3.3.2. Temperature and Humidity Sensor

DHT11 is a temperature and humidity composite digital signal sensor, the sensor to digital acquisition technology and sensor technology integrated development, with years of market reputation and excellent performance. The sensor's clever single-wire serial interface design scheme, so that power consumption, volume has a greater reduction in the extremely demanding scenarios, the sensor can still maintain good performance and stability. The temperature and humidity sensor operation interface is shown in Fig. 9.
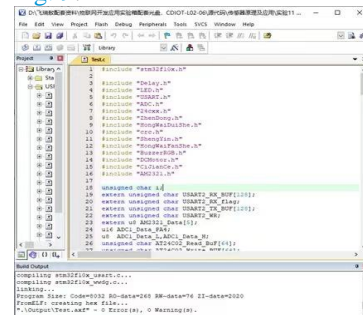
Fig.9 The temperature and humidity sensor operation interface

3.4. Infant Cry Sensing and Recognition Design

Infant cry perception and recognition design uses a cry recognition system, when the baby cries, the detection device will receive the cry signal, and then filtering processing, feature extraction through the basic acoustic features to get the feature data [9]. The results of the human infrared sensor test are shown in Fig. 10.

Fig.10 The results of the human infrared sensor test

3.4.1. Sound Classification Algorithm Based on Inception-v4 network

By studying Dunstan's theory of infant cry classification, our team collected videos and audios of infant cries from video websites and manually edited and labeled them to obtain the speech spectrograms of five categories of infant cry sound samples, which are the speech

spectrograms of sound samples of cries due to wanting to hiccup, being uncomfortable, feeling sleepy and wanting to go to sleep, being hungry and wanting to eat, and being bloated and in pain in the abdomen [10]. The five kinds of speech maps are shown in Fig. 11.
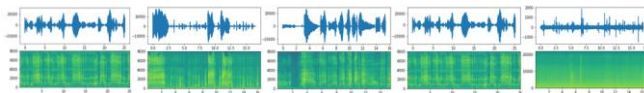
Fig. 11 The five kinds of speech maps

3.4.2. Network Model Structure Details

The database used in the work is the personal self-constructed infant crying database described in this paper, with a total of 7,500 spectrograms converted from infant crying data. In the experiment, the team used 6500 infant crying speech maps as a training set and 1000 infant crying speech maps as a test set. The categories were categorized using Oberschländerstein's theory of baby crying for 5 categories, which were manually labeled.

The tool used to build the network is the Tensorflow deep learning framework, and the task of categorizing different categories of infant cries is accomplished by modifying the Inception-v4 network model. The upper computer gives different feedback for different cries as shown in Fig. 12.

Fig. 12 Different feedback for different cries

3.5. Baby Expression Recognition Design

The baby expression recognition system consists of four main components, namely: baby expression image acquisition and detection, baby expression image preprocessing, baby expression image feature extraction, and matching and recognition.

3.5.1. Image Acquisition and Detection of Baby Expressions

In terms of baby expression image capture, different baby expression images can be captured by the camera lens. When the baby is within the capture range of the capture device, the capture device automatically searches and captures the baby expression image; baby expression detection is mainly used in practice for preprocessing of

baby expression recognition, i.e., accurately calibrating the position and size of the face in the image. Baby expression images contain rich pattern features, such as histogram features, color features, template features, structural features and Haar features [11]. Baby expression detection is to pick out this useful information and use these features to achieve baby expression detection. Baby expression detection is shown in Fig. 13.

Fig. 13 Baby expression detection

Mainstream baby expression detection methods use Adaboost learning algorithm based on the above features, which combines some of the weaker classification methods together to combine new very strong classification methods, which can effectively improve the detection speed of the classifier.

3.5.2. Baby Expression Image Preprocessing

Image preprocessing for baby expressions is a process of processing images and ultimately serving feature extraction based on face detection results. The original image acquired by the system often cannot be used directly due to the limitations of various conditions and random interference, and it must be subjected to image preprocessing such as grayscale correction and noise filtering in the early stage of image processing. For the baby expression image, the preprocessing process mainly includes light compensation, gray scale transformation, histogram equalization, normalization, geometric correction, filtering and sharpening of the baby expression image.

3.5.3. Baby Expression Image Feature Extraction

The methods of infant expression feature extraction are summarized into two main categories, one is the knowledge-based characterization method; the other is the characterization method based on algebraic features or statistical learning.

Knowledge-based characterization methods are mainly based on the shape description of infant expression organs and the distance characteristics between them to obtain feature data that help to classify infant expressions. Infant expression consists of eyes, nose, mouth, chin and

Suqing Duan, Yizhun Peng

other localities, the geometric description of these localities and the structural relationship between them can be used as an important feature for recognizing the face; the basic idea of the algebraic feature-based method is to transform the high-dimensional description of the face in the spatial domain into a low-dimensional description in the frequency domain or in other spaces, and the characterization methods are the linear projection characterization method and the nonlinear projection characterization method.

3.5.4. *Image Matching and Recognition of Baby Expressions*

The feature data of the extracted baby expression image is searched and matched with the feature templates stored in the database, and the result obtained from the matching is outputted by setting a threshold value when the similarity exceeds this threshold value. Baby expression recognition is to compare the face features to be recognized with the obtained baby expression feature templates, and to judge the baby state information based on the degree of similarity.

4. Conclusion

The Modular Photovoltaic Environmentally Friendly Portable Stroller's is dedicated to innovatively solving social problems through a modular photovoltaic smart portable stroller system, while reducing greenhouse gas emissions and enhancing the safety and interactivity of the stroller. The work uses polycrystalline solar panels, aluminum alloy and other environmentally friendly materials, as well as three power modes, fully responding to the initiative of green and low-carbon development. In terms of hardware design, the STM32 development board builds the underlying control center and implements the active safety and motion system. In terms of software, ROS serves as the core, combining with IoT technology, speech recognition and natural language processing to equip the stroller with the ability to intelligently sense and respond to the surrounding environment and the baby's state. Through the Raspberry Pi as the host controller of the upper computer, a cloud-based IoT platform based on the MQTT protocol is built, which realizes the convenient interaction of data and the visualization and monitoring of the mobile terminal. By improving the intelligent level of the stroller, it enables parents to better balance their work and family

responsibilities, and realizes the further penetration of intelligence and advancement in the field of childcare.

References

1. Wan Tao, Feng Junsen, Chen Jin et al. Design and realization of intelligent safety stroller control system[J]. *Internet of Things.* 2022,12(10):102-106.
2. Liu Xiaocheng. China's stroller: the current situation and countermeasure analysis of expanding overseas market[J]. *Foreign Economic and Trade Affairs.*2011.(12):4-7.
3. Liu Yu, Fan Guangjie. Research on the problems and countermeasures of promoting the concept of green low-carbon development into high school civics class[J]. *New curriculum introduction.* 2023(29):9-12.
4. Fu Bairui, Chen Jiaping. Embedded agriculture based on Raspberry Pi and LabVIEW[J]. *Electronic Design Engineering.* 2023,31(20):119-123.
5. Liu Yuxi, Zheng Ming, Wang Xuebin et al. Design of a new type of intelligent stroller[N]. *Proceedings of the Professional Committee on Educational Teaching Innovation of China Society for Educational Development Strategy Volume IV.*2019:311-314.
6. Wang Lufeng. Research on DC drive motor speed regulation for electric vehicles based on PID control[J]. *Automotive practical technology.*2020(10):106-108.
7. Li Minghuan, Peng Jian, Tan Jianfeng et al. Intelligent stroller based on wireless sensor network and human-computer interaction control[J]. *Internet of Things Technology.* 2022,12(09):127-131.
8. Ding Zulei, Jiang Tianze, Wwn Xiuping. Research on autonomous motion navigation system based on GPS positioning[J]. *Automation and Instrumentation.* 2022,37(08):1-4+29.
9. Qu Jin. Design of intelligent interaction system based on speech recognition[J]. *Automation and Instrumentation.*2023, (01):221-225.
10. Yang Han, Wang Miaoqin, Li Yanxi, YANG Chongyao, CHEN Maoqing, LI Tarojun. Knowledge graph analysis based on CiteSpace domestic infant crying research field[J]. *Chinese Journal of Child Health Care.*2022,30(10):1112-1117.
11. Liu Shupeng, Chen Zhiqiang, Chen Na et al. A set of experiments on Tensorflow-based CNN-RNN fusion architecture[J]. *Industrial Control Computer.* 2019,32(08):66-68.

Authors Introduction

Ms. Suqing Duan

She is currently pursuing her undergraduate degree at the School of Electronic Information and Automation, Tianjin University of Science and Technology. Her research field is embedded system.

Dr. Yizhun Peng

He is an Associate Professor in Tianjin University of Science &Technology. He received a doctor's degree in control theory and control engineering from the Institute of Automation, Chinese Academy of Science in 2006.His research field is intelligent robot and intelligent control.

"Teenage Mutant Ninja Turtles" - A Design of a Bionic Quadrupedal Rescue Robot

Hongpi Zhao, Yingfan Zhu, Zhihan Zhao, Xin Liang, Lei Lv, Yizhun Peng*

Tianjin University of Science and Technology, Tianjin, 300222,China

*E-mail: *pengyizhun@126.com*

www.tust.edu.cn

Abstract

"Teenage Mutant Ninja Turtles - Bionic Quadrupedal Rescue Robot is a quadrupedal robot based on the principle of bionics, inspired by the quadrupedal animals in nature. The robot has rescue and life detection capabilities and can perform rescue operations at disaster sites. The design of the robot enables it to operate efficiently and stably in complex environments, and at the same time it has the qualities of bravery and toughness, which are in line with the image of the Teenage Mutant Ninja Turtles. The robot can be used in a wide range of application scenarios, such as earthquake, fire and other disasters, to provide more efficient and safer support for rescue work.

Keywords:Biomimetic Machinery,Artificial Intelligence, Life Detection, Wifi Module

1. Introduction

With the development of society, mankind is faced with more and more natural disasters and man-made accidents, such as earthquakes, fires and mining accidents, which often result in a large number of casualties and property losses, bringing great harm to society. At the scene of these disasters, rescuers often need to risk their lives to enter the danger zone to carry out life detection, environmental exploration, material transportation, etc. Constrained by the number of rescuers, capacity, equipment, time and other factors, it is difficult to achieve high efficiency, safety and comprehensiveness [1].

To solve this problem, a bionic quadruped rescue robot called "Teenage Mutant Ninja Turtle" has been developed. The animated image of "Teenage Mutant Ninja Turtles" is a brave, invincible and courageous little turtle, and the rescue robot designed by the team has the same qualities of bravery and toughness, "Teenage Mutant Ninja Turtles" animated image is a brave, invincible and courageous little turtle, and the rescue robot designed by the team has the same qualities of bravery and toughness essence of the spirit, and therefore the name of this rescue robot for the "Teenage Mutant Ninja Turtles". The legs of "Ninja Turtle" use eight servos as joint actuators, and the legs can be replaced with wheel-type, bionic foot-type, negative pressure suction cup-type and three kinds of modularized structures to cope with different working scenarios, such

as flat, rugged,vertical and muddy.It is equipped with acoustic obstacle avoidance system with superb obstacle-crossing ability.It is also equipped with 4G infrared transmission for working in dark environments and entering collapsed pits for life detection and environmental exploration [2].

The rest of this article is organized as follows. Part II describes the modules used in the hardware system design. Part III describes the design of the bionic mobile joint. Part IV builds the model and tests to determine the foot structure. Part V summarises the main points of the paper [3], [4].

2. The Hardware System Design

In terms of hardware design, the Teenage Mutant Ninja Turtles Bionic Quadrupedal Rescue Robot has been carefully PCB-drawn, including the design and layout of several key modules, to ensure the stable and efficient operation of the robot. The design of the PCB board is shown in Fig. 1.

Hongpi Zhao, Yingfan Zhu, Zhihan Zhao, Xin Liang, Lei Lv, Yizhun Peng
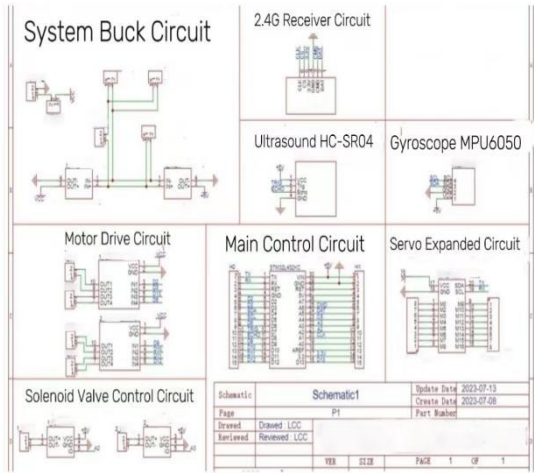
Fig. 1 The design of the PCB board

2.1 Circuit module section

Buck modules are an important part of power management, used to reduce the input voltage to a suitable operating voltage for each electronic component, helping to ensure system stability and power efficiency.

2.2 Motor drive module

Used to control the robot's motors, using PWM to control the speed and direction of the motors.

2.3 Bluetooth module

Enables the robot to communicate with external devices over Bluetooth, providing a convenient way for remote control and data transfer.

2.4 Ultrasonic sensors

Measures distance by sending and receiving ultrasonic signals for obstacle detection, navigation and localisation, providing robots with environmental awareness.

2.5 Wi-Fi wireless mapping module

Using the Wi-Fi wireless mapping module, the robot can transmit real-time images and video streams to remote devices, providing powerful support for teleoperation and environmental exploration [5].

The hardware design is shown in Fig. 2.
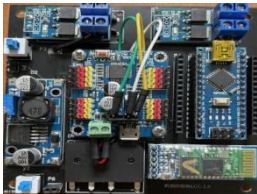
Fig. 2 Hardware design Hardware design

3. Design of bionic mobile joints

The parts design drawing is shown in Fig. 3. Detailed descriptions of functions and uses are shown in Table 1.
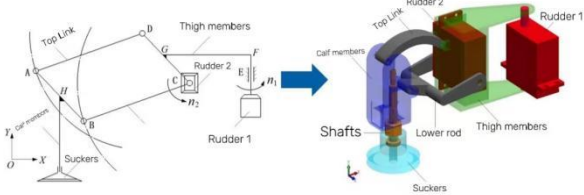
Fig. 3 Component design drawing

Table 1 Functions and uses

Functions and Uses	
Connecting rod	Transmission elements for connecting different joints and actuators to ensure coordinated movement of the robot's limbs
bear	Used to provide a rotational connection that allows the robot's joints to rotate freely without excessive friction
Servo 1	Used to control the swinging motion of the thighs
Servo 2	For controlling the elevation movement of the lower leg
shafts	Allows the robot to control the attachment and release of suction cups by varying air pressure, thus enabling the ability to climb and attach to different surfaces
suction pad	Uses a vacuum or pneumatic system to create a suction force to enure that the robot is firmly attached to the target surface.

A sophisticated system has been designed for Teenage Mutant Ninja Turtles Bionic Quadrupedal Rescue Robot in Pneumatic Adsorption System. The system has a strong adsorption and climbing function and ensures strong adsorption by negative vacuum -55KPa. The suction cups can withstand wall loads from 1.25kg to 1.7kg, allowing the robot to safely support its own weight and other additional loads. The design of the pneumatic adsorption system is complex and requires consideration of material selection, airflow control, negative pressure generation and other factors to ensure that the robot performs reliably in a variety of environments. Such systems are not only used in bionic robots, but are also widely used in industrial automation, glass suction handling, and handling equipment [6]. The pneumatic adsorption system is shown in Fig. 4.
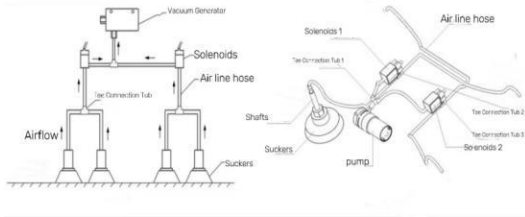
Fig. 4 Pneumatic adsorption system

4.Modelling and Testing

4.1 Multi-body dynamics modelling

In the mechanism dynamics simulation, a multi-body dynamics model is adopted, focusing on the kinematic mechanism of the robot leg, and the swinging of the thigh joint, the lifting and lowering of the calf joint, and the linkage between the thigh joint and the calf joint are simulated and analyzed, respectively [7]. And the model diagram is shown in Fig. 5.
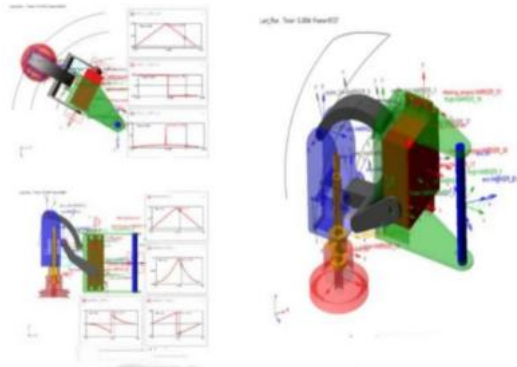
Fig. 5 Multi-body dynamics model

4.2 Stress and deformation maps

The structure of the Teenage Mutant Ninja Turtles (TMNT) bionic quadrupedal rescue robot was studied in depth in a finite element static analysis to evaluate its performance under different workloads and environmental conditions. This analysis includes the generation of stress and deformation maps, as shown in Fig. 6.
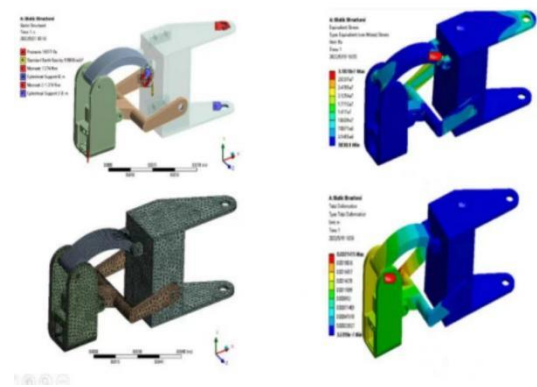
Fig. 6 Stress and deformation cloud diagrams

Stress distributions also help to ensure that the robot does not suffer from material fatigue or damage during operation; deformation clouds show how the robot's structure deforms under different loads. They are presented through colour coding or arrows to visualise the change in shape of the robot under stress.

5.System testing and results

Continuous optimizations and upgrades are performed to ensure good product performance. Through continuous product iterations, system testing and results are closely monitored to validate and improve the effectiveness of each iteration round. Over the course of the program's continuous development, five generations of products have been witnessed with significant progress and improvements.Firstly, the first generation of the product identified a quadrupedal structural solution with a four-link structure, which significantly enhanced the airframe's manoeuvrability. However, this was only the beginning [8]. The first generation is shown in Fig. 7.
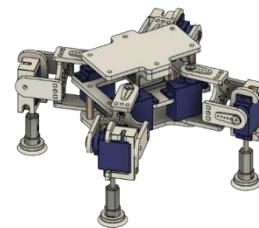
Fig. 7 First generation products

Subsequently, the introduction of the second generation of the robot brought significant improvements with the adoption of a high torque servo, MG996, to make the robot more agile and the introduction of modularly designed legs to improve ease of maintenance. The second generation is shown in Fig. 8.

Hongpi Zhao, Yingfan Zhu, Zhihan Zhao, Xin Liang, Lei Lv, Yizhun Peng
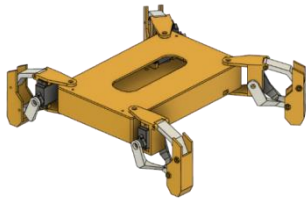
Fig. 8 Second-generation products

The third generation product further extends the function by adding tail and head structure, introducing automatic obstacle avoidance function and wireless mapping technology to make the robot more intelligent and adaptable to different environments. The third generation product is shown in Fig. 9.
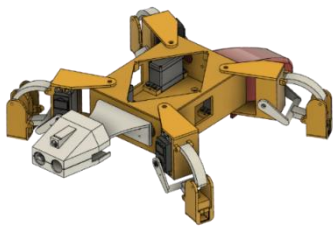
Fig. 9 Third-generation products

The highlight of the fourth generation is the addition of tracks and McNamee wheels, which give the robot the ability to not only move in all directions, but also steer in a more flexible manner. The fourth generation is shown in Fig. 10.
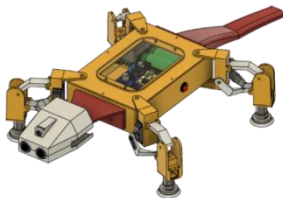
Fig. 10 Fourth generation products

Finally, the fifth generation product has undergone an in-depth industrial design to resemble the appearance of the robot to a turtle, giving it a more vivid and adorable image and enhancing its user-friendliness. The fifth generation product is shown in Fig. 11.

Fig. 11 Fifth generation products

At each iteration stage, we conducted system tests and analysed the results to ensure stability and performance optimisation of the new features. These tests included the robot's performance in various terrains and environments, its manoeuvrability, the reliability of the automatic obstacle avoidance function, the effectiveness of omnidirectional movement, and the aesthetic design.

References

1. Gao Liang; Gao Jingsong; Xu Wei; Zhao Jie. Research Status and Development Prospect of amphibian bionic robot [J]. Robot Industry,2023
2. Research on the Development Status and trend of biomimetic robot [J]. Menbao; Fan Xuekun; Chen Yongxin. Robot Technology and Application,2019
3. Gait planning of gecko Robot based on inverse Kinematics [J]. Li Hu; Sun Shudong; Duan Jun. Mechanical Science and Technology,2011
4. Infrared induction control system [J]. Gong Shulan; Guo Wencheng. Private Science and Technology,2007(09)
5. Application and Development of Rescue Robot Communication System [J]. Chen Wei; Mali; Wang Zhu. Fire Science and Technology,2016(08)
6. Analysis of Gait Control and Switching Strategy for Quadruped bionic Robot [D]. Zhao Shuangyu. University of Electronic Science and Technology of China,2014
7. Coordinated Gait Control of Hexapod Robot [J]. Han Jianhai, ZHAO Shushang, LI Jishun. Mechanical and Electrical Engineering,2004(04)
8. Mechanism Analysis and Design of a New Hexapod Robot [D]. Xie Zhong. Yanshan University,2019

Authors Introduction

Mr. Hongpi Zhao

He is currently pursuing his undergraduate degree at the School of Electronic Information and Automation, Tianjin University of Science and Technology.

Ms. Yingfan Zhu

She is currently pursuing her undergraduate degree at the School of Electronic Information and Automation, Tianjin University of Science and Technology. Her research area is robotics engineering.

Ms. Zhihan Zhao

She is currently pursuing her undergraduate degree at the School of Electronic Information and Automation, Tianjin University of Science and Technology. Her research area is robotics engineering.

Mr. Xin Liang

He is currently pursuing his undergraduate degree in the School of Mechanical Engineering at Tianjin University of Science and Technology. His research area is mechatronics.

Ms. Lei Lv

She is currently pursuing her undergraduate degree in the School of Mechanical Engineering at Tianjin University of Science and Technology. Her research area is industrial design.

Dr. Yizhun Pen

He is an Associate Professor in Tianjin University of Science &Technology. He received a doctor's degree in control theory and control engineering from the Institute of Automation, Chinese Academy of Science in 2006. His research field is intelligent robot and intelligent control.

Design of Grass Lattice Planter for Complex Environment Based on Adaptive Suspension Technology

Shaokai Tian, Wenqi Fu, Yizhun Peng*
Tianjin University of Science and Technology, China

*E-mail: *pengyizhun@126.com*

www.tust.edu.cn

Abstract

The Complex environmental grass grid growers based on adaptive suspension adopted the high-performance stm32 as the master chip, and use the Bluetooth communication module, so that users can easily control the vehicle driving, feeding and the posture of the pressing wheel through the mobile phone application. In addition, we have introduced new mechanical structures such as adaptive suspension to ensure effective planting of grass squares in complex terrain such as sloping land. Our vehicle uses a Mecanum wheel motion system to squeeze seeds by rotating the friction wheels at high speed.

*Keywords:*STM32F103ZET6, Adaptive suspension, Machine vision, Machine learning

1. Introduction

In recent years, severe weather has become more and more frequent. In some inland China, the rainfall is decreasing year by year, and the land desertification is becoming increasingly serious. The sharp decrease of arable land has led to the harsh environment such as yellow sand, which seriously affects People's Daily life. With the scientific concept of development to today's clear waters and green mountains are gold and silver mountains, China has formed a set of localized sand control route: grass square sand fixing. Grass square sand is getting more and more attention because of its extraordinary sand prevention effect [1]. However, most of the existing grass square planting methods in China are still artificial. Although there has been professional grass planting equipment in recent years, the investigation and interview found that most of the machines can only be applied to flat land. Under rolling roads and dune landforms, there is no machine to plant, and it is still manual work. The harsh and changing environment in the desert greatly increases the risk of manual work [2].

In order to conduct the actual test of the developed grass square planting system, the team designed and developed the corresponding mechanical structure that can control the movement and integrate the feeding and planting. In the circuit of this project, the stm32 microchip as the main control chip and the Bluetooth as the communication module can control the walking, feeding and grass wheel posture of the vehicle through the mobile phone APP [3], and then complete the grass square planting in the slope with the new adaptive suspension and other mechanical structures. The method is the McNham wheel motion system, which uses the motor to drive the friction wheel and squeeze the tennis ball at high speed to launch, and uses STM32 to control.

The remainder of this paper is organized as follows. The second part introduces the mechanical structure design for the implementation of the robot function. In the third part, the hardware circuit design of the robot is given. In the fourth part, the electronic control program design of the robot is given. The fifth part summarizes the main content of this article.

2. A Mechanical Structure Design

2.1. An Adaptive structure design

In the four-wheel movement system, the front and rear wheels on the same side are connected by the connecting rod structure shown in the figure, and the left and right side connecting rods are connected through two sets of

Shaokai Tian, Wenqi Fu, Yizhun Peng

swing arm structure. The swing arm can move up and down through the linear bearing, and finally the left front wheel and the right rear wheel move up and down synchronously, and the left rear wheel and the right front wheel move up and down synchronously. The above structure can realize the all-directional adaptive of various slope landforms, and can operate in the desert undulating terrain [4]. Adaptive suspension diagram is show in Fig. 1.

Fig. 1 Adaptive suspension diagram

2.2. An Electric swing arm link design

In order to better adapt to the undulating terrain such as desert and Gobi, the wheel group hanging swing arm is designed with three degrees of freedom, which can meet the rotation of Pitch axis, Yang w axis rotation and Y axis sliding at the same time.

The following is the details of the upper and lower movement of the swing arm: the village sleeve and the screw set the distance between the two steel swing arms, the flange is fixed on the swing arm through the screw, the straight bearing is fixed on the flange by welding, and the steel column is fixed on the steel plate 1 and 2 with the screw, thus fixed on the frame. The linear bearings can move up and down around the steel column or rotate. Structural details of the swing arm connecting rod contact point: the sliding shaft and the crosshead are fixed, through the sliding shaft, the sliding coat can be rotated by a small margin, and the swing arm can be rotated around the shaft through the thrust bearing and steel column. The distance of the distance is fixed between the two links by rigid jacket. Through the above mechanism, the movement realizes the slip and the rotation on the whole node space.[5]Details of the swing arm mechanism is show in Fig. 2.
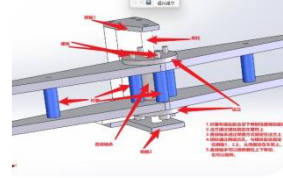
Fig. 2 Details of the swing arm mechanism

2.3. A Material supply design

The double output motor 1 is flange and the power is transferred to the roller 4. The roller is fixed on the aluminum pipe base through the bearing and steel shaft. The roller connected with the motor is the active wheel, and the roller not connected with the motor serves as the driven wheel.

During operation, the motor rotates, the active roller rotates, the active roller and the driven roller pre-tighten, the active rod rotates and drives, and produces friction with the grass between the main and slave roller, so as to realize the feeding function.

The slot explains: adding the slot can realize the adjustment of forage with different thickness and width, so as not to produce excessive stress between the rollers and increase the practical service life of the structure. Material supply mechanism diagram is show in Fig. 3.
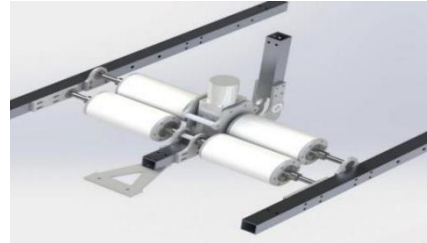
Fig. 3 Material supply mechanism diagram

2.4. A Management mechanism design

This structure adopts multi-link structure, through the rotating motion into linear motion, drive the upper rod 3 rotate around the hinge 8 , the swing of the upper rod 3 will make the shock absorber 4 compression, then transfer the force to the lower rod 6 , the lower rod 6 will rotate around the steel shaft centering by the horizontal support, so as to realize the upper and lower movement of the pressing wheel.

According to the calculation and analysis, the structure has only one degree of freedom, so the control of the structure can be fully realized by controlling the number of circles turned by the motor, and the structure is

simple and easy to use. Managing mechanism diagram is show in Fig. 4.
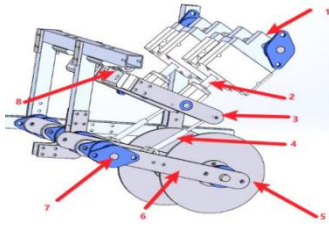
Fig. 4 Managing mechanism diagram

2.5. A Design of shock absorption structure

(1) The function of the shock absorber
- Realize the transmission of force;
- The expansion of the shock absorber 4 can ensure the balance between the grass pressing wheel and the ground of the force, to ensure the quality and efficiency of planting;
- Bear the impact of the irregular ground, protect the vehicle structure, improve the service life;

(2) Advantages: This system does not use the rigid connection, but uses the shock absorber as a flexible connecting rod, thinking from suspension, innovative application solved the rolling road may produce the disadvantages of rigid stress stress structure, and by ensuring the range between the ground wheel always small fluctuations and always fit the ground, so the car can not only in the plane planting, but also can be in the plane planting (slope, tunnel, etc.), and has the very good practical application [6]. Damper is show in Fig. 5.
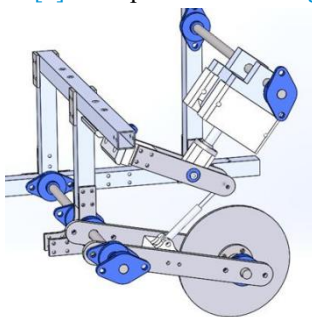
Fig. 5 Damper

3. A Hardware Circuit Design

3.1. STM32F103ZET6 Master control module

Up to 8 timers, containing a separate watchdog timer and a window-type timer, allowing for automatic microreset. This chip has up to 72 MHz working frequency. The bus system is a synchronous serial peripheral interface that enables the MCU to communicate with various peripherals in a serial manner to exchange information. The MCU architecture features an easy-to-use STM 32 platform ideal for motor drive [7]. STM32F103ZET6 Master control module is show in Fig. 6.

Fig. 6 STM32F103ZET6

3.2. A4988 Stepper motor drive module

Can adapt to the effects of various natural environments to better achieve work in harsh environments. The selected A4988 DMOS micro-step drive module with converter and overcurrent protection has more stable performance. A4988 Stepper motor drive module is show in Fig. 7.

Fig. 7 A4988 Stepper motor drive module

3.3. The HC-12 Bluetooth module

The super-long transmission distance and acceptance performance of the Bluetooth module can well realize the

purpose of long-distance control. The HC-12 Bluetooth module is show in Fig. 8.
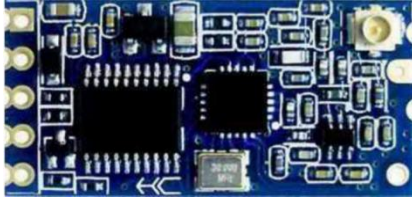
Fig. 8 The HC-12 Bluetooth module

3.4. L298N DC motor drive module

L298N can be used to drive our 12V DC deceleration motor, L298N can directly control the motor, set the control level through the I / O input of the main control chip, can drive the motor forward reversal, simple operation, good stability, can meet the high current driving conditions of the DC motor. L298N DC motor drive module is show in Fig. 9.

Fig. 9 L298N DC motor drive module

3.5. 12V DC deceleration motor

12V DC deceleration motor is widely used, because its excellent performance is often used in smart home, electronic products, smart car driver, precision medical equipment, etc. In this grass square unmanned planting car, the price is appropriate, product depression purchase and assembly and other excellent performance. L298N DC motor drive module is show in Fig. 10.

Fig.10 L298N DC motor drive module

3.6. Circuit logic control

The overall electrical route of the machine is arranged to make the machine line structure stable and safe, and will not be damaged due to external or internal interference, especially for the line layout where the line is concentrated. Circuit chassis design makes most of the circuit placed in the back of the body, increase the convenience of human-machine interaction, while the circuit, convenient line arrangement.

Select the power line, signal line and other electronic components needed for the robot, reasonably arrange the line of the robot, to ensure the stability and safety of the robot operation [8].

3.7. A Chassis module design

(1) Main components: Bluetooth module, STM32 main control, L298N DC motor drive module, 12V DC deceleration motor.

(2) Implementation mode: the main control board receives the chassis control instruction through the Bluetooth module, then calculates the corresponding instruction as the corresponding control information, and controls the motor speed with a PWM control pulse through the IO port.

(3) Electrical connection: the 12V power supply directly supplies power to the L298N DC motor drive module, and the motor drive module is connected to the motor and directly supplies power to the motor. The 12V power supply supplies power to the STM32 main control board after the 12-5V step-down module, and the development board supplies power to the HC-12 Bluetooth module through the 3.3V interface.

3.8. A Material supply module design

(1) Main components: Bluetooth module, STM32 main control, L298N DC motor drive module, 12V DC deceleration motor.

(2) Implementation mode: the main control board receives the control instruction through the Bluetooth module, then calculates the corresponding instruction as the corresponding control information, and issues the PWM control pulse to control the feeding motor through the IO port, so as to realize the material supply and discharge.

(3) Electrical connection: the 12V power supply directly supplies power to the L298N DC motor drive module, and the motor drive module is connected to the motor and directly supplies power to the motor. The 12V power supply supplies power to the STM32 main control board after the 12-5V step-down module, and the development board supplies power to the HC-12 Bluetooth module through the 3.3V interface.

3.9. A Management module design

(1) Main components: Bluetooth module, STM32 master control, A4988 step motor drive module, 42 step motor.

(2) Implementation mode: the main control board receives the control instruction through the Bluetooth module, then calculates the corresponding instruction as the corresponding control information, and sends the PWM control pulse and IO level control the rotation of the stepping motor through the IO port, so as to realize the control of the depth of the pressing structure.

(3) Electrical connection: the 12V power supply directly supplies power to the A4988 stepping motor drive module, and the motor drive module is connected to the stepping motor to directly supply power and control it. After the 12V power supply supplies power to the STM 32 through the 12-5V step-down module, and the development board supplies power to the HC-12 Bluetooth module through the 3.3V interface [9].

4. An Electric Control Program Design

4.1. A System main body logic design

Program flow chart is show in Fig. 11.
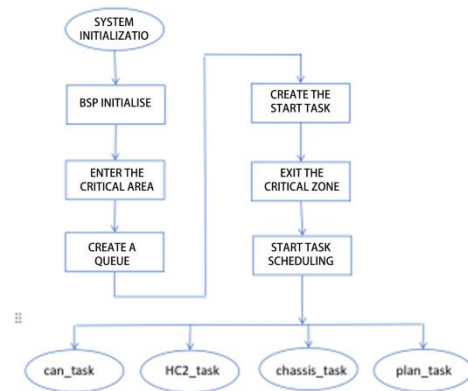
Fig. 11 program flow chart

4.2. A Chassis system logic design

Chassis.c Three subroutines are set up, namely, chassis motor initialization program, stepper motor initialization program and chassis control program.

(1) In the chassis initialization program, complete is the GPIO used for the chassis motor.

(2) In the initialization program of the stepping motor, the initialization of the IO port of the control stepping motor is completed.

(3) In the chassis control program, according to the control information after the incoming processing, set the PWM wave with different duty cycle ratio to complete the control of the motor speed and steering under the corresponding situation.

4.3. A Logical design of the communication system

After receiving the message at the serial port, we need to process it. However, because there is more logic in this project, we set up a separate structure to integrate the information flow and facilitate us to process the data uniformly.

In this structure, we integrate the information of chassis speed, direction, stepping motor status and feeding motor status. In the serial receiving interrupt, we directly judge the received message in the interrupt, a character check function is called to judge the instruction

Shaokai Tian, Wenqi Fu, Yizhun Peng

according to the received characters and rewrite the variables of the structure body.

In addition, in order to facilitate the user's control, we also set a timer interrupt, for timing to clear instructions, can realize the micro operation of the robot, complete fine control.

5. Conclusion

Dunes or gobi desert governance is a long time, now large machinery is mainly plane industry planting, planting can not complete the dunes, moreover, due to the dune itself has strong liquidity, not only lead to human resources in the same repeated continuous consumption, also has a potential danger: quicksand, a threat to people's life and property safety. However, in order to create a long-term sustainable green water peak, we shall not ignore the one in the field of governance, so the team proposed adaptive suspension has been practical, break through the existing laying machinery technical problems, can set foot on the ecological restoration in this field, and the wireless control, no one to close interference do industry, great guarantee the personal safety, the design breakthrough existing laying machinery technical problems.

References

1. Yaozong Wang, Xinbin Yue, Jiali Xie,et al,Desertification evolution in the sandy region to the east of the Yellow River in Ningxia from 2000 to 2020,Journal of Desert Research,2023,4:pp.31-40.
2. Chengzhi Hou, Danqing Huang, Dongwei Gui,et al,Spatiotemporal variations of climate extremes and influential factors in deserts and sandy fields of northern China from 1961 to 2019,Scientia Geographica Sinica,2023,8:pp.1495-1505.
3. Yan Zhang, Jiatong Li, Xiaoyi Song,et al,Survey of IoT Device Security Detection,Journal of Computer Research and Development,2023,10:pp.2271-2290.
4. Yang Liu, Unloading impact simulation and test of luffing jib tower crane, Journal of Jilin University: Engineering and Technology Edition,2022,6:pp.1292-1300.
5. Haimin Han, Design of Improved Damping Device Used in Mechanical Equipment, Mechanical Engineering & Automation,2022,4:pp.113-114.
6. Chunyang Liu,Fan Chen,Zhangfei Wang,et al,Design of lane keeping assistance control system based on STM32,Transducer and Microsystem Technologies,2022,3:pp.68-71.
7. Jinhua Wang, Jia Liu, Yonghe Zhuang, et al, A new logic control method of the high precision current/frequency converter, Microelectronics & Computer,2023,6:pp.85-89.
8. Mei Yang, Muhua Liu, Feiyu Zhu ,et al, Design of Soil Environmental Information Sensor Based on Bluetooth Technology, Acta Agriculturae Universitatis Jiangxiensis,2023,4:pp.963-971.
9. Linfeng Li, Yong Wang, Yongchun Xie, Learning Space Robotic Manipulation via Multi-View Visual Goal Generation, Aerospace Control and Application,2022,2:pp.18-28.

Authors Introduction

Mr. Shaokai Tian

He is currently pursuing his undergraduate degree at the College of Mechanical Engineering, Tianjin University of Science and Technology

Ms. Wenqi Fu

She is currently pursuing her undergraduate degree at the School of Electronic Information and Automation, Tianjin University of Science and Technology.

Dr. Yizhun Peng

He is an Associate Professor in Tianjin University of Science &Technology. He received a doctor's degree in control theory and control engineering from the Institute of Automation, Chinese Academy of Science in 2006. His research field is intelligent robot and intelligent control.

A Design of Intelligent Ecological Multifunctional Plant and Animal Breeding System

Suqing Duan, Yuntian Xia, Siyi Wang, Yizhun Peng*

College of Electronic Information and Automation, Tianjin University of Science and Technology, China
E-mail: *pengyizhun@tust.edu.cn
www.tust.edu.cn

Abstract

The intelligent ecological multi-functional system for animal and plant cultivation integrates PLC programmable controllers and microcontroller-based automatic control technology. This innovative approach amalgamates aquaculture with horticulture, fostering a symbiotic environment for the breeding of animals and plants. The system's operations are categorized into four key components: mechanical structure, environmental perception, automatic control, and intelligent networking. Its comprehensive functionality encompasses plant monitoring, self-checking of temperature and humidity, group control for fertilizer replenishment, consistent temperature regulation, full-spectrum illumination, advanced filtration, fogging supplementation, versatile water management, fertilizer blending, and water purification through sterilization.

Keywords: Intelligent Ecology, PLC programmable controller, Environmental sensing, Eco-agriculture, microcontroller, monitoring

1. Introduction

With the rapid development of science and technology, intelligent technology has emerged in the field of agriculture. Based on Siemens S7-1200 version PLC programmable controller and AT89C51 microcontroller automatic control technology [1]. this thesis aims to integrate aquaculture and horticulture planting, and innovatively realize plant and animal symbiotic aquaculture. The most important feature of the system is that it can realize ecological energy saving, environmental protection recycling and symbiotic breeding of plants and animals, which pushes the intelligent ecological breeding to a new height.

The work is mainly composed of four modules: mechanical structure, environmental sensing, automatic control and intelligent networking. System functions include plant monitoring, temperature and humidity self-checking, group control and fertilizer renewal, constant temperature and heating, full illumination, multiple filtration, fogging supplementation, multi-purpose water, fertilizer mixing and sterilization of water purification and other special functions [2]. In addition, the system introduces an animal feeding system, which includes an adaptive feeding cabin and an intelligent feeding system that monitors the physiological parameters of the animals and automatically feeds them at regular intervals and quantities in order to realize intelligent feeding management.

The rest of this paper is organized as follows. The second part introduces the top ten functional designs. The third part introduces the working principle of the intelligent ecological multifunctional plant and animal breeding system. The fourth part is hardware design, including mechanical structure part design, environment sensing part design, automatic control part design and intelligent networking part design. The fifth part is the software design. The sixth part summarizes the main content of this paper.

2. Ten functional designs

Intelligent ecological multi-functional animal and plant breeding system aims to realize fully automatic group control of plants and aquatic animals, which contains ten key functions such as plant monitoring, temperature and humidity self-checking, group control and fertilizer renewal, constant temperature and heating, full illumination and light supplementation, multiple filtration, and fogging supplementation. The application of these technologies not only significantly reduces manual intervention and improves breeding efficiency and production, but also promotes the recycling of resources and realizes environmentally sustainable production.

This intelligent ecological multi-functional plant and animal breeding system realizes all-round monitoring and management of plants and animals by integrating advanced monitoring and control technologies,

Suqing Duan, Yuntian Xia, Siyi Wang, Yizhun Peng

minimizing manual operations and thus improving overall breeding efficiency and production. The top ten functional diagrams of the system are shown in Fig. 1.
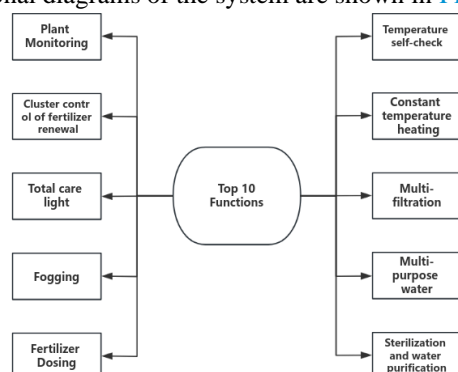
Fig. 1 The top ten functional diagrams of the system

2.1. Aquaponics Technology Design

Based on the concept of energy recycling, the system is guided by the concept of watering vegetables with fish water and improving the growth environment of fish with vegetables, integrating aquaculture and hydroponic cultivation technologies to build a harmonious aquaponics system. In urban family application, this technology can reduce the purchase expenditure of vegetables and fishery products and create a green and energy-saving family living environment.

2.2. Temperature parameter change automatic judgment sliding table contraction design

The system monitors the ambient temperature through an RS485 industrial-grade temperature and humidity transmitter to ensure that the slide table is intelligently triggered to retract automatically in adverse climatic conditions (less than 15 degrees Celsius or more than 30 degrees Celsius) to prevent plants from being damaged by cold or overheating. Meanwhile, through intelligent determination of the current time, the system realizes artificial light supplementation at 18:00 pm to ensure that plants can still obtain sufficient light when it is dark, while it automatically extends at 8:00 am to absorb sunlight to provide a good growing environment for plants. Such time determination and intelligent regulation mechanism effectively guarantees the growth needs of plants in different time periods.

2.3. Plant group control automatic humidification design

The system is equipped with a humidity sensor to monitor the humidity of the plant growth environment in real time. When the humidity is lower than 80%, the intelligent system automatically triggers the humidification device to ensure that the photosynthesis, respiration and transpiration of plants proceed normally. Through time triggering and intelligent determination of air humidity at 8 a.m., 11 p.m. and 5 p.m., the system decides whether or not to humidify to meet the humidity needs of the plants at different times of the day, thus providing a suitable growing environment [3]. This intelligent humidification system helps to maintain the physiological activities of plants and improve the breeding efficiency.

2.4. Plant group control automatic fertilization design

The system performs real-time sensing and calculation through sensors at the height of the plants, and judges the reasonable amount of fertilizer (CM/MG) according to the growth status of the plants. In terms of time triggering, the system intelligently sets 3:00 pm to 10:00 pm every day as the fertilizer application time period to ensure that the plants get the right amount of nutrients during this time period. In addition, soil sensors are used to monitor the levels of nitrogen, phosphorus, potassium and other fertilizers in the soil. The system intelligently determines the soil condition according to the plant species and the seasons, and applies the right amount of fertilizer when needed. Such an intelligent fertilizer application mechanism is designed to ensure that plants are supplied with appropriate nutrients, thereby improving breeding efficiency and yield.

2.5. Plant group control automatic watering design

The system decides whether watering is needed by checking whether the soil moisture reaches the set value through real-time soil moisture monitoring every 10 minutes. In addition, the system intelligently takes into account plant species and seasonal changes, and based on these factors determines whether watering is required for the day, ensuring that soil moisture meets the plant's growth needs. This integrated intelligent watering mechanism aims to provide plants with the right amount of water according to the actual situation, ensuring that they receive appropriate irrigation in different seasons and growth stages, and improving breeding efficiency.

2.6. Plant group control automatic nutrient replenishment design

The system intelligently calculates the corresponding levels of nutrients such as nitrogen, phosphorus and potassium (CM/MG) through plant height sensing. In terms of time triggering, the system is set to automatically replenish nutrients once a month to improve plant resilience and resistance and activate microbial vitality in the soil [4]. This integrated highly sensible and automatic replenishment mechanism helps to

ensure that plants are supplied with appropriate nutrients at different growth stages, improving their resilience and thus enhancing the ecological balance of the whole aquaculture system.

2.7. Automatic Oxygen Supply Design for Aquatic Animals

According to the change of water temperature, automatic oxygen supply is realized to ensure the oxygen demand of fish in different temperature ranges.

2.8. Automatic sterilization design

The system adopts the principle of ozone sterilization, through which ozone acts on the cell membrane of microorganisms, leading to cell death, so as to achieve the purpose of sterilization, anti-inflammation and detoxification. In addition, the system also applies UV germicidal lamps, which utilize UV rays to destroy the DNA or RNA molecular structure in the cells of various microorganisms in the water, effectively killing various pathogens in the water in order to keep the water quality clean. This integrated sterilization system is designed to ensure the hygienic condition of aquaculture water, improve water quality, and guarantee the healthy growth of plants and animals.

2.9. Solar photovoltaic power generation system design

The system is centered on PLC and HMI, and is powered by electricity from the solar photovoltaic power generation system reserve. According to the light condition, the system realizes automatic switching of power supply to ensure the normal operation of the equipment under different light conditions. In addition, the system intelligently utilizes LED plant growth lights for automatic replenishment. When natural light is insufficient, the solar PV system activates the LED plant growth lights to provide an additional light source to promote plant growth. This integrated PLC and solar photovoltaic power generation system, as well as the automatic replenishment function of the LED plant growth lamps, effectively guarantees the sustainable operation of the system and the growth needs of plants under different light conditions.

2.10. Design of LED Plant Grow Lights for Yield Improvement

Spectral regulation is crucial to plant growth, and red light promotes chlorophyll formation and carbohydrate synthesis, fuels the growth of long sunlight plants, delays the growth of short sunlight plants, and promotes seed germination. Theoretical and experimental evidence shows that red light has the strongest photosynthetic

effect, and the use of red-rich light sources for supplemental lighting can advance plant flowering and fruiting and promote organ formation. Blue light helps stomata open, promotes the entry of external carbon dioxide, increases the rate of photosynthesis, and is conducive to protein synthesis. Blue light accelerates the growth of short-daylight plants, slows down the growth of long-daylight plants, and promotes leaf growth. Artificial supplemental light using a blue light-rich light source can delay flowering and allow full plant growth [5].

Plant grow lights are electric light sources designed to stimulate plant growth by emitting electromagnetic waves suitable for photosynthesis, and are used in applications where natural light is not available or where additional supplemental light is required. In winter, when daylight hours are insufficient to meet the needs of plants, plant lights are used to extend light hours and prevent plants from growing unduly. Meanwhile, the plant curtain green light realizes cooling effect in summer by blocking heat, shading strong sunlight and providing a cool indoor environment; in winter, it prevents cold air from intruding and provides thermal insulation. The ambient light is used to provide extra light, highlight the color of the fish, and meet the fish's need for light. In the absence of sunlight, the ambient light creates a comfortable environment for the fish tank, creating a more perfect viewing effect. The finished rendering and exterior design are shown in Fig. 2.

Fig. 2 The finished rendering and exterior design

3. Working Principle

Firstly, the ornamental and convenient management of plants is realized through the design of plant sliding table. The plant sliding table can display the plants in the balcony or outdoor community, and when the environment changes, the plants can be cooled, humidified, and replenished with light through PLC control in order to reduce the labor cost and improve the breeding efficiency. The plant curtains are driven by 24V motor for 90 degree opening and closing. Secondly, the system supports hydroponic plant cultivation, which can save space utilization by putting the plant curtains away. The root system of hydroponic plants can filter the fish tank water flowing through the plant curtains, realizing the recycling of aquaponics resources. Finally, the system

is equipped with an industrial control touch screen to realize real-time feedback on the monitoring of plants and animals and human-computer interaction.

For automatic sensor detection, ultrasonic ranging, soil sensors, environmental sensors, PLC time calculation and other multiple logic judgments are used to improve the accuracy of plant control and help improve the survival rate of plants. Fertilizer pump adopts peristaltic pump, and through the "one drop one number" control method, it realizes the precision of fertilizer application and avoids the situation of using too much or too little fertilizer. In the filtration system, through a variety of filtration methods, such as hydroponic plant stalks, UV lamps, ozone and active filters, etc., the depth of the fish tank water poisoning and filtration, to achieve the healthy operation of the aquaponics system. The working schematic is shown in Fig. 3.
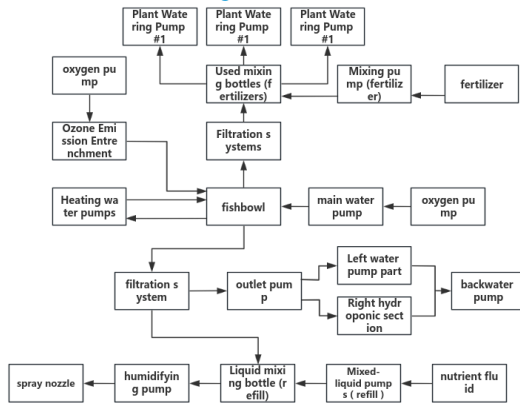
Fig. 3 The working schematic

4. Hardware Design

The overall structure of the intelligent ecological multifunctional plant and animal breeding system is divided into the following four major components. The overall structure is shown in Fig. 4.
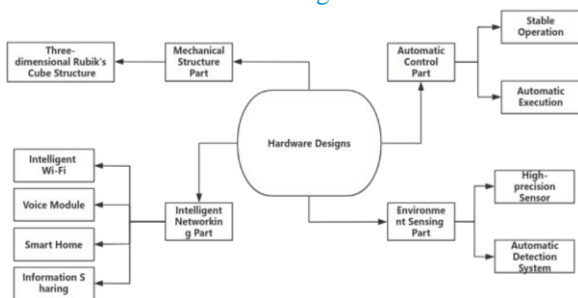
Fig. 4 The overall structure

4.1. Mechanical structure design

The three-dimensional Rubik's Cube structure is adopted, which is inspired by the aerospace station. The

three-dimensional Rubik's Cube structure is characterized by a large available space, which is divided into five different unfolding surfaces, with different functions corresponding to different surfaces, which are specifically divided into the "display and observation surface", "aquatic planting surface", "photovoltaic surface" and "energy collection surface" [6]. The different surfaces have different functions, specifically divided into "display and observation surface", "water planting surface", "photovoltaic surface", "energy collection surface", "sliding platform structure surface" and "control and manipulation surface".The mechanical structure is shown in Fig. 5.

Fig. 5 The mechanical structure

4.2. Environment Sensing Part Design

Composed of various high-precision sensors, the environment sensing part is equivalent to the "eyes" and "ears" of the intelligent ecological multi-functional animal and plant breeding system, which can automatically detect the working state of the system and convert various process parameters such as temperature, humidity, flow rate, liquid level and composition into uniform standard signals. Physical quantities such as temperature, humidity, flow, liquid level, composition, etc. are transformed into uniform standard signals to meet the requirements of information transmission, processing, storage, display, recording and control, etc. It is the first link to realize automatic detection and automatic control. The high precision sensor map is shown in Fig. 6.

Fig. 6 The high precision sensor map

4.3. Automatic control part design

The system is mainly composed of four parts: controller, sensor, actuator and transmitter. The controller is the core

of the system, responsible for automated control tasks without direct human intervention. Through a type of work, a state or a parameter in a machine, equipment or production process, it operates automatically according to pre-set rules. Electricity or other energy sources are used and converted into drive action by means of motors or other devices [7]. Sensors are responsible for detecting process parameters and transmitting the measured values in the form of specific signals for display and regulation. The role of the sensors is to convert various process parameters, such as temperature, flow rate, composition and other physical quantities, into uniform standardized signals, which are then transmitted to the regulator and touch screen for regulation, indication and recording. Such an automatic control system ensures the continuity of the production process and stabilizes the system operation. The controller and controlled object diagram is shown in Fig.7.

Fig. 7 The controller and controlled object diagram

4.4. Intelligent networking part design

As the communication hub of the intelligent ecological multi-functional animal and plant breeding system, the intelligent Wi-Fi module plays an important role in connecting with the outside world, including the close connection with other intelligent devices. As the core component of the smart home LAN, the module is able to independently complete a variety of functions of the home internal network without the intervention of other devices. It converts and shares information between communication protocols, realizing efficient data exchange between networks. Meanwhile, the data exchange function between the smart Wi-Fi module and the external communication network enables the smooth realization of intelligent functions such as remote control, scene control, linkage control and timing control. The design ensures an efficient and seamless connection between the smart farming system and external devices and networks, further enhancing the intelligence of the whole system [8]. The smart Wi-Fi module diagram is shown in Fig. 8.

Fig. 8 The smart Wi-Fi module diagram

5. Software Design

5.1. System logic circuit diagramEnvironment Sensing Part Design

The software system of the Intelligent Ecological Multifunctional Animal and Plant Breeding System is the "brain" of the system, which is responsible for monitoring, controlling and coordinating various components to realize efficient animal and plant breeding. The system logic circuit diagram is a graphical representation of the internal logic structure of the software system, showing the relationship and information flow between different modules. The system logic circuit diagram is shown in Fig. 9.
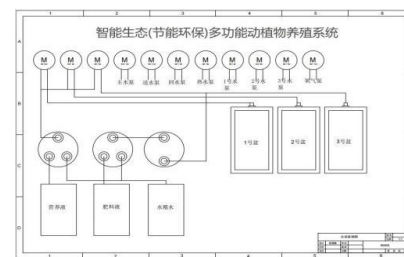
Fig. 9 The system logic circuit diagram

5.2. Human-computer interaction diagrams

The human-computer interaction diagram of the software system of the intelligent ecological multifunctional animal and plant breeding system is the part of the system that interacts with the user directly, providing the user with a friendly operation experience through an intuitive graphical interface. The human-computer interaction diagram is shown in Fig. 10.
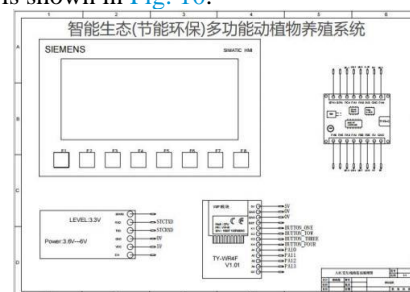
Fig. 10 The human-computer interaction diagram

5.3. Relay and Communication

The relay and communication diagram of Intelligent Ecological Multifunctional Animal and Plant Breeding System software system includes PLC, sensors, relay module, communication module, network connection, data transmission protocol, remote monitoring equipment, etc. PLC realizes equipment control through relay module, sensors monitor environmental parameters, and communication module realizes equipment data transmission and remote monitoring [9]. The relay and communication communication diagram is shown in Fig. 11.
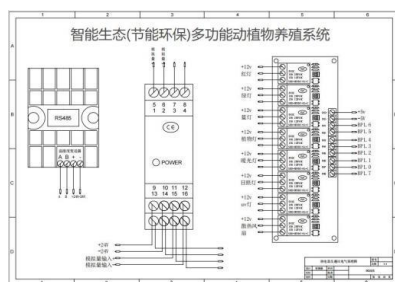
Fig. 11 The relay and communication communication diagram

5.4. Aquaculture System Logic Diagram

In the software system of Intelligent Ecological Multifunctional Animal and Plant Farming System, the logic diagram of aquaculture system is mainly related to the monitoring, control and optimization of aquatic ecology. The logic diagram of the aquaculture system is shown in Fig. 12.
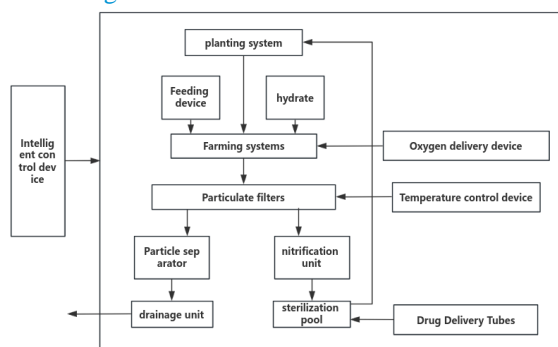
Fig. 12 The logic diagram of the aquaculture system

5.5. Human-computer exchange design

Our system integrates advanced technologies and provides five control modes such as touch, WIFI, voice, Bluetooth, and physical buttons, which provide users with more convenient means of controlling and managing animal and plant breeding. At the same time, the system also introduces intelligent self-learning function, which can automatically update the best planting plan to improve the breeding efficiency and realize the effective use of resources.

5.5.1. Main Picture

The main drawing is shown in Fig. 13.

Fig. 13 The main drawing

(1) Manual mode screen selection
(2) Automatic mode screen selection
(3) Data monitoring screen selection
(4) Running status screen selection
(5) Plant weekly record screen selection
(6) Plant status screen selection

5.5.2. Manual Mode

The manual mode diagram is shown in Fig. 14.

Fig. 14 The manual mode diagram

The system status is divided into running and stopped, and the water colonization system performs sterilization and water filtration every 15 minutes. At the same time, the plant aquatic curtains and outreach sliding table has five states: not in motion (stopped), retracted (being retracted), retracted to complete, stretching out (being stretched out), stretched out to complete. The logic of such states is clear and helps the user to understand the system operation status in real time.

5.5.3. Automatic Mode

The automatic mode diagram is shown in Fig. 15.

Fig. 15 The automatic mode diagram

When the system is running, the user can set the running time of the automatic mode through the timer function to ensure that the system automatically performs the breeding tasks according to the preset plan within a specific time period. At the same time, the system monitors the growth of the plants in real time and ensures that the necessary breeding operations, such as fertilization and watering, are carried out at the right time through the real-time feedback of the plant height. In addition, the system also takes into account the needs of different plants and provides a potted plant species selection function, which allows users to choose the appropriate breeding method according to the specific plant species, ensuring a more intelligent and adaptable system.

6. Conclusion

Our system is of great significance in promoting the development of modern agriculture. Facing the problems of shortage of social agricultural workers, low production efficiency and low value-added of agricultural products, our design builds a unique and independent animal and plant breeding system based on technologies such as artificial intelligence, Internet of Things, cloud computing and big data. In addition, our plant and animal breeding system also has the potential for mixed plant and animal breeding in large-scale agro-ecological farms. Through smart WiFi and 3D management functions, the system is able to realize precise management of plants and animals, which improves the efficiency of agricultural production. More importantly, this system helps to solve the problem of environmental pollution, especially in large-scale agro-ecological farms, where our technology can effectively minimize the impact on the environment and provide more high-quality ecological agricultural and sideline products. By enhancing people's aesthetic requirements for a beautiful ecological environment, our work plays an active role in building a more sustainable and healthy agro-ecosystem.

References

1.Wang Jianju. Siemens S7-1200PLC communication research[J]. *Southern Agricultural Machinery.* 2021,52(18):146-149.
2.Yuan Quan, Liu Yongjun, Huang Weiwei et al. Characteristics of water quality indicators and microbial community diversity in fish-vegetable symbiotic aquaculture system[J]. *Jiangxi Journal of Agriculture.* 2023,35(09):156-161+168.
3.Li Hongmei,Wu Jinji,Wang Liangming et al. Design and test of intelligent cultivation equipment for home fungus[J]. *Chinese Journal of Agricultural Mechanical Chemistry.* 2023,44(08):75-80.
4.Zhu Panpan, Ma Yanping, Zhou Zhongxiong et al. Trace element zinc and plant nutrition and human health[J]. *Fertilizer and Health.* 2021,48(05):16-18+23.
5.Long Jia-huan,Pu Min,Huang Zhi-wu et al. Research progress in spectral regulation of plant growth and development[J]. *Journal of Lighting Engineering.* 2018,29(04):8-16.
6.Zhang Xiaoyu,Liu Chang,Shi Liming et al. Optimized design of mobile deformable assembly with integrated support structure of skinned dot matrix and space station application[J]. *Journal of Solid Mechanics.* 2022,43(05):551-563.
7.Fang Zewen. Application of configuration king and Siemens PLC in wastewater treatment control system[J]. *China Equipment Engineering.* 2022(09):269-271.
8.Song Ruibo,Zhang Yan,Lian Menghui. Design of smart home system based on internet of things[J]. *Modern Industrial Economy and Informatization.* 2023,13(10):99-101+106.
9.Chen Ying. Intelligent monitoring and control of electromagnetic relay based on STM32 and LoRa communication[J]. *Information and Computer (Theoretical Edition).* 2023,35(13):55-57.

Authors Introduction

Ms. Suqing Duan

She is currently pursuing her undergraduate degree at the School of Electronic Information and Automation, Tianjin University of Science and Technology. Her research field is embedded system.

Mr. Yuntian Xia

He is currently pursuing his undergraduate degree at the School of Artificial Intelligence, Tianjin University of Science and Technology. His research area is neural networks.

Ms. Siyi Wang
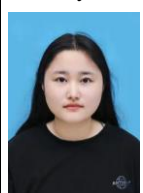
She is currently pursuing her undergraduate degree at the School of Electronic Information and Automation, Tianjin University of Science and Technology. Her research area is automation.

Suqing Duan, Yuntian Xia, Siyi Wang, Yizhun Peng

Dr. Yizhun Pen

He is an Associate Professor in Tianjin University of Science &Technology. He received a doctor's degree in control theory and control engineering from the Institute of Automation, Chinese Academy of Science in 2006. His research field is intelligent robot and intelligent control.

Design of a Fully Automated Logistics Handling Platform

Hongpi Zhao, Jianfeng Qin, Yizhun Peng*

Tianjin University of Science and Technology, Tianjin, 300222, China

*E-mail: *pengyizhun@126.com*

www.tust.edu.cn

Abstract

With the progress of science and technology and the development of society, the world's countries on industrial production efficiency, intelligent manufacturing transformation continues to grow. Therefore, we designed a fully automatic logistics handling platform, which is a multi-modular device based on mechanical design, microcontroller control, visual positioning, the device can realize autonomous identification, autonomous transportation, improve production efficiency, the device has a wide range of applications, can be applied to food packaging, parts processing, intelligent manufacturing, and other automation scenarios.

Keywords: Mechanical design, Circuit design, stepper motor drive, logistics and transportation

1. Introduction

With the development of society and the progress of science and technology, artificial intelligence technology is gradually changing the traditional mode of industrial production [1]. Through the introduction of intelligent manufacturing technology, some advanced products such as unmanned technology, drone courier, intelligent medical care and so on are born. This design is based on the fully automatic logistics handling platform in the context of intelligent manufacturing technology, which effectively solves the problem of time-consuming and labor-consuming manual operation [1], [2], [3], [4].

The rest of this article is organized as follows. the second part introduces the mechanical model and related structural analysis, the third part gives the finite element analysis, and the fourth part shows the related hardware design

2. Mechanical modeling and related structural analysis

Automated item transportation and handling systems are an essential part of the production line in modern factories. Its main purpose is to transport items from one place to another and to perform handling and assembly operations when needed. Such systems usually consist of many different mechanical, electrical and control components to accomplish various tasks. In this product, the transportation and handling of items is mainly achieved by means of a screw slide and rack and pinion mechanism. The main features of this system design are stability and reliability in the face of a variety of items to be transported.

2.1 Design of mechanical parts

The mechanical part of the transportation and handling system of this design is mainly composed of the following parts:
(1) Screw slide: used to transport items from the starting position to the designated position.
(2) Rack and pinion mechanism: used to ensure stability and reliability during transportation.
(3) Pneumatic Cylinder: Used to carry out the transportation and stamping assembly of the articles.
The design of the mechanical part is shown in Fig. 1.
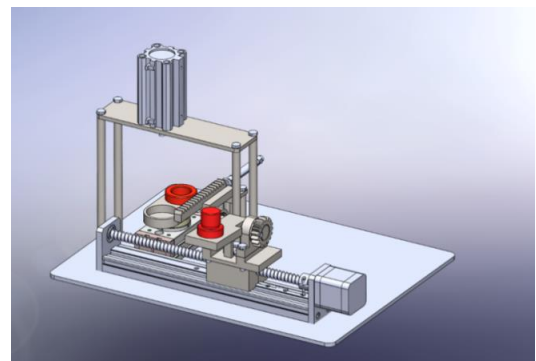
Fig. 1 The Mechanical design drawings

2.2 Component Introduction

(1) Screw slide table: It is a kind of mechanism that utilizes the principle of screw rotation to carry out linear motion, which consists of threaded rod and slider, and pushes the slider to carry out linear motion through the rotation of the threaded rod. Screw slide has the advantages of high precision, good stability and smooth movement, so it is widely used in automated item transportation and handling systems. Screw slide is one of the most important components in the system and is responsible for transporting items from the starting position to the specified position. Screw slides usually consist of a threaded rod, a slider and a guideway. The threaded rod rotates to drive the slide in a linear motion, while the guide rails are used to ensure the orientation and stability of the slide.

(2) Rack and pinion mechanism: it consists of a set of gears, racks and guiding shafts, which is mainly used to realize the overturning and handling of the goods and ensure the stability of the goods in the transportation process. At the same time, the guiding shaft in the rack and pinion mechanism can avoid the lateral displacement of the spiral slider in the transportation process.

(3) Cylinder: It is used to realize the handling and stamping assembly of the articles. When the article reaches the specified position, the cylinder starts to work, clamps the fixture to the article, and moves the article to the stamping assembly area for operation.

The system has a variety of functions, including transportation, handling and stamping assembly of items, etc., and at the same time, it is characterized by stability and reliability, easy operation, safety and high efficiency. It is suitable for various application scenarios in logistics, production and manufacturing industries.

3. Mechanical design

The following focuses on the mechanical design of the handling part of the automated article transportation and handling system.

3.1 Design Idea

The syringe pushes the rack and pinion to mesh with the gear to provide the circumferential force, the gear is connected with the shaft in the handling mechanism through the key to drive the shaft to rotate circumferentially, and at the same time the shaft is connected with the loading fixture in the handling mechanism through the key to realize that the syringe drives the loading fixture to rotate circumferentially to realize the handling of the goods.

3.2 Material selection
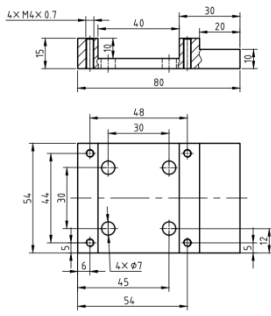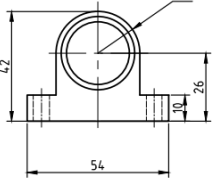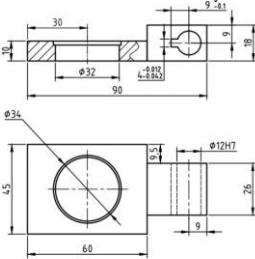
The material selection of the parts of this system is 6063 aluminum alloy, because 6063 aluminum alloy has excellent weldability, extrudability and plating, good corrosion resistance, toughness, plasticity, and can meet

the high requirements of precision machining.

3.3 Manufacturing process

The mechanical parts in this system mainly take two kinds of machining methods: turning and milling, in which turning is mainly used for machining shaft system parts, and milling is used for machining other configurations of parts.

The following table provides a brief parts drawings, drawings include a brief machining accuracy, assembly requirements note. As shown in Table 1.

Table 1 Parts Drawings

Part Name	Summary Drawings
	Machining precision

Part Name	Summary Drawings / Machining precision
Turntable	
Flip Table Stand	
Flip Carrier	

ensure its safety.

Next is the flip shaft, as the most core part, its safety is also the most important part of the flip mechanism. As shown in Fig. 3.

Flap axle

4. Finite Element Analysis and Simulation

Finite element analysis is the use of mathematical approximations to simulate real physical systems. By utilizing simple but interacting elements, it is possible to approximate a real system with a finite number of unknowns to an infinite number of unknowns.

The tipping mechanism is based on mechanical design to analyze the degree of deformation when it is subjected to forces to ensure its safety. Because the material of the device is aluminum alloy 6063, by checking its permissible stress is 160MPa, through solid works finite element analysis, the following diagrams are obtained. As shown in Fig. 2 is the cylinder table.

Fig. 3 Flap axle

From the above figure, it can be seen that the maximum shear stress of this shaft is 2.553×10^7 Pa, while the permissible stress of aluminum alloy 6063 is 1.6×10^8 PA.

Therefore this shaft is safe.

There is also the fit between the rack and pinion, which is the key condition for the flip mechanism to be able to flip. This is shown in Fig. 4.

Fig. 4 Pinion

Fig. 2 Cylinder table

As can be seen from the figure, this design is not safe, although the actual cylinder table is not subjected to such a large force, but it can be added with reinforcement to

As can be seen from the figure, the maximum shear stress is 3.102×10^9 Pa, which is too large compared with the permissible stress, so it is not safe, but it is still feasible in the flip mechanism because it is calculated not

to use such a large force.

Overall, the designed tilting mechanism is completely feasible.

5. Hardware design

To accomplish this, the team designed a PCB motherboard that implements the above features. The PCB schematic diagram is shown in Fig. 5.
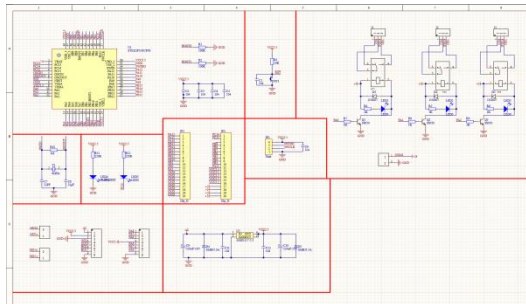
Fig. 5 PCB schematic diagram

The design of this PCB board with STM32F103C8T6 as the main chip, equipped with a stepper motor driver and a three-way relay, realizes precise stepper motor control and solenoid valve on/off. Through programming, the board can control the speed, direction and movement distance of the stepper motor, while it can control the on/off of the circuit through the relays to realize automation control.

5.1 Solenoid Valve Drive:

The Relay is shown in Fig. 6.
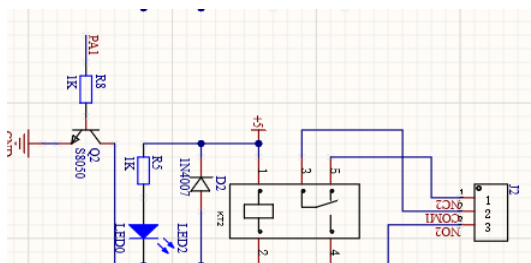
Fig. 6 Relay

Through the main control board to send high and low level signals to control the relay on and off, and then control the solenoid valve on and off.

5.2 Stepping motor drive:

This design mainly adopts two stepper motor driving methods, the first one is stepper motor driver and the second one is TB6612FNG stepper motor driver module [4], [5], as shown in Fig. 7.
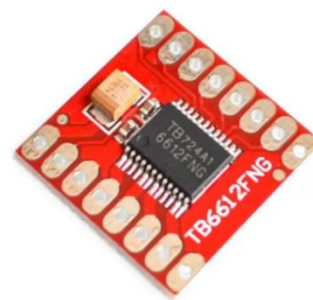
Fig.7 Stepper Motor Driver and TB6612FNG Stepper Motor Driver Module

The actual principle of a stepper motor driver is to convert a pulse signal from a control system into an angular displacement of the stepper motor. Whenever the stepper driver receives a pulse signal, it drives the stepper motor to rotate in a set direction at a fixed angle (called the "step angle"). The stepper motor rotates step by step at a fixed angle. The angular displacement can be controlled by controlling the number of pulses for precise positioning. The speed and acceleration of the motor rotation can also be controlled by controlling the pulse frequency for speed regulation [5], [6], [7], [8], [9].

5.3 Voltage regulator module

This design voltage regulator module mainly uses the chip for AMS117-3.3V for 3.3v voltage regulator design.

6. Product Summary

Through the test, the product can complete the recognition of objects and achieve the desired position to achieve flipping and stamping through the precise control of the stepper motor, which can be adapted to the automatic scenes such as product packaging and parts processing.

References

1. Design and research on control system of handling robot in automation perspective[J]. ZHANG Qiang;PAN

Zhendong;TAO Chenghu; CAO Zhenbo. Science and Technology Innovation Herald,2020

2. A practical design of variable interpolation drive controller for stepper motor[J]. LI Lingjuan; LIU Jinglin;WANG Can. Microtronic Motor,2008

3. Stepper Motor Speed Control Method Based on single chip Microcomputer [J]. Yu Jianguo. Electric Times,2006

4. Overview of stepper motor [J]. Li Lihua; Zhang Yongli. Private Science and Technology,2008

5. Cause Analysis and Solution of stepping motor loss [J]. Xie Min. New Technology and New Process,2008

6. New stepping motor [J]. Wang Cailong. Micro Electric Machine,1984

7. Several Research on Mechanical Design Technology in the Era of Intelligent Manufacturing [J]. Gao Songlin. Hubei Agricultural Mechanization,2019

8. Preliminary Study on Design Problems of stepper Motor Driver [J]. Tuo Chaoyong. Hunan Agricultural Machinery,2010

9. Development and Research of Small Capacity Stepper Motor Intelligent Driver [J]. Jin Hailong. Automation and Instrumentation,2000

Authors Introduction

Mr. Hongpi Zhao

He is currently pursuing his undergraduate degree at the School of Electronic Information and Automation, Tianjin University of Science and Technology.

Mr. Jianfeng Qin

He is an Associate Professor in Tianjin University of Science &Technology. He received a doctor's degree in control theory and control engineering from the Institute of Automation, Chinese Academy of Science, in 2006.His research field is intelligent robot and intelligent control.

Dr. Yizhun Pen

He is an Associate Professor in Tianjin University of Science &Technology. He received a doctor's degree in control theory and control engineering from the Institute of Automation, Chinese Academy of Science in 2006. His research field is intelligent robot and intelligent control.

Design and Application of AI-based Brush Calligraphy and Painting Robot

Haibo Li, Yizhe Sun, Shuxin Wang
College of Artificial Intelligence, Tianjin University of Science and Technology, Tianjin, China
College of Electronic Information and Automation, Tianjin University of Science and Technology, Tianjin, China

Yizhun Peng[*]
College of Electronic Information and Automation, Tianjin University of Science and Technology, Tianjin, China
Email: pengyizhun@tust.edu.cn

Abstract

With the rapid development of digital technology and artificial intelligence, the innovation and exploration of traditional painting forms in the field of digital art have become increasingly captivating. The focus of this research is to create a brush calligraphy and painting robot system based on PLT files. It digitizes user creative instructions to autonomously generate brush calligraphy and landscape paintings. This technology not only advances the deep integration of digital art and traditional culture but also opens new perspectives and vast possibilities for artistic creation.

Keywords: Artificial intelligence, Calligraphy and painting robot, Brush word, Landscape painting, Digital art

Introduction

With the rapid progress of digital technology [1] and artificial intelligence, traditional art forms [2] are experiencing a profound digital change, and brush painting and calligraphy as a precious heritage of traditional Chinese culture is no exception. This study aims to further explore the influence of digital technology on traditional painting, and explore the widespread application of this technology in the field of art creation. The core concept of the brush calligraphy and painting robot system is planted on the basis of the PLT file. It receives the user's creative instructions in a digital manner and aims to independently draws brush characters and landscape painting. The rise of artificial intelligence provides strong technical support for the design of brush calligraphy and painting robots. The algorithm of machine learning such as deep learning enables the machine system to imitate the unique brush strokes of the brush, and even understand the emotion and artistic conception expressed by the artist in landscape painting. This technology not only injects new vitality into traditional art, but also expands the boundaries of art creation, and provides users with more diversified and innovative creative methods. In this study, we will introduce the technical principles, design frameworks and application scenarios of the brush calligraphy and painting robot system. Through in -depth excavation of these aspects, we expect to provide in -depth insights for the integration of digital art and traditional culture, and contribute to the innovation and inheritance of traditional culture of digital art.

1. The main text

The design scheme uses a robotic arm, a set of linear guide systems, as well as painting tools such as pen, ink, paper, and inkstone. The robotic arm is mainly responsible for grasping the brush and realizing dipping ink, writing and painting [3]. The robotic arm is cleverly placed on the linear guide to expand its movement range and improve accuracy and efficiency [4]. For the overall design effect, see Fig. 1.

Fig.1 overall arrangement

2. Core structure of robotics

The robotic arm chooses the desktop four-axis robotic arm of Yuejiang Technology. As shown in Fig. 2, the robotic arm works radius of 320mm, the load is 0.5kg, the repeated positioning accuracy is 0.2mm, and it has high accuracy in the vertical direction. And in the horizontal direction, it can be moved quickly, with a light weight and a small volume. The length and width of the base are 158mm. You can complete the writing or painting work of this robot. The first

freedom of the robotic arm is the rotation freedom composed of a shaft base. The second and third degree of freedom is the rotation freedom composed of large arms and small arms. Four freedom is located at the end of the robotic arm, which is rotating freedom [5]. You can install a fixture or suction cup at this location. The robot chooses a fixture to pinch the brush.

Fig. 2 Manipulator

The core parameters are shown in Table 1.

Table 1. The core parameters of the robotic arm.

Weight	3.4 kg
Base Dimension (Footprint)	158 mm x 158 mm
Materials	Aluminum Alloy 6061, ABS Engineering Plastic
Controller	Dobot Integrated Controller
Robot Mounting	Desktop
Number of Axes	4
Packing Size (L X w X H)	421 mm x 334 mm x 352 mm
Payload	500g
Max.Reach	320mm
Position Repeatability	0.2mm
Communication	USB\WIFI*\Bluetooth*
Power Supply	100-240 V,50/60 Hz
Power In	12V/6.5 A DC

3. The innovation point

The convolutional neural network (CNN) [6] is usually used in tasks such as image recognition and feature extraction, and the PLT file is a format for describing the graphic drawing instruction sequence. In the design of the brush calligraphy and painting robot, the two can be used with each other.

First, convolutional neural networks can be used to generate training of PLT files. The robot can draw a specific pattern or imitate a certain style. It can use convolutional neural networks to train images, and then use the characteristics to generate the corresponding PLT files. CNN can help identify the characteristics of strokes, lines, and shapes in the image, thereby generating the corresponding drawing instructions. Fig. 3-1 is the "chrysanthemum picture" entered by the user, Fig. 3-2 is the picture form in the PLT file generated after program processing.

Fig. 3-1 Original Picture   Fig. 3-2 Processed pictures

Secondly, convolutional neural networks can be used for image recognition and PLT file generation. When the robot receives the image input provided by the user, the convolutional neural network can be used for image recognition and identify the content and characteristics of the image [7]. Then, the corresponding PLT file is generated according to the recognition result, so as to guide the robot to complete the drawing task. Fig. 4 is the calligraphy of the robot

Fig. 4 the calligraphy

3.1. The use of PLT files

The PLT file provides a digital approach for describing painting instructions. By using PLT files, you can digitize traditional paintings and landscape painting such as landscape painting to make it meet the needs of modern digital art.

The PLT format is a vector graphic format, also known as HPGL (Hewlett Packard Graphics Language) format. It is a drawing language developed by HP, which was originally used to control the output of the drawer. This format is now widely used in fields such as drawing, engineering and manufacturing.

The PLT file is a text file that can use any text editor to open the viewing and editing. It contains drawing commands and coordinates, which can be drawn through these commands.

3.2. Convolutional neural network

Convolutional neural network (CNN) is a feeding neural network structure that is designed to process large image data and performs well in the fields of image and voice recognition. The network structure consists of one or more convolutional layers, full connection layers, pooling layers, and linear rectifier layers. It has the ability to capture the two -dimensional structure of the input data, especially in the field of image processing.The components of convolutional neural networks include:

(1) Convoiced layer: The convolutional layer can generate a set of parallel feature maps, and performs operations by sliding the convolution nuclear on the input image. The convolution nucleus and the input image perform the computing of the component of the element and the computing of the element. This process is called the rod. The size of the convolution nucleus is small. Enter the image through overlapping or parallel effects to share the same weight and bias item to effectively extract image features.

(2) Linear rectifier: Linear rectifier uses linear rectifier (RELU) as an incentive function to enhance the non -linear characteristics of the entire neural network without changing the output of the convolution layer.

(3) Pondalization layer: Pondization is a non -linear form of downgrade, of which the largest pooling is a common form. This layer divides the input image into a rectangular area, outputs the maximum value of each sub -area to achieve the purpose of reducing the data dimension.

(4) Complete connection layer: After the convolution and pooling layer, the neural network is high -level reasoning through the complete connection layer. The neurons in the complete connection layer are connected to all the activation in the previous layer, and activation is calculated through imitation transformation, including the addition of matrix multiplication and bias items.

4. Conclusion

This research focuses on the design and development of a brush calligraphy and painting robot system based on PLT files. By digitizing user creative instructions, the system autonomously generates brush calligraphy and landscape paintings. The rise of artificial intelligence and convolutional neural networks provides robust support for the robot's design, enabling it to mimic unique brush strokes and comprehend the emotions expressed by artists in landscape paintings. This technology not only deepens the integration of digital art and traditional culture but also explores new avenues and vast possibilities for artistic creation.

5. Future outlook

The design and application of brush calligraphy and painting robots have opened a new era of deep integration of digital art and traditional culture. In the future, the development of brush calligraphy and painting robots will inject new vitality into the field of digital art, while promoting traditional culture in the digital age. The continuous evolution of this technology will bring more possibilities to artists, educators, and cultural and creative industries, and bring a more colorful future to our art world.

Acknowledgements

References

1. Li Pengyue. The study of the four -freedom robotic arm drawing vector character and graphics [D]. Taiyuan University of Technology, 2018.
2. Zhang Yu. The application of traditional Chinese visual elements in digital media art design [J]. Toys world, 2023, (06): 82-84.
3. Wang Miao. Innovation research in digital media art in the era of artificial intelligence [J]. Toys World, 2023, (06): 8-10.
4. Ding Hao. Study on image recognition and processing algorithm based on neural networks [C] // Tianjin Electronic Society. Conference paper set. Tianjin City Investment Group Asset Management Co., Ltd.;, 2023: 3.Doi: 10.26914/C.CNKIHY.2023.022845
5. Feng Kaili. Research on image dehuminating methods based on convolutional neural networks [D]. Shandong NormalUniversity,2023.DOI:10.27280/d.cnki.gsdsu.2023.000567
6. Lin Qianqi, Wang Yibo, Zhan Qiucheng, etc. The four-freedom mechanical arm sports algorithm and steering gear slowdown design [J]. Mechanical engineer, 2023, (11): 64-66.
7. Lu Ran. Six -axis calligraphy and painting robot trajectory optimization control method Study [D]. Taiyuan University of Science and Technology, 2021.DOI: 10.27721/d.cnki.Gyzjc.2020.000297

Authors Introduction

Mr. Haibo Li

He is currently pursuing his undergraduate degree at College of Artificial Intelligence, Tianjin University of Science and Technology. His research field is to Reinforcement learning and intelligent robots.

Mr. Yizhe Sun

He is currently pursuing his undergraduate degree at College of Artificial Intelligence, Tianjin University of Science and Technology. His research field is to ROS and intelligent robots.

Mr. Shuxin Wang

He is currently pursuing his undergraduate degree at College of Artificial Intelligence, Tianjin University of Science and Technology. His research interests are image processing and neural networks.

Dr. Yizhun Pen

He is an Associate Professor in Tianjin University of Science &Technology. He received a doctor's degree in control theory and control engineering from the Institute of Automation, Chinese Academy of Science in 2006. His research field is intelligent robot and intelligent control.

Intelligent Wheelchair System：Non-contact Heart Rate and Body Temperature Measurement

Dongpo Ma

Tianjin University of Science and Technology, 1038 Dagu South Road, Hexi District, Tianjin, 300222, China

Junsheng Zhang

Tianjin University of Science and Technology, 1038 Dagu South Road, Hexi District, Tianjin, 300222, China

Yizhun Peng[*]

Tianjin University of Science and Technology, 1038 Dagu South Road, Hexi District, Tianjin, 300222, China

Email: pengyizhun@126.com

Abstract

With the continuous development of society and human civilization, people, especially the disabled and the elderly, need more and more to use modern high technology to improve their lives and improve the comfort of life. Therefore, this paper introduces a single-chip microcomputer-based non-contact body temperature and facial recognition-based heart rate measurement method and applies it to the intelligent wheelchair system. Monitor the vital signs of the elderly, and give early warnings in time when abnormal values occur, so that family members or caregivers can rescue them in time.

Keywords: Smart wheelchairs，Non-contact，heart rate，Body temperature measurement

1. Introduction

Population aging has become a global phenomenon. China's elderly population is huge, the aging rate is fast, and the problem of elderly care and daily care for the elderly is becoming increasingly severe. Wheelchairs are indispensable means of daily life for the elderly and people with lower limbs, and a cost-effective, comprehensive and reliable intelligent wheelchair can not only improve the quality of life of the disabled and the elderly, but also reduce the burden on caregivers. Smart wheelchairs. Research plays an important role in the development of elderly care projects in China [1]. Since the development of the first intelligent wheelchair in the United Kingdom in the 80s of the 20th century, the United States, Japan, France and other countries have also carried out research on intelligent wheelchairs [2], [3], [4], [5], [6]. The research on intelligent wheelchairs in China started late, but in recent years, some research results have been achieved. The research of intelligent wheelchairs based on visual tracking [7], autonomous driving [8], electromyography [9], head [10], multi-sensor fusion [11], and voice [12] has also been progressing, but these wheelchairs are still in the laboratory stage, which are expensive, and their safety and real-time performance cannot meet the demand. Therefore, this study takes the above problems as the starting point, and carries out the system improvement design on the basis of the existing wheelchair machinery to meet the current needs.

2. Infrared body temperature measurement

Based on the stm32 microcontroller, the B-1 infrared temperature measurement module is used to complete the non-contact body temperature measurement. The physical diagram is shown in Fig. 1.
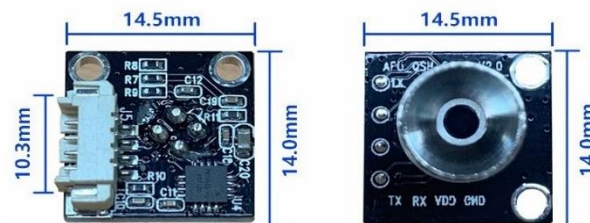
Fig. 1 Photographs of the B-1

The module has a built-in 24-bit high-precision AD conversion chip, and its working principle is that when the object is higher than absolute zero (-273℃), it will radiate infrared energy outward, and the infrared sensor receives the infrared light reflected back by the object, and obtains the induced temperature data through the high-precision program algorithm in the CPU.

Table 1 Temperature measurement results

Temperature	1	2	3	4
Measure the temperature/℃	36.7	36.2	37.3	35.9
True temperature/℃	36.8	36.0	37.1	36.2

3. Non-contact heart rate measurement

3.1 Technical principle

There are two commonly used methods to measure heart rate in clinical practice: electrocardiogram (ECG) and optical volume scanner (PPG) PPG is a method of detecting blood volume pulses (BVP) in human blood tissues using photoelectric means, such as fingertip pulse oximeters, exercise bracelets, etc

Remote photoplethysmography (RPPG): Diffuse and specular reflections of light after passing through the skin, both of which can be captured by the camera. According to Lambert-Beer's law, the strength of a substance's absorption of a monochromatic light is related to the concentration of the absorbent substance and the thickness of its liquid layer. Since the periodic beating of the heart causes the same periodic change in the concentration of blood in the blood vessels and capillaries, it leads to a periodic change in the concentration of light absorbed by the blood, so that the intensity of the diffuse reflected light received also changes periodically, that is, the pulse wave information.

When the heart contracts, the concentration of blood increases, the human skin absorbs more light and reflects less light; When the heart stretches, the concentration of blood decreases, the human skin absorbs less light, and the reflected light increases [13]. The schematic diagram is shown in Fig. 2.
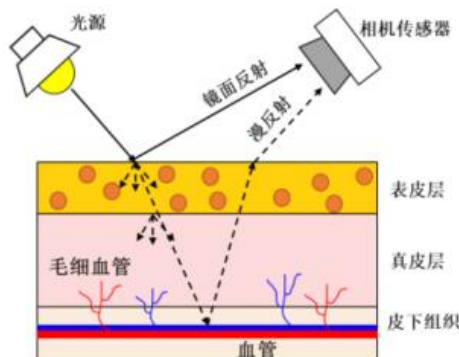
Fig. 2 Schematic diagram of the principle

3.2 Data set

This experiment uses a UBFC dataset of a total of 10 RGB videos, recorded in an indoor environment, including light changes and head movements, with a pixel resolution of 640x480. At the same time, the fingertip oximeter was used to synchronously record the real BVP signal and heart rate value as label data.

3.3 Test results

According to the distribution of the Bland-Altman diagram, the blue scatters are basically within the red

dotted line and around the black dashed line. According to the distribution of the Fig. 3 scatter plot, the distribution of the sample points of the predicted value and the true value is basically around the straight line y=x, which indicates that the consistency between the predicted value and the true value is good. Fig. 4 shows that the experimental consistency is good
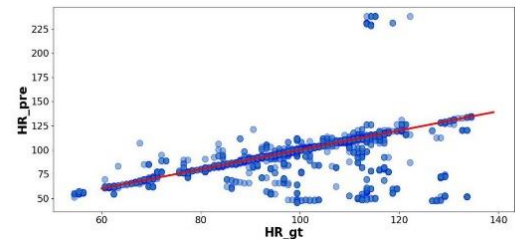
Fig. 3 Scatter plot
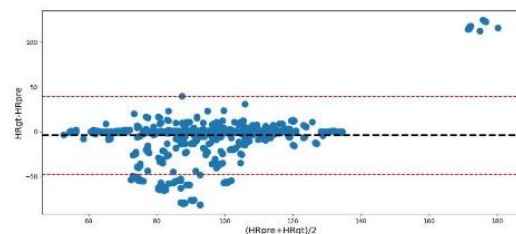
Fig. 4 Bland-Altman

4 Experimental analysis
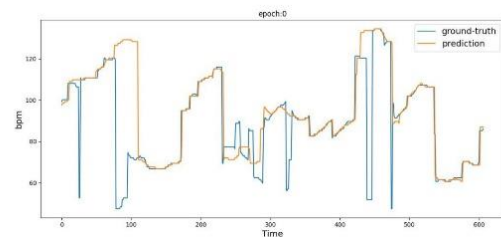
Fig. 5 Comparison chart

Fig. 5 shows that the true value curve is in good agreement with the predicted value curve. Table 1 shows that the error between the measured and true values of body temperature is within acceptable limits.

5 Conclusion

In addition to the non-contact heart rate and body temperature measurements presented in this article, a smart wheelchair system is a highly intelligent device that integrates a variety of sensors, controllers, and actuators to help improve the quality of life for the elderly and people with disabilities with reduced mobility. At the same time,

the human-computer interaction module enables users to easily communicate with the wheelchair. In the future, with the continuous development of technology, smart wheelchairs will be able to provide convenient and safe travel solutions for more people in need.

References

1. Zhen Zhang, Xinmei Liu, Junling Yin ,etc. Design of ARM-based intelligent wheelchair safety monitoring system [J]. Foreign electronic measurement technology, 2021,40(6):138-143.
2. Lamia B, Noureddine E. Implementation of a biometric interface in voice controlled wheelchairs[J]. Sound and Vibration, 2020, 54(1):1-15.
3. Pawuś D, Paszkiel S. BCI wheelchair control using expert system classifying EEG signals based on power spectrum estimation and nervous tics detection[J]. Applied Sciences, 2022,12(20):1-22.
4. Rosero P, Peluffo D, Batista V, et al. Intelligent system for identification of wheelchair user's posture using machine learning techniques [J]. IEEE Sensors Journal, 2019, 19 (5):1-8.
5. Abolfazl M. Human-robot interaction in rehabilitation and assistance: a review[J]. Current robotics reports,2020, 1(3) : 131-144.
6. Timofei I V, Eugene V C, Gleb A U. The control system based on extended BCI for a robotic wheelchair[J]. Procedia Computer Science,2018,123: 522-527.
7. Yan Li, Mingda Tang, Yingling Zhou, etc. Computer vision-based wheelchair following system design [J]. Computer Engineering and Applications,2021,57(10):163-172.
8. Chongyue Bai, Jianjun Wang, Xiaoxiao Cheng, etc. Fusion laser SLAM realizes the optimization of spatial positioning of unmanned wheelchairs [J]. Advances in Lasers and Optoelectronics, 2022, 59(2):485-493.
9. Weijian Yang, Qingshan She, Qizhong Zhang, etc. Intelligent wheelchair control method for shoulder and neck electromyography signal recognition [J]. Journal of Shanghai Jiao Tong University,2016,50(6):949- 956, 962.
10. Cong Wu, Kun Zhang, Licai Yang. Design of attitude monitoring device for the elderly based on inertial sensor [J]. Biomedical engineering research,2018,37(2):210-214.
11. Zhixin Han, Xiuwu Sui. Research on intelligent control technology of myoelectric wheelchair based on multi-source information fusion [J]. Modern Manufacturing Engineering,2020(9):136-144.
12. Jie Hu, Hongliu Yu, Ping Shi, etc. A kind of multifunctional intelligent wheelchair control system design [J]. Chinese Journal of Rehabilitation Medicine,2016,31(11):1246-1249.
13. Weixiang Gao, Qing Li, Yuzhu Yang. Non-contact heart rate detection based on PTZ camera [J]. Computer applications and software,2023,40(09):43-50+98.
14. Weihua O, Longbao C, Jiahao X, et al.Non-contact heart rate measurement from face video based on group sparse representation[J].Digital Signal Processing,2024,144

Authors Introduction

Dongpo Ma

He received his bachelor's degree in physics and optoelectronic engineering from Taiyuan University of Technology in 2019 and is now a graduate student in Tianjin University of Science and Technology

Mr. Junsheng Zhang

In 2021, he will be admitted to the School of Electronic Information and Automation, Tianjin University of Science and Technology, China, majoring in electronic information. He is currently a master's student at Tianjin University of Science and Technology, China

Dr. Yizhun Pen

He is an Associate Professor in Tianjin University of Science &Technology. He received a doctor's degree in control theory and control engineering from the Institute of Automation, Chinese Academy of Science in 2006. His research field is intelligent robot and intelligent control.

Recognition and Localization Method for Automotive Axle Holes in Assembly Robots

Junsheng Zhang

Tianjin University of Science and Technology, 1038 Dagu South Road, Hexi District, Tianjin, 300222, China

Dongpo Ma

Tianjin University of Science and Technology, 1038 Dagu South Road, Hexi District, Tianjin, 300222, China

Yizhun Peng[*]

Tianjin University of Science and Technology, 1038 Dagu South Road, Hexi District, Tianjin, 300222, China

Email: pengyizhun@126.com

Abstract

During the production of automotive axle holes, the roundness error at the pipe opening leads to low detection efficiency due to manual measurements, rendering real-time inspection unfeasible. This paper proposes a method for detecting and sorting the roundness of automotive axle holes based on visual inspection. Ensuring accuracy in grasping, it establishes the kinematic model of the robot Aubo_I5. Targeting automotive axle holes for grasping, it employs adaptive threshold segmentation to highlight the section features of the axle hole. The Canny algorithm is then used to extract edge information, and finally, the least squares method is utilized to detect roundness errors for sorting the axle holes based on this error.

Keywords: Axle Hole, Assembly, Vision Deep, Learning

1. Introduction

The assembly of mechanical components constitutes the subsequent phase in the entire manufacturing process. Manual assembly is characterized by low efficiency, and the assembly speed is affected by factors such as worker labor intensity. Traditional assembly robots have mostly been controlled through teaching or offline programming [1], [2]. Consequently, these robots strictly follow predefined assembly procedures, lacking the ability to promptly respond to changes in the surrounding environment or objectives, thus exhibiting poor robustness. With the advancements in computer and visual technologies, automation machinery on production lines gradually possesses visual perception capabilities. This enables real-time dynamic recognition and localization of targets, facilitating visual-guided adjustments of robot poses and grasping [3].

The emergence of deep learning has improved issues regarding detection accuracy in visual recognition. However, due to the complexity of the assembly process, uncertainties such as lighting conditions, backgrounds, workpiece sizes, shapes, and other factors in the working environment are critical issues affecting target recognition and localization.

2. Kinematic Modeling

Establish the joint coordinate system of the Aubo_I5 robot as shown in Fig. 1, and the structural model and dimensions of each joint as illustrated in Fig. 2. To more accurately define the sorting model coordinate system, the gripper coordinate system G_{xy} is established at the end joint.
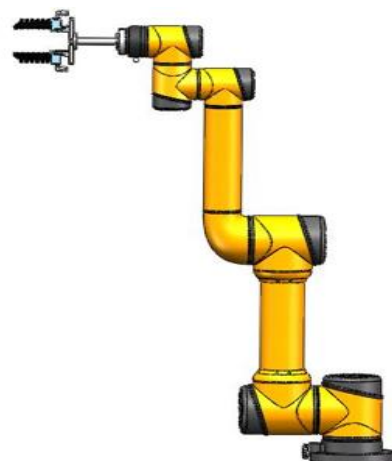
Fig. 1 Structure of the Aubo_I5 Robot

Junsheng Zhang, Dongpo Ma, Yizhun Peng
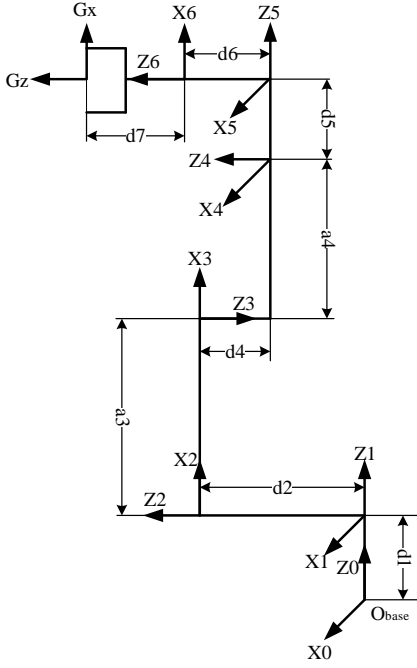
Fig. 2 Robot Joint Coordinate System

Fig 2: Establishing the Joint Coordinate System of Aubo_I5

Based on the joint coordinate system, the corresponding D-H parameter table is shown in Table 1 [4], where each link parameter is defined as follows:

α i-1 represents the distance moved along Xi-1 axis from Zi-1 to Z_i;

ai-1 represents the angle rotated around Xi-1 axis from Zi-1 to Z_i;

d_i represents the distance moved along Z_i axis from Xi-1 to X_i;

θ _i represents the angle rotated around Z_i axis from Xi-1 to X_i.

Table 1 D-H Parameter Table for AUBO_I5 Robot

i	α_{i-1}/deg	a_{i-1}/mm	d_i/mm	θ_i/deg
1	0	0	$d_1 = 98.5$	θ_1
2	90°	0	$d_2 = 140.5$	θ_2
3	180°	$a_3 = 408$	0	θ_3
4	180°	$a_4 = 376$	$d_4 = -19$	θ_4
5	−90°	0	$d_5 = 102.5$	θ_5
6	90°	0	$d_6 = 94$	θ_6

2. 1 Robot Forward Kinematics Analysis

Forward kinematics analysis involves determining the

pose relationship between the tool coordinate system of a robot relative to the polar coordinate system, given the known motion angles of each joint. This is accomplished using the homogeneous transformation matrix $^{i-1}_iT$ between adjacent link pairs, as shown in Equation (1):

$$^{i-1}_iT = \begin{bmatrix} C\theta_i & -S\theta_i & 0 & a_{i-1} \\ S\theta_iC\alpha_{i-1} & C\theta_ic\alpha_{i-1} & -S\alpha_{i-1} & -S\alpha_{i-1}d_i \\ S\theta_iS\alpha_{i-1} & C\theta_iS\alpha_{i-1} & C\alpha_{i-1} & C\alpha_{i-1}d_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

In the equation：$C\theta_i = \cos\theta_i, S\theta_i = \sin\theta_i$

According to the D-H parameter table of Aubo_I5, calculate the homogeneous transformation matrices for each of the seven joints ^0_1T、^1_2T、^2_3T、^3_4T、^4_5T、^5_6T、^6_7T As shown below：

$$^0_1T = \begin{bmatrix} C1 & -S1 & 0 & 0 \\ S1 & C1 & 0 & 0 \\ 0 & 0 & 1 & d_1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2)$$

$$^1_2T = \begin{bmatrix} C2 & -S2 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ S2 & C2 & 0 & d_2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3)$$

$$^2_3T = \begin{bmatrix} C3 & -S3 & 0 & a_3 \\ -S3 & -C3 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4)$$

$$^3_4T = \begin{bmatrix} C4 & -S4 & 0 & a_4 \\ -S4 & -C4 & 0 & 0 \\ 0 & 0 & -1 & d_4 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5)$$

$$^4_5T = \begin{bmatrix} C5 & -S5 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -S5 & -C5 & 0 & d_5 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (6)$$

$$^5_6T = \begin{bmatrix} C6 & -S6 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ S6 & C6 & 0 & d_6 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (7)$$

Combining the above formulas, the robot's forward kinematics equation can be obtained by simultaneous.

$$^0_6T = {}^0_1T{}^1_2T{}^2_3T{}^3_4T{}^4_5T{}^5_6T = \begin{bmatrix} n_x & o_x & a_x & p_x \\ n_y & o_y & a_y & p_y \\ n_z & o_z & a_z & p_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (8)$$

2.2 Robot Inverse Kinematics Analysis

Inverse kinematics involves calculating the joint angles θ_i of a robot given the pose of the tool coordinate system. Considering the robot's link coordinate systems, the axes of joints 4, 5, and 6 intersect at one point, satisfying the Pieper criterion, indicating the existence of a closed-form

solution. Assuming the end pose ${}^{0}_{6}T$ in Equation (8) is known, solving is performed using an analytical method, and the inverse solutions for each joint are presented in Table 2.

Table 2: Inverse solution formula of joint

$$\theta_1 = arctan \frac{u_1}{\pm\sqrt{1 - u_1{}^2}} - arctan \frac{p_y}{p_x}$$

$$\theta_5 = arctan \frac{\pm\sqrt{1 - u_2{}^2}}{u_2}$$

$$\theta_6 = arctan \frac{o_x S_1 + o_y C_1}{n_y C_1 + n_x S_1}$$

$$\theta_3 = arctan \frac{\pm\sqrt{1 - g^2}}{g}$$

$$\theta_2 = arctan \frac{k}{\pm\sqrt{1 - k^2}} - arctan \frac{u_4}{v_4}$$

$$\theta_4 = arctan \frac{v_5}{u_5} - \theta_3 + \theta_2$$

Among which:

$$u_1 = \frac{d_4 - d_5}{\sqrt{p_x{}^2 + p_y{}^2}}$$

$$u_2 = -a_x S_1 - a_y C_1$$

$$u_3 = C_1(p_x + d_6 o_x - d_5 C_6 a_x - S_6 n_x) + S_1(d_5 C_6 a_y - p_y - d_6 o_y - S_6 n_y)$$

$$g = \frac{u_3{}^2 + v_3{}^2 - a_3{}^2 - a_4{}^2}{2a_3 a_4}$$

$$u_4 = a_3 + a_4 C_3 - a_4 S_3$$

$$v_4 = a_3 + a_4 S_3 + a_4 C_3$$

$$k = \frac{u_4 + v_4}{\sqrt{u_4{}^2 + v_4{}^2}}$$

$$u_5 = C_1(o_x S_5 + n_x C_5 C_6 + a_x S_6) - S_1(o_y S_5 + n_y C_5 C_6 + a_y S_6)$$

$$v_5 = o_z S_5 + C_5(n_z C_6 + a_z S_6)$$

2.3 Mage Preprocessing

Before conducting roundness measurement, in order to eliminate the effects of metal debris, noise from the surrounding environment, and surface wear on the workpiece, the captured images need preprocessing. This involves removing noise and the influence of lighting while retaining the edge features of the workpiece [5]. Image preprocessing mainly includes:

1）Image Denoising: Using specific filtering algorithms to reduce noise in the image while preserving edge features.

2）Threshold Segmentation: Highlighting the target contour features more effectively.

(1) Image Filtering

Image filtering refers to the suppression of noise in an image while preserving detailed features, and its effectiveness directly impacts subsequent image processing. Common filtering algorithms include mean filtering, median filtering, Gaussian filtering, and others.

1）Mean filtering is a typical linear filtering algorithm that replaces the grayscale value of a particular pixel with the average pixel value in its surrounding area. Its mathematical expression is as follows:

$$g(x, y) = \frac{1}{mn} \sum_{(i,j) \in S_{xy}} f(i, j) \qquad (9)$$

Where represents the center point at, and is the size of the filtering window. By controlling the window size, the filtering effect can be adjusted. A larger value for results in a more pronounced filtering effect.

2）The median filtering algorithm is a common non-linear filter that calculates the median value of points within a neighborhood range in an image for a particular pixel, replacing the original pixel value. The mathematical expression is as follows:

$$g(x, y) = median_{(i,j) \in S_{xy}} \{f(i, j)\} \qquad (10)$$

Median filtering controls the size of to adjust the filtering effect. Compared to mean filtering, median filtering not only removes salt-and-pepper noise but also better preserves edge information.

3) Gaussian filtering is also a type of linear smoothing filter. It involves the weighted averaging process of the entire image and is primarily used to remove uneven lighting or sensor noise.

To compare the effects of the three filtering methods, experiments were conducted on the images, and the results are shown in Fig. 3.

From the results, it's apparent that after applying median filtering, not only has the original noise in the image and the interference from oil stains on the workpiece surface reduced, but it has also effectively preserved the edge features. Hence, median filtering is employed to filter the images of the pipe end face.

(2) Threshold Segmentation

Image threshold segmentation is a crucial step in segmenting the target region. Determining the threshold is a key aspect of image segmentation. If the threshold is set too high, target points may be considered as background

Junsheng Zhang, Dongpo Ma, Yizhun Peng

and ignored, while setting it too low might lead to background points being identified as target features. Common threshold segmentation methods include adaptive thresholding algorithms, Otsu's method, iterative threshold segmentation, and others. Due to varying surface qualities of the hole, there's no fixed threshold selection. Therefore, this paper conducts comparative experiments among several commonly used segmentation algorithms to choose the appropriate threshold segmentation algorithm suitable for the cooling pipe imaging environment [6].

(a)Mean Filtering

(b)Median Filtering

(c)Gaussian Filtering

Fig. 3 Comparison of filtering algs

After applying median filtering to the captured borehole regions, comparative experiments of three segmentation algorithms were conducted, and the results are shown in Fig. 4.

From the segmentation results above, it's evident that the adaptive algorithm has better segmentation results, preserving the edge features of the pipe end more completely. Therefore, this paper adopts the adaptive segmentation algorithm.

From the segmentation results above, it's evident that the adaptive algorithm has better segmentation results, preserving the edge features of the pipe end more completely. Therefore, this paper adopts the adaptive segmentation algorithm.

(a)Adaptive Algorithm

(b)OTSU Algorithm

(c)Iterative Algorithm

Fig. 4 Comparison of threshold segmentation

2.4 Edge Detection

After threshold segmentation, the next step is edge detection for the borehole. To ensure better contour detection, reducing noise interference is essential to maintain the integrity and clarity of edge information. Commonly used edge detection algorithms include the Canny algorithm, Sobel algorithm, Laplacian algorithm, etc. Compared to other edge detection algorithms, the Canny algorithm is less susceptible to noise interference [7].

The Canny algorithm first applies a Gaussian filter to smooth the image, reducing its sensitivity to noise. Then, it calculates the gradients in the X and Y directions using the Sobel operator, performs non-maximum suppression, and finally uses dual thresholds to determine edges. The Canny edge detection effect for the port block is shown in Fig. 5.

Fig. 5 Canny detection algs

2.5 Borehole Circular Detection

After edge extraction, detecting circles is a crucial part of roundness measurement. The most commonly used methods for circular detection are the Least Squares Circle Fitting and Hough Transform [8]. To measure circles accurately, it's essential to filter out the circular contours for measurement. As shown in the image, after pre-processing the image, the contour of the borehole is extracted using Canny edge detection. The contour points of the borehole are then sorted for subsequent roundness measurement.

Using the detected circle center as the reference [9], [10], the roundness error is represented by the difference between the maximum contour circle radius Rmax and the minimum contour circle radius Rmin, expressed as Rerr= Rmax − Rmin.

$$erro = \sqrt{(x_{max} - x_c)^2 + (y_{max} - y_c)^2} - \sqrt{(x_{min} - x_c)^2 + (y_{min} - y_c)^2} \quad (11)$$

among which : $(x_{max}, y_{max}), (x_{min}, y_{min})$ The coordinates of the farthest and nearest points from the fitted circle center, respectively, correspond to the contour.

The principle of the least squares fitting circle involves several steps: 1) detecting the actual contour, 2) applying the least squares fitting to find a circle, 3) creating a circle with the least squares circle center and the maximum distance to the contour, and 4) creating a circle with the least squares circle center and the minimum distance to the contour. The difference between the maximum and minimum radii of these circles is used as the basis for evaluating roundness.

The circle fitted using the least squares circle algorithm for the axle hole contour is depicted in Fig. 6. The coordinates of the fitted circle's center are (232.87733, 85.63195), and the fitted circle's radius is R = 63.36658.3
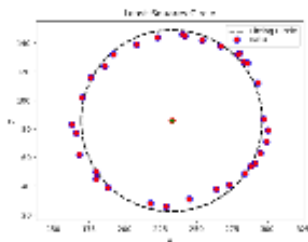
Fig. 6 Least squares circle fitting

3. Target recognition network based on YoloV4

The YoloV4 algorithm can be divided into four parts: input end, backbone feature extraction network, enhanced feature extraction network and output end. The input end includes image preprocessing, converting the size of the input image to the specified input size, and normalizing the image information. Secondly, it is the backbone feature extraction network. CSPDarknet53 formed after YoloV4 is added to CSPNet serves as the backbone feature extraction network. YoloV4 's enhanced feature extraction network consists of a PANet structure, which is used to improve the diversity of feature expressions and strengthen the fusion of feature information; the last is the output end, which is used to perform classification and regression tasks and output the final prediction results.

YoloV4 has three outputs after passing through the backbone feature extraction network, namely L3, L4 and L5. Among them, L3 and L4 will be output to the enhanced feature extraction network for corresponding feature fusion after one 1 × 1 convolution; L5 will be input to the spatial pyramid pooling layer SPP after three convolutions, and then enter the enhanced feature Extract the network for feature fusion. Finally, the final detection result can be obtained by sending it to the output terminal.

YoloV4''s backbone extraction network CSPDarknet53 is an improvement based on Darknet53. The original residual module is changed to the CSP structure as shown in Fig. 7. The CSP structure changes the original residual module into two parts, one part goes through the residual network and the other part Directly merging with the output of the residual network, this method can reduce the among of parameters and calculations while maintaining a high accuracy.
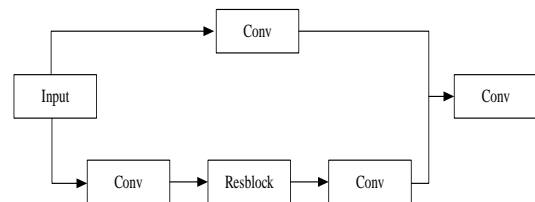
Fig. 7 CSP structure block

The PANet network is used to build a feature pyramid, analyze semantic features from top to bottom, and use the feature layer results obtained by down sampling and up sampling to reduce the loss of underlying feature information during the propagation process of the FPN

algorithm by increasing the bottom-up feature fusion path. problem [69], which improves the richness of the feature map and the detection accuracy of the algorithm. The feature map after feature fusion is predicted in three different sizes by Yolo-head. The prediction results include offset x, y, detection frame size w, h, and confidence level c of target information. Finally, redundant ones are eliminated through IOU and NMS. Detect the and NMS. Detect the frame and complete the target prediction.

3.1 Shaft hole type identification

As shown in Fig. 8, after the roundness detection of shaft holes of different sizes is completed, the YoloV4 algorithm needs to be processed to predict the category and position frame of the shaft hole. Before attitude detection, in order to remove the shaft hole target in the image, The effects of other complex backgrounds, noise, and light damage on the target surface existing in the sorting environment can be used for subsequent shaft hole type identification. We have preprocessed the image above, and then use YoloV4 to predict the image.

Fig. 8 Shaft hole image after inspection by YoloV4

4 Experiments and Analysis

4.1 Dataset and environment configuration

The data set used in the experiment of this article contains a total of 6 categories, with 600 data for each category. In order to prevent overfitting due to too little data, this article uses Gaussian noise and random color transformation for data enhancement. The final data the quantity is 3600 sheets.

The training and testing environment of this article is running on a Windows 10 system, NVIDIA GeForce RTX 4060TI GPU, and 8G video memory. The 3600 data sets

are divided, 90% of which are used for training and 10% for verification. Train for 100 epochs, and freeze the next 50 epochs. The initial value of the learning rate for the first 50 epochs is 0.01, and the batch size is 32. The initial value of the learning rate for the last 100 epochs is 0.001, and the batch size is 16. The Adam optimizer is used during training, and the cosine annealing learning rate decay strategy is adopted.

4.2 Experimental results and analysis

From the global loss function curve in Fig. 9, we can see that the loss curve drops rapidly in the early stages of iteration, indicating that the model is fitting quickly and the model's learning efficiency is high. When epoch is 40, the network model gradually becomes stable. After 100 iterations of training, the loss function of the final model converged to 1.50. In addition, by observing the consistency of the loss curves of the training set and the verification set, it can be seen that the generalization ability of the model has reached the best state.
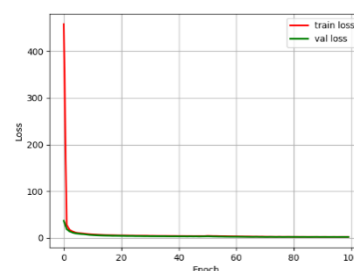
Fig. 9 Loss value curve

Fig. 10 shows the detection effect of the Yolov4 model. It can be seen that the network model accurately identified all buckles without missing detection. It can be seen from the above results that this method has a high accuracy in detecting targets and is practical.
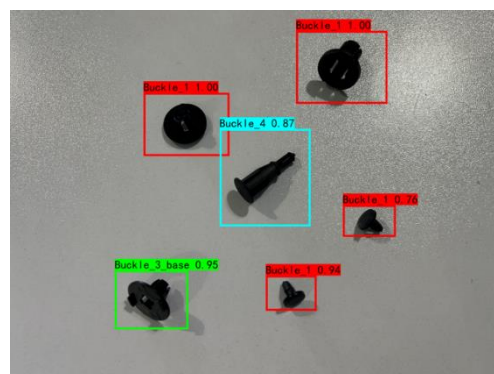
Fig. 10 Yolov4 detection results

In this experiment, a plastic bottle was placed next to the shaft hole as an interference object through the above steps, and 50 sorting tests were conducted, in which the positions of the plastic bottle and the shaft hole were randomly placed. The robot grabbing effect is shown in Fig. 11. The robot successfully avoided other interference and grasped the target shaft hole.

Fig. 11 Robotic arm grabs shaft hole

5 Conclusion

The vision-based automobile shaft hole roundness detection system proposed in this article improves the applicability of detection and can be widely used in roundness error detection of various workpiece nozzles. In the image pre-processing stage, by comparing various threshold segmentation algorithms and filtering algorithms, the adaptive threshold segmentation algorithm combined with the median filter algorithm was selected, which can not only remove the noise in the image, but also retain the edge features of the nozzle. The Canny detection algorithm is used to identify the contour of the nozzle, and the least squares method is used to detect the roundness error of the nozzle, and then combined with the YoloV4 algorithm to complete type recognition, and finally the robot completes the shaft hole identification and grabbing work. The results show that the designed detection system can not only improve the real-time performance and efficiency of detection, but also complete the subsequent grabbing work.

References

1   CHURSIN Y A，REDKO L A，FEDOROV E M. Enlargement of measuring zone in laser gauges without sacrificing measurement accuracy ［J］. Measurement，2019，131: 647－653.

2   LIAN M . Methods and key technologies of automatic ultrasonic thickness measurement of large thin－wall parts ［D］. Dalian: Dalian University of Technology，2019 .

3   URBAN S，WURSTHORN S，LEITLOFF J，et al . MultiCol bundle adjustment: a generic method for pose estimation， simultaneous self－calibration and reconstruction for arbitrary multi － camera systems ［J］. International Journal of Computer Vision，2017，121(2) : 234－252 .

4   CRAIG J J. Introduction to Robotics［M］. Yun Chao, waiting for translation. Beijing: Machinery Industry Press, 2006.

5   CHILLARÓN M，VIDAL V，VERDÚ G. Evaluation of image filters for their integration with LSQR computerized tomography reconstruction method［J］. PLOS One，2020，ne，2020，15(3) : e0229113.

6   FU S J，ZHONG W B，LIANG L Z. Optical fiber diameter non－contact measurement with computer image processing ［J］. Acta Scientiarum Naturalium Universitatis Sunyatseni，2000，39(2) : 24－27.

7   LI Y M，TU J K，XIANG H Z，et al . Measurement of optical fiber geometric parameters with Canny operator and binarization filtering ［J］. Optical Technique，2018，44(5) : 513－518.

8   SHEN X J，DUAN X Y，YUAN W L，et al. Research on circle detection algorithm based on connected region labeling algorithm［J］. Computer Engineering and Applications，2018，54(21) : 95－98.

9   LEI X Q，ZHANG Y D，MA W S，et al. Least square fitting and error evaluation of the convex contour of bearing roller ［J］. Optics and Precision Engineering，2018，26(8) : 2039－2047.

10 CAI Z，WANG J L，LYU M F，et al. Roundness error assessment based on improved cuckoo search algorithm ［J］. Modular Machine Tool & Automatic Manufacturing Technique，2020(7) : 40－44.

Authors Introduction

Junsheng Zhang

In 2021, he will be admitted to the School of Electronic Information and Automation, Tianjin University of Science and Technology, China, majoring in electronic information. He is currently a master's student at Tianjin University of Science and Technology, China.

Mr. Dongpo Ma

H He received his bachelor's degree in physics and optoelectronic engineering from Taiyuan University of Technology in 2019 and is now a graduate student in Tianjin University of Science and Technology

Dr. Yizhun Peng

He is an Associate Professor in Tianjin University of Science &Technology. He received a doctor's degree in control theory and control engineering from the Institute of Automation, Chinese Academy of Science, in 2006. His research field is intelligent robot and intelligent control.

Evaluation Standard of Error Recovery Planning Focused on Revival Process from Failures in Robotic Manufacturing Plants

Akira Nakamura

Department of Information Systems, Faculty of Engineering, Saitama Institute of Technology
1690 Fusaiji, Fukaya, Saitama 369-0293, Japan

Kensuke Harada

Robotic Manipulation Research Group, Systems Innovation Department
Graduate School of Engineering Science, Osaka University
1-3 Machikaneyama, Toyonaka 560-8531, Japan

Abstract

In recent times, intelligent robots have found applications across diverse fields. In scenarios demanding both repetitive and non-routine tasks, the likelihood of work errors increases. To address this, we have proposed methods encompassing both forward and backward recovery. Forward recovery suits minor modifications, while our focus is on backward recovery for substantial failures. Our study introduces a novel evaluation method to discern the optimal recovery path among various options.

Keywords: error recovery, task stratification, error classification, evaluation standard

1. Introduction

At the present day, information technology, particularly generative AI, is making rapid strides, a trend mirrored in the progress of robotic intelligence. As a consequence, robots face an increasing array of tasks during execution, often grappling with challenging assignments prone to errors and failures. This underscores the pressing need for effective methods to address these issues [1], [2], [3], [4], [5].

Over several years, our research [6], [7], [8], [9], [10] has focused on systematizing error recovery theory, resulting in a method based on task stratification and error classification concepts. The primary components of the robot system include sensing, modeling, planning, and execution sequences (Fig. 1). When an error occurs, the process transitions to the recovery phase. This section involves estimating the error's cause, classifying it, and correcting the original system. The refined process then operates on an enhanced, reliable system.

The proposed error recovery technology returns to the process before the step where the failure occurs and starts over from there. In practice, not only this type of backward recovery but also forward recovery, which moves forward after failure, is used. This study considers various paths from failure to recovery execution.

The concept of skills, which are motion primitives comprising a task, is described in Section 2. The fundamental technique for generating an error-recovery path is described in Section 3. Multiple possibilities for various recovery paths are presented in Section 4. Various evaluation standards and methods for selecting the most suitable path using the new evaluation method are considered in Section 5, and examples are presented in Section 6.

2. Concept of Skill

This section provides a brief overview of essential aspects of these skills [11], [12], [13].

2.1. Skill primitives

Motion primitives constituting tasks, termed "*skills*," were derived by analyzing the human behavior. Three crucial skills, "*move-to-touch*," "*rotate-to-level*," and "*rotate-to-insert*" (Fig. 2), are highlighted. A person's behavior, including representative and similar skills, serves as a model for the robot's motion primitives.

2.2. Stratification of tasks

The use of task hierarchies, as shown in Fig. 3, is effective for the execution of automated plants. The layer "$task^{(i+1)}$" occurs one tier above the layer "$task^{(i)}$," and the layer "*skill primitive*" is represented by the lowest layer "$task^{(0)}$."

Akira Nakamura, Kensuke Harada

3. Error Recovery

In an actual work environment, contrasting to an ideal case, various factors can lead to errors in the execution of a robot. This section outlines the error classification concept and error recovery technique [6], [7], [8], [9], [10].

3.1. Error classification

The errors can be categorized into several groups based on their possible causes. We considered four error groups: execution, planning, modeling, and sensing (Fig. 4).

3.2. Error recovery based on classification

First, if an error occurs, the cause is determined. Next, appropriate corrections are made to the system based on the tentative causes. The process returns to the previous step, and the task is executed again in this step (Fig. 4). Since a modified process was executed, the same error was less likely to occur. If the error scale is small, the process returns to the previous step in the lowest hierarchy (Figs. 4 and 5). Conversely, if the error scale is large, the process returns to the previous step in the highest-ranking layer of the hierarchy and is executed again from that step (Fig. 5).

4. Various recovery paths

The error recovery methods we proposed have primarily centered on rerunning the process by reverting to the step before an error occurs, constituting a backward error recovery process. However, alternative recovery procedures exist, including backward recovery for significant errors (failures) and forward recovery for minor errors. Additionally, errors may impact the environment, altering the arrangement or shape of surrounding objects, necessitating modified recovery processes. This section introduces several recovery procedures focusing on the degree of destruction of the environment surrounding an object according to [10].

Consider an indicative task sequence composed of n subtasks from the start to the goal (Fig. 6(a)). Suppose a failure occurs in the qth subtask during the process (Fig. 6(b)).

4.1. Recovery sequence I (RS-I): Complete restart

This method restarts work from the original starting point using the same process (Fig. 6(c)). If needed, the environment is restored to its original state, and if necessary, the original object or part is replaced with a new one and executed.

4.2. Recovery sequence II (RS-II): Restart from the middle of a prior process (without another process)
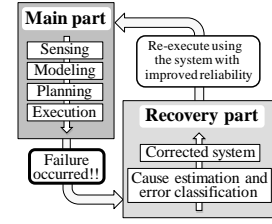
Fig. 1 Robot task system with an error recovery function

(a) Move-to-touch skill   (b) Rotate-to-level skill

(c) Rotate-to-insert skill
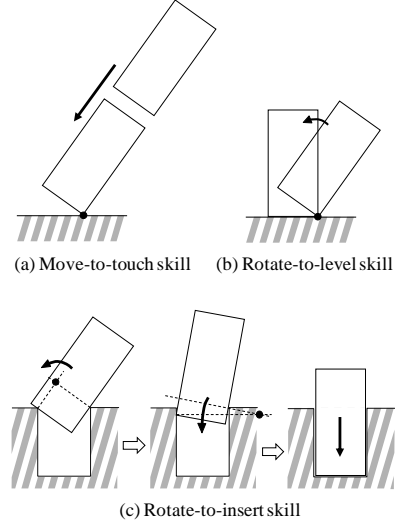
Fig. 2 Three fundamental skills
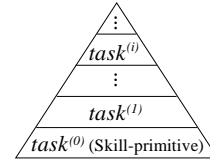
Fig. 3 Hierarchy of taskS

This method resumes work in the same process from subtask$_p$, the point in the middle of the process before the failure in the original process occurs (Fig. 6(d)).

4.3. Recovery sequence III (RS-III): Restart from the middle of a prior process (with another process)

This method resumes work from a point in the middle of the original process before a failure occurs (Fig. 6(e)). Unlike RS-II, a sequence from subtask$_t$ to subtask$_u$, not included in the original planning, restores the environment without causing problems for subsequent work.

4.4. Recovery sequence IV (RS-IV): Restart from the middle of a process that was supposed to come up later (with another process)

This method restarts work from the middle of the original process (Fig. 6(f)), differing from RS-III by reverting to its original process with subtasks scheduled after the original process failure. It includes a sequence from subtask$_v$ through subtask$_w$, not in the original plan.
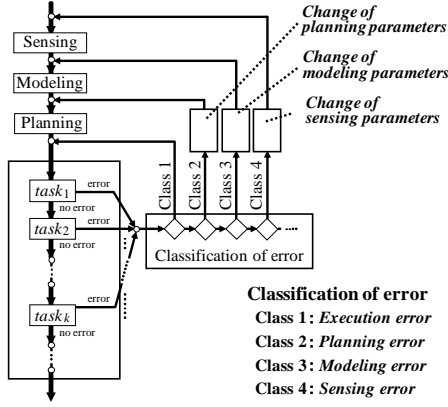
Fig. 4 Fundamental process flow
with error recovery

Classification of error
Class 1: *Execution error*
Class 2: *Planning error*
Class 3: *Modeling error*
Class 4: *Sensing error*
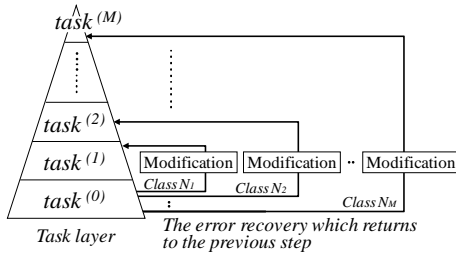
Fig. 5 The expression of task stratification and
the process flow of the error recovery

Table 1 Degree of correlation for each criteria

	Manufacturer ／Operator	Consumer ／User
(i) Cost	△	◎
(ii) Time	○	○
(iii) Reliability	◎	△
(iv) Safety	◎	△
(v) Finishing	△	◎
(vi) Recovery data	◎	－
(vii) Tool	◎	－
(viii) Operator skill	◎	－
(ix) Work efficiency	◎	－
(x) Environmental impact	◎	－
(xi) Damege	○	○

(a) Original task sequences from start to goal

(b) Failure Occurrence at $subtask_q$

(c) Original reproduction from start step

(d) Original repetition from step along the way (without another process)

(e) Original repetition from step along the way (with another process)

(f) Restart from step in the middle of the original process scheduled to run later

(g) Continuation of work, including consideration of any recovery process that may be required later

Fig. 6 Various error recovery sequences

4.5. Recovery sequence V (RS-V): Continuation of work including consideration of recovery needed later

In this method, work continues as long as possible after a failure. When further work becomes impossible, the environment is corrected, and work progresses (Fig. 6(g)). This modification involves a sequence from $subtask_x$ through $subtask_y$, not included in the original planning.

4.6. *As mentioned above, this study briefly categorizes the recovery sequences into five types. Recovery sequences RS-I through RS-III are backward recovery processes, whereas recovery sequence RS-V is a forward recovery process. Recovery sequence RS-IV is not a*

forward recovery process because it does not continue working after failure has occurred. However, because recovery sequences are subtasks scheduled to be executed after failure of the original process, this study included the recovery sequence RS- IV in the forward recovery process for convenience.

5. Selection of a recovery path

As explored in Section 4, a single failure often presents multiple possible recovery processes. The study advocates the use of evaluation standards to discern the most suitable process from several candidates. This research incorporates 11 evaluation standards, expanding on the four to eight standards introduced in [9].
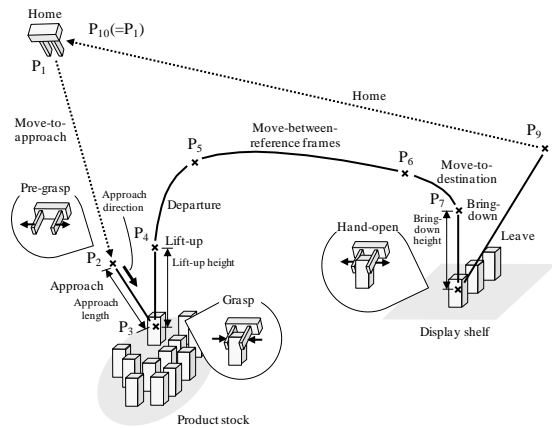
Fig. 7 Picking and placing task using a gripper

5.1. Various evaluation standards

(i) Cost

Cost is considered the most important evaluation standard. A recovery process with the minimum practical cost is selected.

(ii) Time

Time is an important evaluation standard. Priority is given to paths that require shorter recovery times.

(iii) Reliability

Reliability is considered a representative evaluation standard. The path with a high success rate is prioritized for accomplishing the recovery task.

(iv) Safety

Safety is considered a representative evaluation standard. Priority is assigned to paths that are less likely to harm people.

(v) Finishing

Finishing is considered an important evaluation standard. It prioritizes paths displaying excellent completion of the operation on the target object.

(vi) Recovery data

Recovery data is considered as the evaluation standard. Processes with a lot of useful data are prioritized.

(vii) Tool

The tool used for recovery is evaluated based on workability, with priority given to tools displaying high efficiency.

(viii) Operator skill.

Operator skill during the recovery process is an evaluation standard, favoring processes managed by a larger number of skilled experts.

(ix) Work efficiency.

Work efficiency is considered as an evaluation standard. Selecting an efficient work plan, that is, a route that is not unreasonable, wasteful, or uneven, is important. This index is common to (i) cost and (ii) time, but even if it is an efficient route, it may not necessarily be cheap or fast; therefore, it was set as a separate indicator.

(x) Environmental impacts

Environmental indicators, including noise levels, are considered, accounting for overall environmental problems such as air, ocean, and water pollution.

(xi) Damage

Damage is a measure of the difference between an object without and with errors. Although it may not be readily visible externally, the quality of an object could change owing to errors in the production process.

5.2. Selection of a recovery path using evaluation standards for each section

Reference [9] details how to choose the best recovery path for each evaluation standard. In contrast, this study proposes specifying indicators for each section of the recovery process and selecting the path with the superior total indicator. This prevents situations where a path chosen by a single indicator may be partially inappropriate.

Furthermore, evaluations may vary depending on the recipient of the standard. Table 1 delineates whether the evaluation is from the Manufacturer/Operator side or the Consumer/User side, indicating the degree of involvement with symbols ◎, ○, △, and - in descending order. For instance, a process with high manufacturer/operator involvement may be chosen in the first half of the recovery stage, while a process with high consumer/user involvement may be selected in the second half. This ensures optimal restoration on the manufacturer/operator side and maximum satisfaction with the finished object on the consumer/user side.

6. Example of error recovery in product display

The task involves utilizing a manipulation robot to arrange products on shelves or in the display window of a convenience store. The objective is to position and posture the products optimally, as a tidy arrangement significantly boosts customers' inclination to make purchases, as opposed to disorganized placements. Fig. 7 displays an image illustrating a pick-and-place task that transfers a single object from a product stock to a display space. Furthermore, Fig. 8 showcases the command sequence constituting a singular pick-and-place task. In this instance, the task revolves around evenly arranging 12 identical rectangular products in a layout of three rows
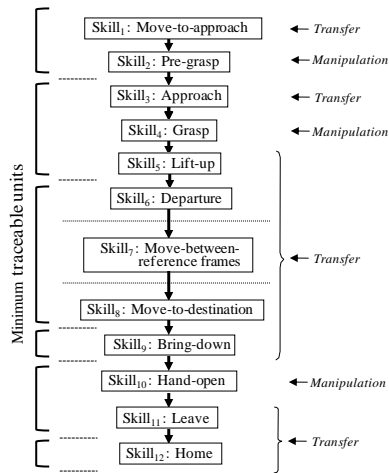
Fig. 8 Task sequence of picking
and placing a product
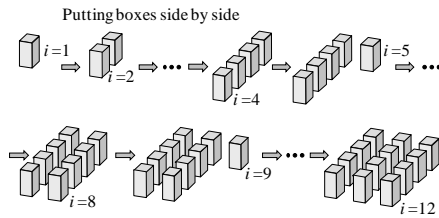
Fig. 9 Sequence of tasks related
to display of products

(i) Correct placement task

(ii) Failure to shift adjacent products

(iii) Failure to knock down adjacent product
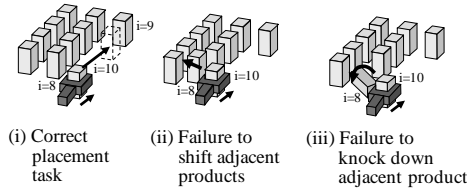
Fig. 10 Successes and Failures in tasks of
display of products

(a) Original task sequences {subtask$_i$}(i=1,2,···,12)

(b) Failure Occurrence at a moving task in subtask$_{10}$

(c) re-excution from start step

(d) Original repetition from the first step of subtask(i=10) (with another process)

(e) Original repetition from step along the way (with another process)

(f) Continuing working to the end and aligning if necessary

Fig. 11 Various recovery patterns

and four columns, as depicted in Fig. 9. Fig. 10 illustrates an error scenario; Fig. 10(i) demonstrates a carry-and-place operation for a pre-planned 10th object. Subsequently, Fig. 10(ii) portrays an instance where a hand interferes with the 10th object during the carry-and-place operation, causing the displacement of the 8th item from its original position. Finally, Fig. 10(iii) presents a scenario wherein the hand impacts the 8th item during the carry-and-place operations for the 10th item, resulting in the toppling of the 8th item.

Fig. 11 shows various recovery patterns. Fig. 11(a) illustrates normal planning in which no errors occur. Fig. 11(b) depicts the point at which the aforementioned failure occurred during the pick-and-place operation of the 10th object. Figs. 11(c)–11(f) show the respective
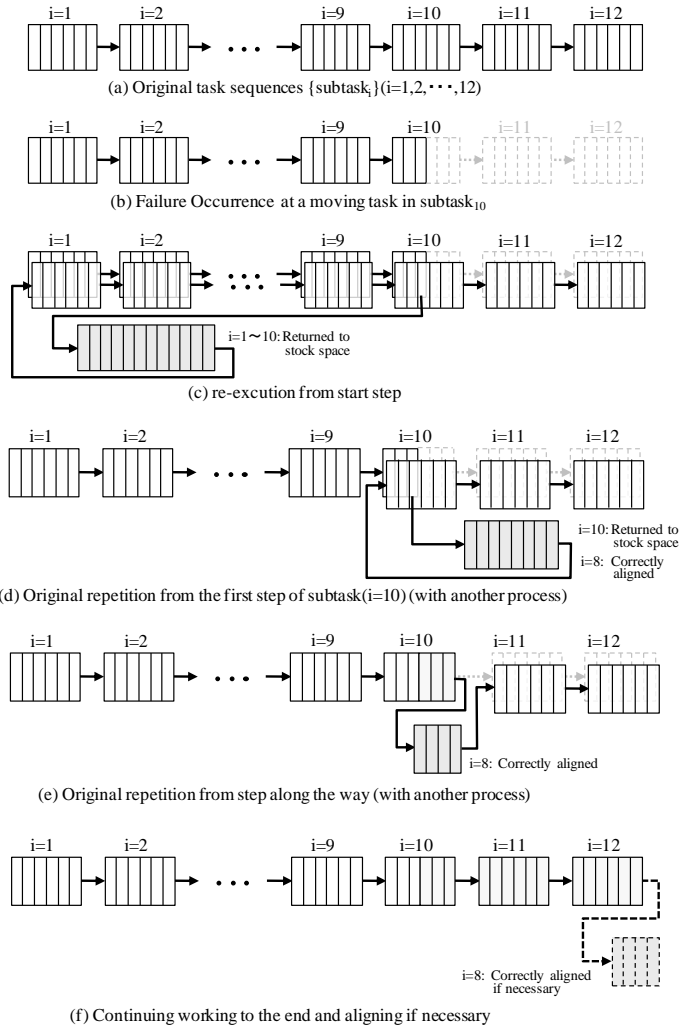
recovery patterns. Fig. 11(c) demonstrates the error recovery of the backward type back to the start, in which the 1st object through the 8th object are returned to their original stock space, and the task is rerun from the beginning. Fig. 11(d) displays the error recovery of the backward type, where the recovery task is to return the 10th item to its original stock space, restore the 8th item to its original position and orientation, and pick and place the 10th item again. Fig. 11(e) shows the forward error recovery, which executes a recovery task that places the 10th item in its designated position even after an error occurs and then returns the 8th item to its original position and orientation, followed by the 11th pick-and-place. Fig. 11(f) shows a special type of forward error recovery, which is a recovery task that continues the work to the end even after an error occurs and returns the eighth

Akira Nakamura, Kensuke Harada

product to its original position and orientation, if necessary. If a product is displayed in a store, it may not matter much if it is misplaced or even fallen over; that is how to deal with it. However, if it resonated with a customer's willingness to buy, it is executed. Although this is a particularly obvious example, it is sometimes desirable to perform the first half of the recovery task based on the operator's criteria and the second half of the recovery task based on the user's criteria.

7. Conclusion

When an error arises during the primary task, the process transitions into the recovery phase. This section explores diverse recovery paths utilizing our proposed error recovery method, grounded in both task stratification and error classifications. We introduced a method to systematically derive the optimal recovery path, aligning with evaluation standards for each section.

As discussed, numerous recovery paths exist, and choosing the right one poses challenges. While this paper introduced a method for optimal path selection, it does not delve into defining evaluation standards for each section. Addressing this aspect becomes imperative for future research endeavors.

References

1. B. R. Donald, Planning multi-step error detection and recovery strategies, Int. J. Robot. Res., 9(1) (1990) 3-60.
2. T. Niemueller, G. Lakemeyer and S. S. Srinivasa, A Generic Robot Database and its Application in Fault Analysis and Performance Evaluation, in Proc. IEEE/RSJ Int. Conf. on Intell. Robots Syst., (Vilamoura, Portugal, 2012), 364-369.
3. E. D. Lello, M. Klotzbucher, T. D. Laet and H. Bruyninckx, Bayesian Time-Series Models for Continuous Fault Detection and Recognition in Industrial Robotics Tasks, in Proc. IEEE/RSJ Int. Conf. on Intell. Robots Syst., (Tokyo, Japan, 2013), 5827-5833.
4. E. Krabbe, E. Kristiansen, L. Hansen and D. Bourne, Autonomous Optimization of Fine Motions for Robotic Assembly, in Proc. IEEE Int. Conf. Robot. Autom., (Hong Kong, China, 2014), 4168-4175.
5. A. S. Wang and O. Kroemer, Learning Robust Manipulation Strategies with Multimodal State Transition Models and Recovery Heuristics, in Proc. IEEE Int. Conf. Robot. Autom., (Montreal, Canada, 2019), 1309-1315.
6. A. Nakamura, K. Nagata, K. Harada, N. Yamanobe, T. Tsuji, T. Foissotte and Y. Kawai, Error recovery using task stratification and error classification for manipula-tion robots in various fields, in Proc IEEE/RSJ Int. Conf. on Intell. Robots Syst., (Tokyo, Japan, 2013), 3535-3542.
7. A. Nakamura, K. Nagata, K. Harada and N. Yamanobe, Technique of Recovery Process and Application of AI in Error Recovery Using Task Stratification and Error Classification, J. Robotics, Networking and Artificial Life, 5(1), 2018, pp. 56-62.
8. A. Nakamura, N. Yamanobe, I. R. Alpizar, K. Harada and Y. Domae, Cost-oriented Planning for Error Recovery in an Automation Plant, J. Robotics, Networking and Artificial Life, 6(4), 2020, pp. 225-230.
9. A. Nakamura, N. Yamanobe, I. R. Alpizar, K. Harada and Y. Domae, Selection of Optimal Error Recovery Process using Evaluation Standards in Automated Plants, J. Robotics, Networking and Artificial Life, 8(3), 2021, pp. 211-217.
10. A. Nakamura and Harada, Error Recovery Patterns Focusing on the Revival Process from Failures in Manipulation Tasks, J. Advances in Artificial Life Robotics, 3(4), 2023, pp. 230-237.
11. T. Hasegawa, T. Suehiro and K, Takase, A robot system for unstructured environments based on an environment model and manipulation skills, in Proc. IEEE Int. Conf. Robot. Autom., (Sacramento, USA, 1991), 916-923.
12. A. Nakamura, T. Ogasawara, T. Suehiro and H. Tsukune, Fine motion strategy for skill-based manipulation, Artificial Life and Robotics, Springer, 1(3), 1997, pp. 147-150.
13. A. Nakamura, K. Kitagaki and T. Suehiro, Using simplified geometric models in skill-based manipulation, Advanced Robotics, 18(8), 2004, pp. 835-858.

Authors Introduction

Prof. Akira Nakamura

He received the Ph.D. degree in Electrical Engineering from Keio University in 1991. From 2021, he has been working as a Professor at Faculty of Engineering of Saitama Institute of Technology. His research interests include robot planning, vision and control system.

Prof. Kensuke Harada

He received his Doctoral degrees in Mechanical Engineering from Kyoto University in 1997. From 2016, he has been working as a Professor at Graduate School of Engineering Science, Osaka University.

Robotic Food Handling Utilizing Temperature Dependent Variable-Stiffness Material

Rozilyn Marco
Division of Engineering Science, University of Toronto

Prashant Kumar, Xinyi Zhang, Weiwei Wan, Kensuke Harada
Robotic Manipulation Research Group, Osaka University
Osaka, Japan
E-mail: rozilyn.marco@mail.utoronto.ca, {kumar@hlab.,chou@hlab.,wan@,harada@}sys.es.osaka-u.ac.jp

Abstract

This paper presents a robotic end-effector that addresses challenges of automated robotic food handling. We utilize a variable stiffness fabric on the finger surfaces of the gripper. Depending on its temperature, the finger stiffness changes to tackle grasping food items with different physical characteristics. Hard objects can be grasped with the hard mode of the finger while a fragile object can be safely grasped with the soft mode. This gripper design was validated empirically through force and object-grasping experiments.

Keywords: robotic end-effector, food picking, variable stiffness, temperature dependence

1. Introduction

Within industry, the adoption of robotic automation is driven by the promise of heightened efficiency and the reduction of laborious, repetitive tasks that can exert both mental and physical strain on humans [1]. The food industry is a continuous source of opportunities for advancements in automation. Successful automated food handling can improve product quality and production rates [2], however, is a challenge due to inherent inconsistencies in food size, texture, and delicacy [3]. The automated processes may take significant time or fail to adjust for these variances. This paper addresses these challenges by utilizing a compact, mechanically simple design that takes advantage of the dynamic control of its applied force and stiffness variability. The adjustable compliance of the device's material is used in combination with force feedback sensors to successfully eliminate the chance of bruising that pinching or other enclosing end-effectors can cause [2]. Force feedback or tactile sensors are commonly used in industrial robotics [4], allowing the design to remain industry compatible. The goal of this paper is to demonstrate the potential of

utilizing this material to handle items requiring different amounts of applied force by only altering the temperature exposed to the gripper - eliminating the need for extensive or challenging control procedures required by other soft robotic grippers [5]. We attached the variable stiffness material to the finger surface of a two-fingered gripper and the compliancy increases as the temperature becomes higher. We demonstrate the gripper can successfully adapt to grasping a variety of food materials by changing the finger stiffness according to the stiffness properties of food. The effectiveness of the proposed gripper is verified by experimental results.

2. Related Work

Food items that are heavy, slippery, fragile, or have a minimal-height profile commonly experience grasping difficulties with currently proposed gripper designs [6], [7]. This gripper is compatible with objects having these attributes that other developed variable-stiffness general grippers are not [8]. The design proposed by A. Pettersson et al. utilizing magnetorheological fluid [9] also maintains a secure grasp through a gripper with high and adaptable compliance. However, an improvement

Rozilyn Marco, Prashant Kumar, Xinyi Zhang, Weiwei Wan, Kensuke Harada

this paper introduces is the compactness of the design — a crucial attribute for practical deployment in confined spaces [10]. Other grippers that grasp food items by the side surfaces require the object's size to be predetermined [10]. The paper by G. Endo et al. [11] required the height of the samples to be known and constant. Food items can vary in size and shape even between samples of the same food, different apples for example, and this requirement limits the success and robustness of the gripper. The design proposed in this paper eliminates that requirement through its dynamic force application and compatibility with force feedback implementation. The gripper by R. Maruyama et al. [12] was successful in handling fragile and soft food items but required predetermination of the required force to apply to obtain a successful grasp. The design of this gripper allowed for more innate tolerance in this aspect, allowing it to work with food items not encountered previously.

3. Design

The goal of the design was to achieve both simplicity and effectiveness. Simple mechanical design minimizes the risk of contamination during usage due to component separation [10] and the gripper's achieved compactness, increases the ease of adoption in practical scenarios [10]. Relying on the variability of the gripper's compliance from temperature exposure as the means of adapting the force application decreases the requirement for precise mechanical human adjustments thus not requiring an upskill in labour [13].

3.1. Concept

The robotic end-effector incorporates a variable stiffness fabric as its primary functional component and is the main contributor to the gripper's low mechanical complexity. This fabric exhibits temperature-dependent properties, becoming more flexible as the temperature rises and stiffer as it decreases. When controlled, this dynamic behaviour can be used to handle objects with different structural integrities. Explicitly, the end-effector can be subjected to higher temperatures when the target for transportation is fragile or soft as the fabric would apply less force while in this more compliant state. Likewise, when the fabric stiffens when exposed to lower temperatures, the gripper can exert greater force, making it suitable for firmly grasping rigid food items. This ultimately allows for a secure hold and movement

without causing deformation or damage to samples in both scenarios.

Between grasps, warm temperature can be applied to return the fabric to its neutral position. This methodology of subjecting the gripper to warm temperature after each grasp will be applied during the experiments used in this paper and can be seen in the flow chart of Fig. 1.
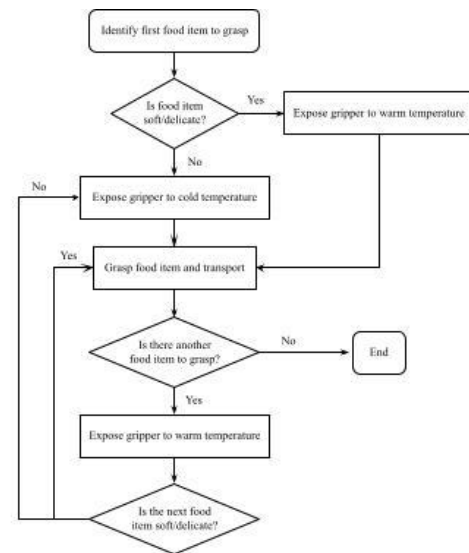
Fig. 1. Decision flow chart when picking an item.

3.2. *Implementation*

The base structure of the gripper was created through CAD software and was 3D printed (Fig. 2) with the fabric attached via sewing. Aside from providing a strong foundation, the base material is inconsequential to the functioning of the gripper. It can thus be substituted for a hygienic material such as stainless steel [14].
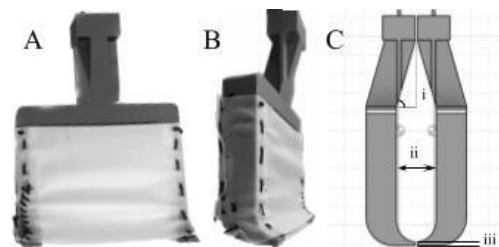
Fig. 2. Gripper physical implementation and CAD model: (A) Front Profile, (B) Three Quarter Profile, and (C) CAD side profile of gripper (i) 73.7°, (ii) 20.5mm, and (iii)2mm.

An essential feature of the gripper base is the curved nature at the bottom which guides the fabric beneath the item and provides support to the bottom when the gripper encloses it. This creates a secure position between the fabric and the food item and prevents the item from slipping through the bottom during transportation. To prevent the loss of the curved formation (Fig.2.B), the fabric was attached to the base while exposed to warm temperature. This was a method of pre-tensing the fabric to ensure the curved position remains taut as the fabric's innate stiffness is reduced at higher temperatures.

The 2mm between the bottom face and the fabric (Fig. 2.C.iii) enables effective grasping of items with narrow height profiles. The angle of the stem and the radius of the inner curvature of the base are 16.3° and 11mm respectively. These values were chosen to minimize the gap between the hands of the gripper and maximize the amount of fabric held vertically. Minimizing the gap between the hands increased the minimum width of food items that could demonstrate the enclosing property of the gripper i.e. in Fig. 2.C.ii. Maximizing the amount of vertical fabric increases the area that applies uniform force along the side of an object. Future design iterations will investigate methods of allowing these dimensions to be variable, increasing the versatility of the gripper.

4. Experiments

The first experiment executed was a force gauge experiment to observe the force applied to the fabric depending on displacement and fabric temperature. The second was a set of demonstrative experiments to display the success of the gripper in a variety of arrangements and sample types. Water was selected as the stimulus medium as it possesses a high heat capacity, meaning it requires a relatively high amount of energy to change its temperature by one degree [15]. This ensured it maintained at a considerably stable temperature throughout the execution of an experiment sequence.

4.1. *Force gauge experiment*

The force gauge experiment was conducted with the gripper to analyze the force applied to an object over time under two temperatures and at three final displacement positions. Specifically, 30°C and 67°C, and 10mm, 20mm, and 30mm respectively. The experiment setup consisted of the force gauge sensor mounted onto a cart of a linear motion apparatus with a blocking cart positioned at the desired displacement away from the

force gauge's starting position (Fig. 3). To depict the behaviour throughout the different segments of a standard pick-and-place operation, the carts were:

- manually moved the same velocity until the final displacement to mimic the initial grasping,
- maintained its position for approximately 10 seconds to simulate the effect the gripper would have during the time required for transportation,
- and then released, as if the object has arrived at its final location.
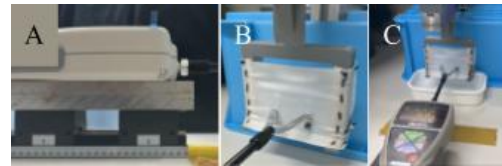
Fig. 3. Force gauge experimental set-up: (A) gauge cart and displacement blocker cart on track, (B) gripper and gauge prong, and (C) entire force gauge setup.

4.2. *Food picking experiment*

The goal was to showcase and verify the final design's practical ability to handle a variety of characteristics - heavy, slippery, fragile, brittle, flat, and rigid. Accordingly, corresponding food items were used during testing as seen in Table 1 and Fig. 4.

Table 1. Food Item Characteristics and Water Temperature Assignment

Item	Characteristics	Water Temperature
A	Heavy – 66 g	Hot
B	Hard, non-uniform vertical profile	Cold
C	Malleable, slippery plastic surface	Hot
D	Soft, narrow	Hot
E	Fragile	Hot
F	Hard, spherical	Cold
G	Brittle	Hot
H	Hard, narrow – 12.95 mm	Cold
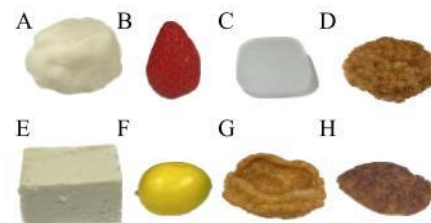
Fig. 4. Experimental food items.

Rozilyn Marco, Prashant Kumar, Xinyi Zhang, Weiwei Wan, Kensuke Harada

The gripper was exposed to temperatures depending on whether the physical properties of the target object are better suited for greater or weaker applied force during grasping. Each food item was tested individually on the platform as well as in more complicated combinations including multiple items on the platform grasped in continual sequence and objects in a stacked formation (Fig. 5). A successful grasp is defined in this paper by moving a sample from its initial to final location without causing damage or the item slipping mid task.
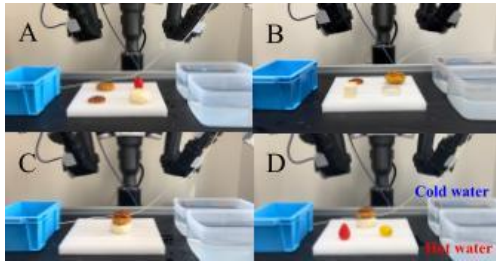
Fig. 5. Experiment arrangements: (A) Sequential (B) sequential (C) stacked, and (D) combined stacked and sequential.

The gripper was attached to an industrial robot and integrated with software that controlled grasp planning and execution while employing a depth camera for the identification of item locations. The gripper's linear actuator operates using force feedback to determine when an adequate clamping width has been achieved.

5. Results

5.1. *Force gauge*

The results are seen in Table 2 and Fig. 6 for each final displacement value, the fabric exposed to the cooler temperature applied more force compared to when exposed to a higher temperature. This confirms the premise of the gripper i.e. force applied increases as the temperature of operation decreases and vice versa.

Table 2. Force Gauge Experiment Key Values

Displacement	Max Force	Plateaued Force	Velocity
Hot Temperature - 67°C			
10 mm	0.6 N	0.4N	1.429 mm/s
20 mm	1.8 N	1.3 N	1.818 mm/s
30 mm	3.8 N	2.8 N	1.887 mm/s
Cold Temperature - 30°C			
10 mm	2.8 N	0.7 N	1.613 mm/s
20 mm	6.5 N	2.5 N	1.538 mm/s
30 mm	9.6 N	3.8 N	1.657 mm/s
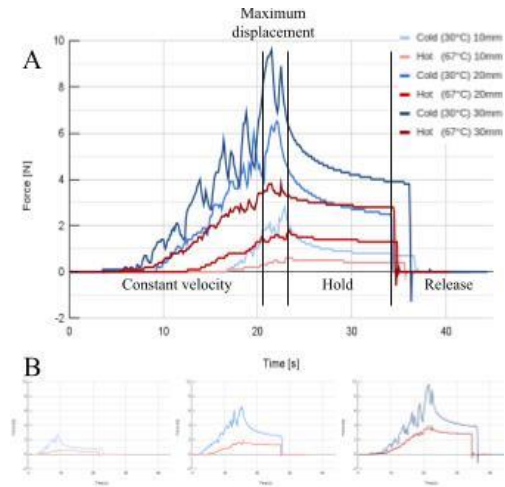
Fig. 6. Results of the force gauge experiment plotted: (A) all displacements and (B) each displacement individually.

The constant velocity of the gripper movement meant that time and displacement differed by a constant factor. Therefore, while accounting for this factor, modelled relationships with respect to time can proxy relationships with respect to displacement. The constant velocity portion of the experiment was isolated and time vs. force - or equivalently displacement vs. force - was analyzed using both linear and second-order polynomial models. A strong second-order polynomial relationship was determined between the displacement of the gripper and the force applied as seen with the R-squared values in Table 3. In future work, an analytical model will be formalized to depict the force a sample will experience depending on the item's width (i.e. the gripper displacement) and considering the temperature.

Table 3. Force Gauge Experiment Model Performance

Displacement	Linear Model	Polynomial Model
Hot Temperature - 67°C		
10 mm	0.909	0.927
20 mm	0.971	0.971
30 mm	0.969	0.983
Cold Temperature - 30°C		
10 mm	0.931	0.932
20 mm	0.957	0.962
30 mm	0.888	0.936

When the final displacement is achieved, the force quickly decreases and then plateaus. This is ideal as there is no prolonged period of unnecessary high force applied after the fabric has molded to the shape of the target. As seen in Table 2, there is a directly proportional

relationship between this plateaued value and the displacement of the gripper and an inversely proportional relationship between the force and temperature. Given that in practice, the displacement of fabric - i.e. a target's width - is immutable, plateaued force can be controlled with temperature. The relationship between this plateaued force, the temperature of the fabric, and the final displacement value can be further investigated to optimize and formalize the temperature used for various targets with different ideal handling requirements.

5.2. *Food picking*

The gripper was successful in grasping and transporting each food sample from its respective initial positions to their final destinations. Fig. 7 shows the end-effector gripping pose for each food item from Fig. 4 and Fig. 8 shows the gripper throughout experiment D from Fig. 5.
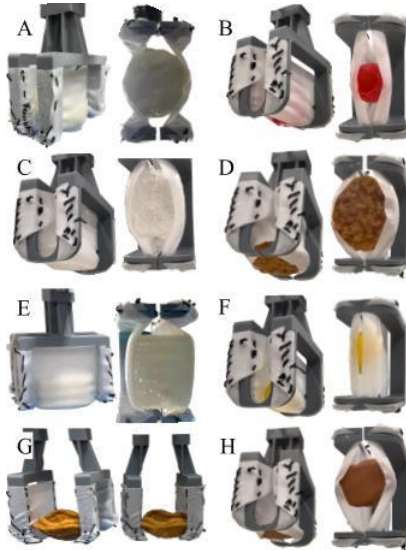
Fig. 7. Gripping snapshots of each food item.

Fig. 8. Snapshots during the execution of experiment D.

Unsuccessful grasping attempts were seen with samples with widths comparable to the maximum open position width of the gripper and with the samples in experiments with. These failures can be attributed to incorrect positioning of the gripper prior to closure. This could be corrected by further developing the method of identifying the desired initial location through the calibration matrix of the software. Additionally, the gripper can be modified to have a bigger difference in the open position width of the gripper compared to the size of the samples to have larger tolerance limits to account for the margin of error between generated and desired locations.

6. Conclusion

The robotic gripper proposed in this paper was designed to utilize its variable stiffness property to successfully grasp food samples with commonly difficult-to-handle features without causing damage and while maintaining mechanical simplicity and compactness. This concept was formed on the basis that allowing the material of the gripper to do the bulk of adapting to the shape and delicacy of a given sample permits the gripper to be functional without human intervention or immediate control. The gripper's variable force application was validated in a force gauge experiment conducted under two temperature conditions and at three displacement values, confirming the viability of our design concept. The practical capabilities of the gripper were demonstrated experimentally with the gripper performing food-picking sequences involving a diverse range of items, each possessing distinct properties and arranged in both sequential and stacked configurations.

In future developments, enhancing the gripper's versatility could involve the implementation of variable width and height dimensions to accommodate an even broader range of items. Moreover, expanding the range of temperatures utilized for more customized force applications can be investigated for more tailored handling of samples.

Rozilyn Marco, Prashant Kumar, Xinyi Zhang, Weiwei Wan, Kensuke Harada

References

1. Tomei G, Cinti ME, Cerratti D, Fioravanti M. Attenzione, lavori ripetitivi, fatica e stress [Attention, repetitive works, fatigue and stress]. Ann Ig. 2006 Sep-Oct;18(5):417-29. Italian. PMID: 17089957.

2. F. Bader and S. Rahimifard, "A methodology for the selection of industrial robots in Food Handling," Innovative Food Science and Emerging Technologies, vol. 64, p. 102379, 2020.

3. C. Blanes, M. Mellado, C. Ortiz, and A. Valera, "Review. Technologies for Robot Grippers in Pick and place operations for fresh fruits and vegetables," Spanish J. of Agricultural Research, vol. 9, no. 4, p. 1130, 2011.

4. P. Li and X. Liu, "Common sensors in Industrial Robots: A Review," J. Physics: Conference Series, vol. 1267, no. 1, p. 012036, 2019.

5. D. Rus and M. T. Tolley, "Design, fabrication and control of Soft Robots," Nature, vol. 521, no. 7553, pp. 467–475, 2015.

6. S. Ma, L. Du, E. Tsuchiya and M. Fuchimi, "Paper-Made Grippers for Soft Food Grasping," 2020 17th Int. Conf. on Ubiquitous Robots (UR), 2020, pp. 362-367.

7. Z. Wang, H. Furuta, S. Hirai, and S. Kawamura, "A scooping-binding robotic gripper for handling various food products," Frontiers in Robotics and AI, vol. 8, 2021.

8. D. Cardin-Catalan, S. Ceppetelli, A. P. del Pobil, and A. Morales, "Design and analysis of a variable-stiffness robotic gripper," Alexandria Engineering J., vol. 61, no. 2, pp. 1235–1248, 2022.

9. A. Pettersson, S. Davis, J. O. Gray, T. J. Dodd, and T. Ohlsson, "Design of a magnetorheological robot gripper for handling of delicate food products with varying shapes," J. of Food Engineering, vol. 98, no. 3, pp. 332–338, 2010.

10. Z. Wang, S. Hirai, and S. Kawamura, "Challenges and opportunities in Robotic Food Handling: A Review," Frontiers in Robotics and AI, vol. 8, 2022.

11. G. Endo and N. Otomo, "Development of a food handling gripper considering an appetizing presentation," 2016 IEEE Int. Conf. on Robotics and Automation, 2016, pp. 4901-4906.

12. R. Maruyama, T. Watanabe and M. Uchida, "Delicate grasping by robotic gripper with incompressible fluid-based deformable fingertips," 2013 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, Tokyo, Japan, 2013, pp. 5469-5474.

13. L. N. K. Duong et al., "A review of Robotics and autonomous systems in the food industry: From the supply chains perspective," Trends in Food Science andTechnology, vol. 106, pp. 355–364, 2020.

14. C. Faille, C. Cunault, T. Dubois, and T. Bénézech, "Hygienic design of food processing lines to mitigate the risk of bacterial food contamination with respect to environmental concerns," Innovative Food Science and Emerging Technologies, vol. 46, pp. 65–73, 2018.

15. S. Jordan, Energy and Light. Milton Keynes: Open University, 2009, p. 179.

Authors Introduction

Ms. Rozilyn Marco

She is currently an undergraduate student at the University of Toronto studying Engineering Science. She was a visiting student at Harada Lab at Osaka University.

Mr. Prashant Kumar

He received his B.Tach degree in Mechanical Engineering in 2018 from the SRM University in India and his M.E. in System Innovation from Osaka University in 2023. He is currently a Doctoral student at Osaka University.

Dr. Xinyi Zhang

She received the Ph.D. degree from Osaka University in 2023 while a member of the Harada Lab.

Dr. Weiwei Wan
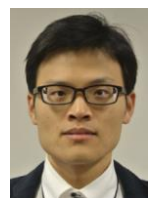
He is currently an Associate Professor with the School of Engineering Science at Osaka University. He received the Ph.D. in robotics from The University of Tokyo. His research interests include smart manufacturing and robotic manipulation.

Dr. Kensuke Harada

He is currently a Professor with the Graduate School of Engineering Science, Osaka University. He received his Ph.D. degree from Kyoto University. His research interests include mechanics and control of humanoid robots and robotic hands.

Fruits Maturity Classification for Harvesting Robot

Reno Muhammad Fadilla*1, Tokuo Tsuji*2, Tatsuhiro Hiramitsu*3, Hiroaki Seki*4, Kaihei Okada*5
Kanazawa University, Kakumamachi, Kanazawa, Ishikawa, 920-1192, Japan
*Email: renofadilla01@stu.kanazawa-u.ac.jp*1, tokuo-tsuji@se.kanazawa-u.ac.jp*2, thiramitsu@se.kanazawa-u.ac.jp*3,*
*hseki@se.kanazawa-u.ac.jp*4, kaihei2112@stu.kanazawa-u.ac.jp*5*

Abstract

This research proposes to create a machine vision using YOLOv5 model to detect maturity of the tomatoes on farm field which categorized by raw, half-ripe, and ripe. The use of multiple color space RGB, HSV, and LUV dataset aims to overcome external disturbances such as light and shadows. The results are the machine vision can differentiate fruits by its category with minor bias and obtaining a good mAP50 rates with the greatest result is LUV color spaces model. This machine vision with custom datasets in multiple color spaces can be implemented on farm field which is considered based on the purposes, environment and external disturbances. These models are developed for autonomous harvesting robot on farm field.

Keywords: Machine Vision, YOLOv5, Color Space

1. Introduction

The gradual maturity of technology has developed in many sectors, one of them is in the agricultural sector called as precision agriculture. Many agricultural systems still using traditional method, particularly in the maturity detection level. The research on precision agriculture aims to transform traditional method into digitalization precisely [1]. Robotics and automation are expected to have a significant impact on farms of the future by increasing efficiency in production for fruits grading which are gradually shifting to automation through a Computer Vision based on fruits detection. In addition, image processing can provide understanding of individual health, nutritional condition, and maturity level. However, the complexity of the field environment as well as the unstructured features of fruits resulting a great challenge to the target detection, particularly for maturity level measurement [2]..

Computer Vision (CV) is a key point for future implementation on farming technology, particularly in the fruits grading and classification, yield estimation counting, health condition, and maturity. Fundamentally, CV generated using Machine Learning (ML) algorithm which enable the analysis of massive quantities of data more rapidly and precisely, instead traditional or manual method. CV traditionally execute image by primary color Red, Green, and Blue (RGB), so the images collected by RGB based camera at RGB wavelength. In addition, grading fruits can be classified by using color, texture, and share feature descriptors. RGB color detection has a good level accuracy when using CV, however, at some condition, CV cannot detect object which has rough or irregular environment, such as on farm field. On the fam field, the objects can be disturbed by external factor, for example, sun rays, leaf, and shadows [3], [4]. To overcome this issues, execution process of CV can be implemented using multiple color space identification.

Color space basically can also describe as the way in which human eyes can perceive color vision. Color space shaped from the development of color model which the color is associated with an accurate definition of the way the components are to be inferred and the situations are viewed with the set of resulting colors. There are many different color spaces such as RGB, NTSC, LUV, HSV, CMYK, and HSI. Each color space has specified function and color model [5].

In the recent years, many researches have conducted to accomplish the best algorithm for ML, basically in image processing, particularly in the fruit's detection based on its color space. Detection of seedling using machine-transplanted rice has been developed using machine vision [6] which aimed to detect paddy seedlings using many color spaces and plant the seeds using a paddy machine. Research [7] conducted research by tomato recognition and localize the object using YOLOv5 algorithm only in RGB images. Research [8] discussed about a quality evaluation for two type of grapes using random forest model to identify quality of each fruit type. Another implementation of ML in image processing on farming activity also developed for palm fruit [9] which purposed to classified the oil palm fresh fruit based on the textures and palm's skin.

This research proposed to determine maturity level for tomato fruits based on three different color space using ML and YOLOv5 model on the real farm implementation to develop the previous similar researches. The maturity level of the tomato was divided into three class those are raw, half ripe, and ripe. In addition, this research was classified the greatest result of dataset's performance based on its color space.

Reno Muhammad Fadilla, Tokuo Tsuji, Tatsuhiro Hiramitsu, Hiroaki Seki, Kaihei Okada
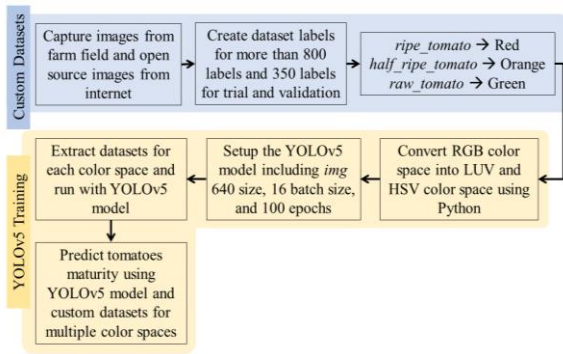
Fig. 1. Flowchart of proposed system.

2. Methodology

2.1. Proposed System

The overall flowchart of the proposed system is presented on the Fig. 1. The specified fruits which used as the main object is tomato. The system was proposed using Machine Learning based on YOLOv5 Algorithm with custom datasets. The used of YOLOv5 in this research rather than another version of YOLO has recently analyzed based on recognition and location of on-farm fruits. The previous research [7] explained about the YOLOv5 accuracy for on-farm moving robot. YOLOv5 has a high-level accuracy rate for rapid image detection with only 9.0 time/ms. Besides the accuracy level can be reach for more than 95%. In addition, the number of objects detected by YOLOv5 is the highest then other YOLO model. YOLOv7 is the latest model of YOLO version however the average detection time is 13.59 time/ms and with a little bit lower accuracy rate compared with YOLOv5. Based on this literature, originally, this research purpose is to develop a machine vision for on-farm robot. Thus, it is required a machine vision to detect fruits rapidly and precisely. All of the test has the same requirements such as 100 epochs, 32 batch with custom datasets in 640 x 640 image size, and three object classes.

This research was conducted in many trials and three sequences for identifying fruits based on three color spaces those are RGB color space trial, HSV color space trial and LUV color space trial. This research conducted to explain scientifically based on the data training. The purpose of each trial is to achieve the best success rate. Then, arrange each dataset depends on the color spaces and implement it into real robot vision on-farm for future concept.

2.2. Custom Datasets

Images of tomatoes used in training and testing were collected from open-source image from internet and manually captured using phone camera on the tomato farm field located in West Java, Indonesia. The total image of the data was obtained and store basically as RGB format. The datasets for both trial and validation

containing many numbers of samples which consist of single tomato for either red or green or orange tomato, group of tomato with the same color, multiple mixed color type of tomato. In addition, the datasets also consist of multiple small sized images for either red or green or orange tomato. The small sized images used to recognize fruits with far distance of camera or the fruits with a tiny shape. To increase the diversity of the samples, the datasets also contain multiple image characteristics such as blur, dim, front-facing, upper-part, bright, and overlapping objects.

The datasets consist of three classes those are *ripe_tomato*, *raw_tomato*, and *half_ripe_tomato*. The *ripe_tomato* represent to mature fruit, the *raw_tomato* represent to unripe fruit, and the *half_ripe_tomato* represent the fruit is half mature [10]. The datasets also distinguished based on the purpose those are train images and validation images. The total sample images of trial and validation for *ripe_tomato* is 192 and 136 respectively, the total samples of trial and validation for *raw_tomato* class are 243 and 101 samples respectively, and the total samples of trial and validation for *half_ripe_tomato class* are 403 and 139 respectively. The total samples from 334 images datasets are approximately 838 samples for trial and 376 samples for validation. Each class basically is RGB color space and converted to another color spaces HSV and LUV. These variety of class and color spaces appointed to determine the best color space to detect fruits maturity level.

3. Results and Discussion

3.1. RGB Trial

This trial consists of several tests and each test has its main purpose. The first test purposed the machine vision to detect fruits and classified them into two classes *ripe_tomato and raw_tomato.* However, the test result has low quality mean Average Precision (mAP50) for only 54% and 52% respectively with total 49%. These issues happened because the machine vision cannot detect small-sized fruits for further scope of the images. In the second trial, more images added to the dataset, specifically for small-sized tomato for either red or green tomato class with total additional datasets are 90 images for both. The additional datasets including single small sized red and green tomato, and multiple objects of small-sized tomato. The results of the second trial for *ripe_tomato* and *raw_tomato* class have mAP50 92% and 83% with overall mAP50 is 87%. However, in the second trial, the machine cannot detect orange tomato which is neither categorized as red nor green tomato. Thus, the final trial conducted to add more images and one dataset labelled as *half_ripe_tomato*. The total images added for the *half_ripe_tomato* class are 133 images containing single tomato, multiple tomato, multiple mixed color tomato, and the small-sized tomato. The total additional images are 133, 328 images for overall class, and 823 samples in total for trial and 360 samples in total for
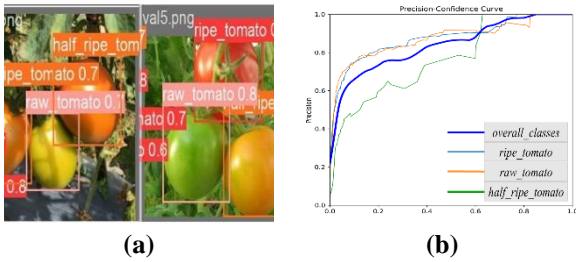
(a)   **(b)**
Fig. 2. (a) Sample results RGB, and (b) P-C Curves.
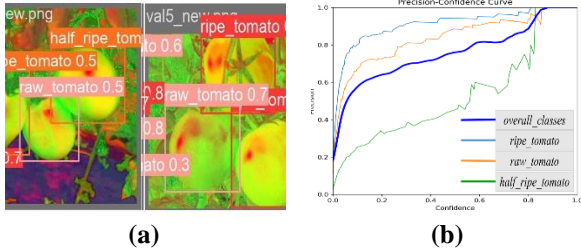
(a)   **(b)**
Fig. 3. (a) Sample results HSV, and (b) P-C Curves.
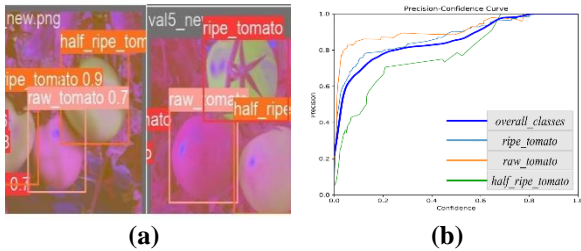
(a)   **(b)**
Fig. 4. (a) Sample results LUV, and (b) P-C Curve.

validation. The sample images of final results shown on the Fig. 2(a) and Precision-Confidence Curve (P-C Curve) shown on the Fig. 2(b).

3.2. HSV Trial

HSV trial basically used the same number of total datasets with RGB datasets. In the HSV trial, the images converted into HSV images using Python code. The HSV trial also use three classes called as *ripe_tomato, raw_tomato,* and *half_ripe_tomato*. The trial test of this HSV datasets conducted in 100 epochs. Brief results of the HSV trial are the machine vision cannot differentiate *raw_tomato* and *half_ripe_tomato* class. This issue happened because the HSV color space conversion generates those tomato's skin look exactly the same. The sample images of final results shown on the Fig. 3(a) and Precision-Confidence Curve (P-C Curve) shown on the Fig. 3(b).

3.3. LUV Trial

LUV trial basically used the same datasets from the converted RGB color space to LUV color space using Python code. The LUV datasets also consist of three classes *ripe_tomato, raw_tomato,* and *half_ripe_tomato*. The Trial test conducted in 100 epochs with the detail information of sample final results shown on the Fig. 4(a) and Precision-Confidence Curve (P-C Curve) shown on the Fig. 4(b).

3.4. Performance Evaluation from All Datasets

Based on this research, there are many advantages and disadvantages to use multiple color space in a dataset. Each color space has specific purpose to be used in machine vision. HSV color space can be used to detect color-based image. HSV separates luma, or the image intensity from chroma or color information. HSV can be useful for histogram equalization, removing shadows, lightning disturbance, etc. Besides, RGB color is the basic color space in every image and captured by camera. RGB color can be easily to use and formulated because this color space not necessary to be converted from the original image. However, RGB color space might be also easily distracted with external factors such as shadow, lightning, sunrays, etc. Meanwhile the LUV color space can be used to decouple color with the U and V components which represent chromaticity values of color image. Besides the L or Luminance can improve the RGB picture which distracted from over light disturbance, etc. The comparison from three color spaces which used in this research shown in Table 1.

Based on the Table 1, the LUV and RGB color space has similar mAP50 rates with the overall value 85% and 83% respectively. On the other hand, the LUV datasets have a good performance in object detection proven by the mAP50 rate shown on the Table 1 because LUV color space are fundamentally suitable for restoration error estimation, fast detection, and impulsive noise removal in real-time color imaging. Based on the Table 1, *ripe_tomato* class and *raw_tomato* class for LUV color space's mAP50 rate is higher than RGB color space because some of the images in RGB overlapped by shadows and light as shown in the Fig. 6(a) and Fig. 5(a) respectively. The explanation of comparison between RGB and LUV datasets also visualized in P-C Curve as shown in the Fig. 2(b) and Fig. 4(b) which shows the RGB curve relatively fluctuating than LUV graph. Nevertheless, the graph shows a great stability and raise gradually with small fluctuations. In contrary, the HSV has the lowest of all datasets training. The HSV datasets conversion from RGB color space has the typical color tone with its backgrounds such as leaf, trees, and grass. The HSV datasets, either *ripe_tomato* or *raw_tomato* class has a high mAP50 rates for more than 70%. However, there is an issue regarding the *half_ripe_tomato* class, the machine vision cannot detect the object significantly resulting only 57% of mAP50 value as shown on the Table 1. This value is nearly with 50% and the results should be evaluated to increase the rates [11]. The P-C curves Fig. 3(b) showing the relatively unarranged graph for the *half_ripe_tomato* and the prediction line increased gradually.

The shadow and lightness disturbances occurred during the dataset collection in this research as shown in Fig. 5 and Fig. 6. RGB image reflect lightness, such as sunrays,

Reno Muhammad Fadilla, Tokuo Tsuji, Tatsuhiro Hiramitsu, Hiroaki Seki, Kaihei Okada

Table 1. Final results comparison.

Class	Total Samples		mAP50 (*Epochs = 100*)		
	Trial	Validation	RGB	HSV	LUV
ripe_tomato	192	136	88%	83%	95%
raw_tomato	243	101	81%	77%	84%
half_ripe_tomato	403	139	80%	57%	77%
Overall	838	376	83%	72%	85%

(a) **(b)** **(c)**
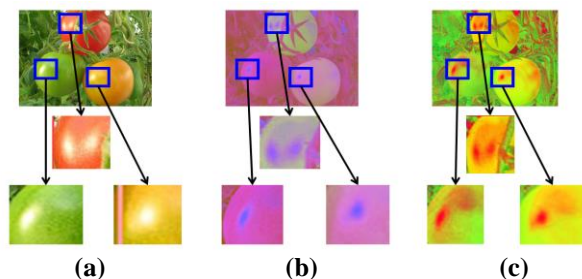
Fig. 5. Lightness disturbance: (a) RGB, (b) LUV, (c) HSV.

(a) **(b)** **(c)**
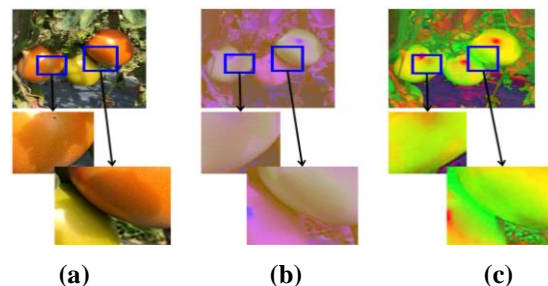
Fig. 6. Shadow disturbance: (a) RGB, (b) LUV, (c) HSV.

into white spot, whereas LUV image converts the white spot into warm blue color as shown in the Fig. 5(b) which was not distract the real color of the image and affecting the machine vision detection more precisely. HSV images turn the white spot into dark red color as shown in Fig. 5(c) and affect the real color of the images. HSV images turn the shadow into lighter color compared with the basic RGB image as shown in the Fig. 6(c) and Fig. 6(a) consecutively. This result influence the machine vision to detect object more precisely. Moreover, LUV also turn the shadows into lighter color than RGB shadow, however, the shadows still configurable and not affected much to fruit's basic color as shown on the Fig. 6(b). Basically, YOLOv5 model trained using a pre-training model originally from RGB color which trained in another color space LUV and HSV. Hence, the model generated results in relatively similar to RGB color space. Overall, LUV color space was the greatest datasets for tomato classification. Therefore, the LUV can be suitable for harvesting robot despite each color space has an advantages and disadvantages in certain occasions and purposes.

4. Conclusion

This research has shown the model of ML which used to determine the best datasets differentiated from color space RGB, HSV, and LUV. According to the final results, RGB and LUV datasets generate an outstanding typical result for the machine vision, on the contrary, the HSV datasets obtain a relatively poor result. The best class in this research from overall datasets evaluation were *ripe_tomato* and *raw_tomato* class. Meanwhile, the *half_ripe_tomato* class acquire intermittent results for all datasets color space depends on the backgrounds and external distraction from the objects. Moreover, HSV datasets need some additional images and adjustment of the "Hue" value as required for increasing the success

rate. In conclusion, all of the datasets in this research can be used depending on the purposes, and object's circumstances which correlated with external disturbances such as backgrounds, lightness, and shadows of the images. LUV color space has shown its performance as the best datasets for rapid segmentation of tomato on farm and more suitable in many external disturbances such lightness and shadows. On the other hand, HSV datasets can be used to recognized tomatoes on farm which has high light intensity and shadow disturbance. Therefore, this color space can be more suitable for harvesting robot with rapid segmentation.

References

[1] E. Mavridou, E. Vrochidou, G. A. Papakostas, T. Pachidis, and V. G. Kaburlasos, "Machine vision systems in precision agriculture for crop farming," Journal of Imaging, vol. 5, no. 12. 2019. doi: 10.3390/jimaging5120089.

[2] W. Jia, Y. Tian, R. Luo, Z. Zhang, J. Lian, and Y. Zheng, "Detection and segmentation of overlapped fruits based on optimized mask R-CNN application in apple harvesting robot," Comput Electron Agric, vol. 172, 2020, doi: 10.1016/j.compag.2020.105380.

[3] A. Bhargava and A. Bansal, "Fruits and vegetables quality evaluation using computer vision: A review," Journal of King Saud University - Computer and Information Sciences, vol. 33, no. 3. 2021. doi: 10.1016/j.jksuci.2018.06.002.

[4] N. Ismail and O. A. Malik, "Real-time visual inspection system for grading fruits using computer vision and deep learning techniques," Information Processing in Agriculture, vol. 9, no. 1, 2022, doi: 10.1016/j.inpa.2021.01.005.

[5] H. Rubesh, D. Hema, and S. Kannan, "Interactive Color Image Segmentation using HSV Color Space," Science and Technology Journal, vol. 7, p. 1, 2020, doi: 10.22232/stj.2019.07.01.05.

[6] H. Li et al., "An automatic approach for detecting seedlings per hill of machine-transplanted hybrid rice

utilizing machine vision," Comput Electron Agric, vol. 185, 2021, doi: 10.1016/j.compag.2021.106178.

[7] T. Li et al., "Tomato recognition and location algorithm based on improved YOLOv5," Comput Electron Agric, vol. 208, 2023, doi: 10.1016/j.compag.2023.107759.

[8] D. Pietro Cavallo, M. Cefola, B. Pace, A. F. Logrieco, and G. Attolico, "Non-destructive and contactless quality evaluation of table grapes by a computer vision system," Comput Electron Agric, vol. 156, 2019, doi: 10.1016/j.compag.2018.12.019.

[9] A. Septiarini, A. Sunyoto, H. Hamdani, A. A. Kasim, F. Utaminingrum, and H. R. Hatta, "Machine vision for the maturity classification of oil palm fresh fruit bunches based on color and texture features," Sci Hortic, vol. 286, 2021, doi: 10.1016/j.scienta.2021.110245.

[10] Ifmalinda, Andasuryani, and I. Rasinta, "Classification of tomato (Lycoersicon Esculentum Miil) ripeness levels based on HSV color using digital image processing," in IOP Conference Series: Earth and Environmental Science, 2022. doi: 10.1088/1755-1315/1116/1/012062.

[11] Q. Zheng, Z. Peng, P. Zhu, Y. Zhao, R. Zhai, and W. Ma, "An Object Recognition Grasping Approach Using Proximal Policy Optimization with YOLOv5," IEEE Access, vol. 11, 2023, doi: 10.1109/ACCESS.2023.3305339.

Authors Introduction

Mr. Reno Muhammad Fadilla

He is currently a Master Student in the Division of Frontier Engineering, Kanazawa University, Japan since 2023 with research theme in the fields of computer vision, and robotics. He received his Bachelor of Engineering from UIN Bandung, Indonesia in 2022.

Dr. Tokuo Tsuji

He is currently an Assoc. Professor in the Faculty of Frontier Engineering, Kanazawa University with research area in the fields of informatics, and intelligent robotics. He received his Doctoral Degree in Engineering.

Dr. Tatsuhiro Hiramitsu

He is currently an Assistant Professor in the Division of Frontier Engineering, Kanazawa University, Japan since 2019 with a research area in the fields of soft robotics and intelligent robotics. He received his Doctoral Degree in Engineering.

Dr. Hiroaki Seki

He is currently a Professor in the Faculty of Frontier Engineering, Kanazawa University, specialty in the fields of robotics and mechatronics automation. He received his Doctoral Degree in Engineering from The University of Tokyo in 1994.

Mr. Kaihei Okada
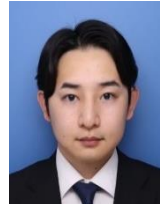
He is currently a master student in the Division of Frontier Engineering, Kanazawa University, Japan since 2023 with research theme in the fields of computer vision and robotics. He received his Bachelor Degree in Engineering from Kanazawa University, Japan.

YOLO real-time object detection on EV3-Robot using FPGA hardware Accelerator

Dinda Pramanta

Kyushu Institute of Information Sciences, 6-3-1, Saifu, Dazaifu, 818-0117, Japan
E-mail: pramanta@kiis.ac.jp

Ninnart Fuengfusin, Arie Rachmad Syulistyo, Hakaru Tamukoh

Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku, Kitakyushu, 808-0196, Japan
E-mail: ninnart@brain.kyutech.ac.jp, syulistyo.arie-rachmad788@mail.kyutech.jp, tamukoh@brain.kyutech.ac.jp

Abstract

The growing demand for robots necessitates faster and more precise processing. However, running large Artificial Intelligence (AI) models from cloud data centers to mobile robots via inference models uses considerable computation resources, which leads to power limitations, particularly for mobile robots. The use of reconfigurable semiconductor devices at the hardware level is a promising solution to this problem. We introduce the educational kit EV3-Robot with a co-design methodology utilizing Field-programmable Gate Arrays (FPGA) Kria KV260 as a hardware accelerator specifically for object detection. We apply the You Only Look Once (YOLO) model for object detection, which provides real-time results for practical applications. Additionally, we analyze the processing times of the local PC and EV3-Robot.

Keywords: EV3-Robot, FPGA, Object detection, YOLO.

1. Introduction

Japan has an ageing population and a decreasing number of young people [1]. To address this issue, mobile robots have gained attention [2]. For example, robots are expected to perform daily indoor tasks, such as picking up and delivering objects. To accomplish such tasks, robots must be capable of recognizing and detecting specific objects [3].

In the field of Artificial Intelligence (AI), Neural Networks (NN) have been extensively researched [4], [5], [6]. Neural networks (NN) are equipped with state-of-the-art perceptual abilities, adaptive learning, and sophisticated human interaction. However, processing this information requires fast and accurate computation, especially in mobile robots where real-time interaction is almost unavoidable. This highlights the increasing need for robots equipped with reliable NNs. In contrast, You Only Look Once (YOLO) has gained attention for its fast and high accuracy [7].

Reconfigurable semiconductor devices at the hardware level offer a promising solution to the challenge of running large YOLO models from cloud to mobile robots via inference models, which require significant computation resources and result in low power consumption. We present the educational kit EV3-Robot, which utilizes a co-design methodology that employs local-PC and Field-programmable Gate Arrays (FPGA) Kria KV260 as a hardware accelerator for the YOLO model.

In our previous study on hardware-based communication [8], [9], the communication processes played a pivotal role. The first step was to design the client-server communication model, where the EV3-Robot served as the server and the local PC as the client. Next, we set the feedback for motor movements during specific tasks, in this case, using Ev3dev Python Socket Connection via Bluetooth to be able detect humans as objects. After that, we execute the deep-learning processor unit (DPU) of the YOLO model on the KV260 board. Finally, we evaluate the processing time.

2. Methodology

2.1. *You Only Look Once version 3 (YOLOv3) as detection system*
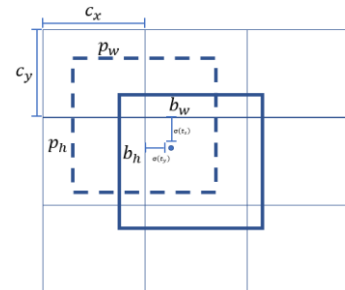
Fig.1 Bounding box [10]

Fig. 1 illustrates how YOLOv3 predicts the anchor box by using the bounding box and size clustering. The feature map cell is represented by c_x, c_y, and the pre-selected bounding box size is p_w, p_h. The predicted coordinates are b_x, b_y, b_w, b_h which can be calculated

Dinda Pramanta, Ninnart Fuengfusin, Arie Rachmad Syulistyo, Hakaru Tamukoh.

through Eq. (1), (2), (3), and (4) where the $\sigma(t_x)$ and $\sigma(t_y)$ are pixel values.

$$b_x = \sigma(t_x) + c_x \qquad (1)$$
$$b_y = \sigma(t_y) + c_y \qquad (2)$$
$$b_w = p_w e^{t_w} \qquad (3)$$
$$b_h = p_h e^{t_h} \qquad (4)$$

2.2. *System architecture*
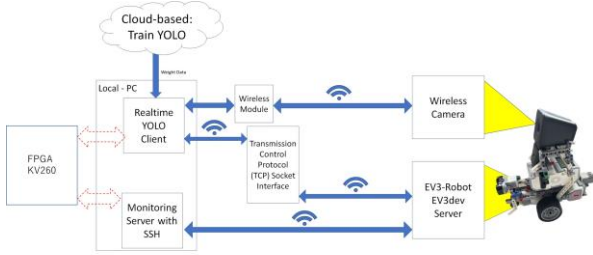
Fig. 2. YOLO real-time object detection on EV3-Robot using FPGA hardware accelerator.

Fig. 2 displays a real-time implementation of YOLO for object detection using an EV3-Robot. The EV3-Robot is equipped with various sensors, although we exclude the input sensor functions in this system. The operating system (OS) used is ev3dev, which is Linux-based and stored on a Secure Digital (SD) card. The connection to the computer is established through Bluetooth and operates at 2.4GHz. The robot is powered by an ARM-9 processor and includes a built-in mini-LCD display. It has four input ports for sensor functions and four output ports for actuators or motors. The EV3-robots motor is connected using the output ports only for feedback.

2.3. *Ev3dev Python Socket Connections for* **feedback**
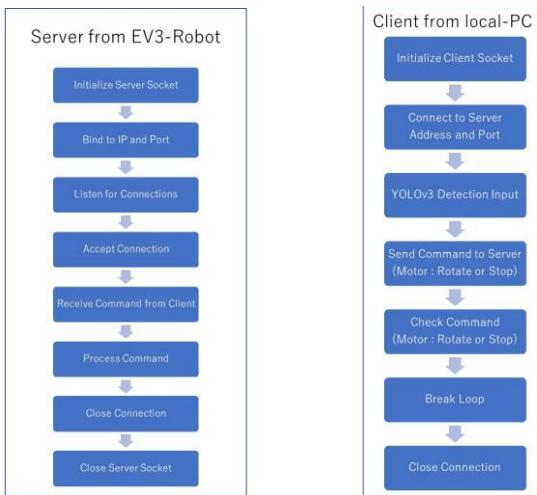
a.   Server diagram flow   b.   Client diagram flow
Fig. 3. Client-server communication systems.
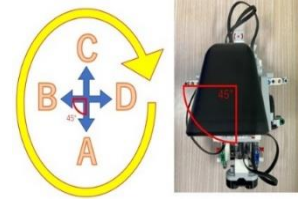
a.   Clockwise rotating for scanning "person" and "A" is an initial direction.
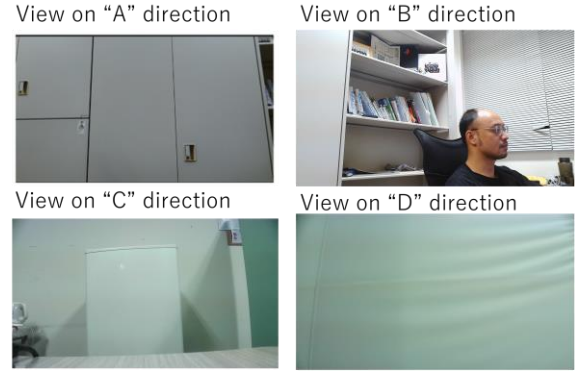
b.   Camera views on each direction
Fig. 4. EV3-Robot motor rotation
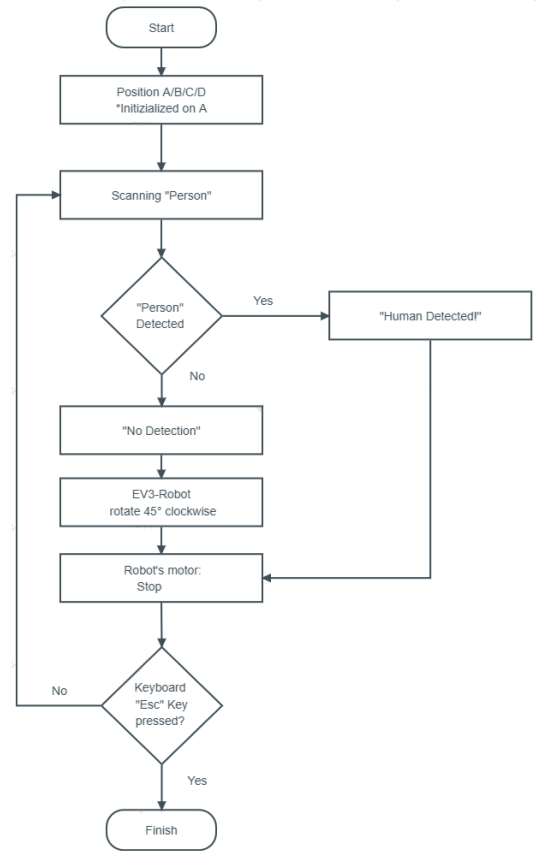
Fig. 5. Feedback sequences.

Fig. 3 illustrates a remote-control flow that utilizes python socket connections [11] to set motor movements as feedback. The communication involves the use of Internet Protocol (IP) addresses to establish connections. The local PC acts as a client and sends commands to an EV3-Robot, which serves as the server, enabling control over the robot.

Fig. 4 depicts a feedback design for detection that utilizes the motor functions of the EV3-Robot. The surrounding environment is scanned by setting up four directional points, denoted as "A", "B", "C", and "D". The direction is rotated 45° clockwise at regular intervals, in this case, every 4 seconds.

Fig. 5 displays the sequence of the motor rotation process. The robot is equipped with a standby mode, referred to as 'scanning for person', which enables it to scan while executing real-time object detection processes until the desired class, 'person', is detected. Finally, we have implemented an exit process triggered by the 'Esc' command from keyboard inputs to conclude the sequence.

2.4. *Getting the weight data from cloud to local-PC*

The pre-trained COCO weight dataset [12] was utilized. As we were testing real-time communication during detection, we only focused on the specified class from COCO, which is "person" (class_id = 0).

This was added to the client part and a real-time processing loop was performed for 'person' object detection. The wireless camera is mounted to perform object detection using the YOLOv3 model. It focuses on specific classes and determines whether a class_id = 0 (person) is detected. If so, it sends a message via the socket connection.

2.5. *Deep-learning Processor Unit (DPU) in KV260*

Fig. 6. DPU flow.

Fig. 7. YOLOv3 xmodel.

Fig. 6 displays the application of DPU and its integration into the KRIA KV260 board FPGA. Initially, we prepare the YOLOv3 model [13]. Next, we proceed to the pre-processing stage to install and verify the dependencies or libraries. Subsequently, we quantize the model from 32-bit floating-point weights to 8-bit.

Fig. 7 shows that we successfully generated the specified model for the KV260 architecture model, referred to as the "xmodel". Finally, we load "xmodel" into DPU deployment for real-time hardware acceleration.

3. Results and Discussion

3.1. *Real-time remote-control communication of client-server systems*

Table 1. The latency results of the first four command keys.

Keyboard Key	Message	Latency (*ms*)
Up	"going UP! received"	522
Right	"going RIGHT! received"	652
Down	"going DOWN! received"	511
Left	"going LEFT! received"	661

Latency is a priority in client-server communications, especially for YOLO applications that require real-time interaction and responsiveness from the EV3-Robot. Achieving low-latency communication is crucial for current system, so we define latency as the round-trip time (RTT) divided by two in milliseconds (*ms*) as shown in (5).

$$Latency\ (ms) = \frac{RTT}{2} \qquad (5)$$

The outcomes of the initial four command keys are displayed in Table1. Each key is associated with a confirmation message, allowing us to verify when the instruction was issued. Our findings indicate that the latency remains below $1,000\ ms$.

Fig.8. FPGA hardware accelerator task

Dinda Pramanta, Ninnart Fuengfusin, Arie Rachmad Syulistyo, Hakaru Tamukoh.
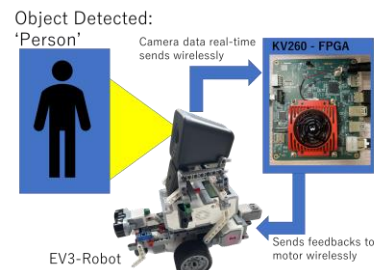
Table 2. Result comparison during N-step = 1.

Local-PC	FPS	Inference-time (*ms*)	Hardware Accelerator	"person" Object Detection
13th Gen Intel(R) Core (TM) i7-13700F 2.10 GHz.	2.63	270.3	None	Success
	3.78	264.4	FPGA KV260	Success

3.2. *Real-time visualizations*
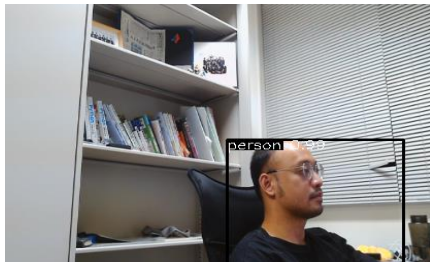
a. Local-PC without hardware accelerator.
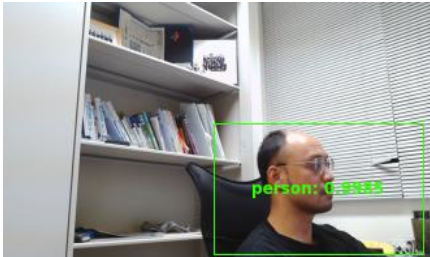
b. Local-PC with hardware accelerator FPGA.
Fig. 9. EV3-Robot real-time visualizations

Fig. 8 shows a task that FPGA hardware accelerator task. Fig. 9(a) and Fig. 9(b) are displaying the real-time visualization from the EV3-Robot's perspective. The class detection for 'person' was achieved in the "B" direction.

3.3. *FPGA hardware accelerator*

Table 2 compares the local-PC and EV3-Robot's performance during the detection task. The use of hardware accelerators resulted in an increase in the number of frames per second (FPS) and a reduction in inference time.

4. Conclusion

Real-time object detection on the EV3-Robot using FPGA hardware acceleration has been successfully implemented. Equipped with the FPGA Kria KV260 accelerator and the YOLO model, the EV3-Robot demonstrates a practical and effective solution for real-time object detection. This paves the way for enhanced performance and broader applications in the field of mobile robotics.

References

1. Coulmas, F., 2007. Population decline and ageing in Japan-the social consequences (Vol. 16). Routledge.
2. Yoshimoto, Y. and Tamukoh, H., 2021. FPGA implementation of a binarized dual stream convolutional neural network for service robots. Journal of Robotics and Mechatronics, 33(2), pp.386-399.
3. Hori, S., Ishida, Y., Kiyama, Y., Tanaka, Y., Kuroda, Y., Hisano, M., Imamura, Y., Himaki, T., Yoshimoto, Y., Aratani, Y. and Hashimoto, K., 2017. Hibikino-musashi@ home 2017 team description paper. arXiv preprint arXiv:1711.05457.
4. Hinton, G., 2018. Deep learning—a technology with the potential to transform health care. Jama, 320(11), pp.1101-1102.
5. LeCun, Y., Bengio, Y. and Hinton, G., 2015. Deep learning. nature, 521(7553), pp.436-444.
6. Paugam-Moisy, H. and Bohte, S.M., 2012. Computing with spiking neuron networks. Handbook of natural computing, 1, pp.1-47.
7. Redmon, J., Divvala, S., Girshick, R. and Farhadi, A., 2016. You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).
8. Pramanta, D., Morie, T. and Tamukoh, H., 2017. Synchronization of Pulse-Coupled Phase Oscillators over Multi-FPGA Communication Links. J. Robotics Netw. Artif. Life, 4(1), pp.91-96.
9. Pramanta, D. and Tamukoh, H., 2019, December. High-speed synchronization of pulse-coupled phase oscillators on multi-FPGA. In International Conference on Neural Information Processing (pp. 318-329). Cham: Springer International Publishing.
10. Redmon, J. and Farhadi, A., 2018. Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767.
11. R. Hempel and D. Lechner, "ev3dev," 2023. Available online: https://www.ev3dev.org/, accessed on 30 January 2023. [Online]. Available: https://www.ev3dev.org/
12. Lin, T.Y., et al, 2014. Microsoft coco: Common objects in context. In Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13 (pp. 740-755). Springer International Publishing.
13. Darknet model, accessed on 30 January 2023. [Online] Available: https://pjreddie.com/darknet/yolo

Authors Introduction

Dr. Dinda Pramanta

He earned his Bachelor's degree from Telkom University in 2013 and completed both Master's and Doctoral degrees at the Kyushu Institute of Technology in 2016 and 2020. After a post-doctoral fellowship with NEDO, Japan, from 2020 to 2021, he is currently an Assistant Professor at Kyushu Institute of Information Sciences since 2021. His research interests include neural networks, hardware, and AI for educational applications.

Dr. Ninnart Fuengfusin

He earned his B.Eng. degree from King Mongkut's University of Technology Thonburi, Thailand, in 2016, and his M.Eng. and D.Eng. degrees from Kyushu Institute of Technology, Japan, in 2018 and 2021, respectively. Currently a postdoctoral researcher at the Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology, his research focuses on deep learning, efficient neural network design, and digital hardware design.

Mr. Arie Rachmad Syulistyo

A member of IEEE, he earned his bachelor's degree in computer science from the University of Brawijaya, Indonesia, in 2011, and his master's degree from the University of Indonesia in 2015. Currently pursuing a Ph.D. in life science and system engineering at Kyushu Institute of Technology, Japan, his research interests encompass computer vision and machine learning.

Dr. Hakaru Tamukoh

A member of IEEE, he obtained his B.Eng. degree from Miyazaki University in 2001 and completed his M.Eng. and Ph.D. degrees at the Kyushu Institute of Technology in 2003 and 2006, respectively. After serving as a Postdoctoral Research Fellow, he became an Assistant Professor at the Tokyo University of Agriculture and Technology from 2007 to 2013. Currently a Professor at the Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology, his research interests encompass digital hardware design, soft computing, and home-service robots.

A Low Computational Cost Hand Waving Action Recognition System
with Echo State Network for Home Service Robots

Hiromasa Yamaguchi
School of Computer Science and Systems Engineering, Kyushu Institute of Technology,
680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan

Akinobu Mizutani
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu 808-0196, Japan

Arie Rachmad Syulistyo
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu 808-0196, Japan

Yuichiro Tanaka
Research Center for Neuromorphic AI Hardware, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu, 808-0196, Japan

Hakaru Tamukoh
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu 808-0196, Japan

E-mail: yamaguchi.hiromasa512@mail.kyutech.jp
https://www.brain.kyutech.ac.jp/~tamukoh/

Abstract

This study proposes a low computational cost hand-waving action recognition system for non-verbal communication in home service robots. The system is based on an echo state network, which requires lower computational costs than that of deep neural networks (DNNs), and processes time-series data of skeletal coordinates of humans to recognize hand-waving actions. Additionally, this study proposes and compares two types of Preprocessing ing methods of the skeletal coordinates to ensure the robustness of the human positions on the frame: one method extracts shoulder and arm angles, which are invariable regardless of the humans' positions and the other normalizes the skeletal coordinates. The experimental result shows that the proposed system has competitive accuracy and is robust to varying human positions.

Keywords: Action recognition, Home service robot, Low computational cost, Echo state network

1. Introduction

The declining birthrate and aging population have become social problems in many countries, not only Japan, and one of the solutions is the use of home service robots [1]. For this reason, research on home service robots has been active [2], [3], [4]. Home service robots can expand the range of their use by acquiring not only verbal but also non-verbal information such as human facial expressions and movements. In addition, since home service robots are required to judge situations and perform actions in real-time, computers for their information processing should be mounted on the home service robot. Therefore, the resources in the computers are limited and the computational cost of the process must be low.

In this study, we focus on hand-waving action recognition, which indicates that a human is calling the home service robot and aim to construct a system with low computational cost. The proposed system recognizes hand-waving action by providing skeletal information extracted from videos using MediaPipe [5] to an echo state network (ESN) [6], a lightweight machine learning model.

Hiromasa Yamaguchi, Akinobu Mizutani, Arie Rachmad Syulistyo, Yuichiro Tanaka, Hakaru Tamukoh
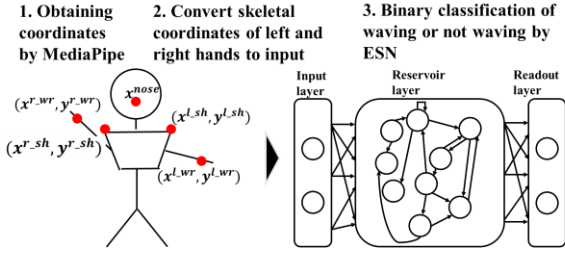
Fig. 1. Process flow of the proposed hand-waving
action recognition.

Table 1. Experimental results

	Ours (A)	Ours (B)	Slow Only	UniFormer V2
Accuracy	0.667	**0.771**	0.625	0.646
Precision	0.875	**0.928**	0.667	0.622
Recall	0.500	0.683	0.714	**1.000**
processing time[s]	3.445	**3.446**	20.519	29.607

Table 2. Accuracy per subject for the proposed
method (Preprocessing B)

Subject	S1	S2	S3	S4
Accuracy	0.813	0.863	0.775	0.662

2. Related Works

MMAction2 [7], provided by OpenMMLab, is a library for action recognition. This library estimates human behaviors such as handshaking, hugging, hand waving, etc. The tool offers powerful deep neural network (DNN)-based models for action recognition such as SlowFast [8], UniFormerV2 [9], and VideoMAE V2 [10], however, the three models listed as examples have a huge number of parameters (more than 20M) and are computationally expensive.

While the methods [8], [9], [10] directly feed input images into DNNs, Doan proposed an action recognition system with a long short-term memory model that processes skeletal information [11]. Compared to the end-to-end processing models, the method is more lightweight but still requires high computational costs because it is based on the DNN.

3. Proposed Methods

We propose a lightweight hand-waving action recognition system with MediaPipe [5] and an ESN [6]. The processing flow of the proposed method is shown in Fig. 1. For each video frame, the system extracts skeletal coordinates by using MediaPipe The skeletal coordinates are then given to an ESN, which performs a binary classification of whether the hand is waving or not. Before feeding the coordinates to an ESN, we apply preprocessing to skeletal coordinates to classify the action not depending on the human position on the frame. We propose and compare two types of preprocessing in this study.

Preprocessing A extracts the angles of the shoulders and arms. Preprocessing A shown in Eq. (1) and (2) use the *x*-coordinate of the left and right wrists (x^{l_wr}, x^{r_wr}), the left and right shoulder (x^{l_sh}, x^{r_sh}), and the y-coordinate of the left and right wrists (y^{l_wr}, y^{r_wr}), and the left and right shoulder (y^{l_sh}, y^{r_sh}) respectively. X^{in} in Eq. (3) is the input to an ESN.

$$X^{l\_in} = \tan^{-1} \frac{(y^{l\_wr} - y^{l\_sh})}{(x^{l\_wr} - x^{l\_sh})} \quad (1)$$

$$X^{r\_in} = \tan^{-1} \frac{(y^{r\_wr} - y^{r\_sh})}{(x^{r\_wr} - x^{r\_sh})} \quad (2)$$

$$X^{in} = \begin{pmatrix} X^{l\_in} \\ X^{r\_in} \end{pmatrix} \quad (3)$$

Preprocessing B shown in Eq. (4) and (5) use the *x*-coordinate of the left and right wrists (x^{l_wr}, x^{r_wr}), the nose (x^{nose}), and the left and right shoulder (x^{l_sh}, x^{r_sh}), respectively. The coordinates of the wrists are relative to the coordinates of the nose to ensure robustness against parallel movement of the recognition target. In addition, the wrist coordinates are normalized by the length of the shoulder. The robustness against human size in the frame is ensured by normalizing the length of the shoulder. X^{in} in Eq. (6) is the input to an ESN.

$$X^{l\_in} = \frac{(x^{nose} - x^{l\_wr})}{|x^{l\_sh} - x^{r\_sh}|} \quad (4)$$

$$X^{r}_{in} = \frac{(x^{nose} - x^{r\_wr})}{|x^{l\_sh} - x^{r\_sh}|} \quad (5)$$

$$X^{in} = \begin{pmatrix} X^{l\_in} \\ X^{r\_in} \end{pmatrix} \quad (6)$$

4. Experiments

4.1 Setup

We collected a total of 360 video files of actions from six persons, of which 240 data points were used as training data from one person, 24 data points were used as validation data from another person, and 96 data points were used as test data from four persons. The CPU of the computer that performed the processing was a CPU (Intel (R) Core (TM) i7-6700K).

To verify whether the proposed method can be a substitute for the prior cases, we conducted similar tests using the behavior recognition models of the prior cases (however, these models can recognize multiple

behaviors). We used the SlowOnly [8] and the UniFormerV2 [9] models provided by the MMAction2 model trained on the Kinetics-700 dataset [12] of 700 behaviors to predict the behaviors of the test data.

4.2 Result

The averages of the accuracy, precision, recall rates, and processing time for the ten estimation results for the test data are shown in Table 1. The proposed method has a higher accuracy rate in hand-waving action recognition than the other two methods. In addition, we can see that Preprocessing B is superior to A in all indices. The proposed method also has a shorter processing time than the other two methods.

Table 2 shows the accuracy for each subject using the proposed method with Preprocessing B, which had the highest accuracy. These results show that the accuracy varies greatly depending on the subject.

5. Discussion

We consider the Preprocessing B shown in Eq. (4), (5) and (6) with normalization to the wrist coordinates to be more accurate than the Preprocessing A using the angles shown in Eq. (1), (2) and (3), because the rate of fluctuation of the values is larger. In particular, the rate of fluctuation of the values for Preprocessing A is small when the elbow is bent while hand waving. Therefore, we consider that Preprocessing B is superior.

The proposed method has a low recall for Preprocessing B. In Preprocessing B, we consider the influence of the data from the wrist-fixed hand wave to be significant. Preprocessing B makes predictions based on fluctuation in wrist coordinates. Therefore, if the hand is waving without wrist fluctuation, it cannot be recognized. Therefore, it is necessary to devise a system that allows recognition without wrist fluctuation.

As shown in Table 2, variations in accuracy were observed among the subjects. This is expected to be caused by the body size and waving habits of the subjects. It is necessary to verify whether this problem can be solved by using data from multiple subjects in the training data in the future.

6. Conclusion

In this study, we proposed a hand-waving action recognition system based on an ESN to construct a low-computational-cost hand-waving action recognition system to be introduced into home service robots. The results show that the accuracy of the lightweight model is comparable to that of deep learning-based multiple-action recognition models.

In the future, it will be necessary to construct a system that can recognize a waving person even if there are multiple people in the frame because there is a possibility that multiple people may be in the image in the real world.

Acknowledgment

References

1. T. Yamamoto, K. Terada, A. Ochiai, F. Saito, Y. Asahara and K. Murase, Development of Human Support Robot as the research platform of a domestic mobile manipulator, Robomech Journal, Vol. 6, No. 4, 2019.
2. K. Isomoto, Y. Yano, Y. Tanaka, and H. Tamukoh, Robust trash can lid opening system, Proceedings of the 2023 International Workshop on Smart Info-Media Systems in Asia (SISA), 2023.
3. T. Shiba, T. Ono and H. Tamukoh, Object Search and Empty Space Detection System for Home Service Robot, Proceedings of the 2023 International Conference on Artificial Life and Robotics (ICAROB2023), 2023.
4. A. Mizutani, Y. Tanaka, H. Tamukoh, K. Tateno, O. Nomura and T. Morie, A knowledge acquisition system with a large language model and a hippocampus model for home service robots, Proceeding of the Institute of Electronics, Information and Communication Engineers (IEICE) Rep. vol. 123, no. 208, 2023, pp. 13–18.
5. C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C. Chang, M. G. Yong, J. Lee, Wan-Teh Chang, W. Hua, M. Georg and M. Grundmann, MediaPipe: a framework for building perception pipelines, 2019.
6. H. Jaeger: Short term memory in echo state networks, GMD Report 152, 2002.
7. MMAction2 Contributors, OpenMMLab's Next Generation Video Understanding Toolbox and Benchmark, https://github.com/open-mmlab/mmaction2, 2020.
8. C. Feichtenhofer, H. Fan, J. Malik and K. He, SlowFast Networks for Video Recognition, Proceedings of the IEEE international conference on computer vision, 2019, pp. 6202-6211.
9. K. Li, Y. Wang, Y. He, Y. Li, Y. Wang, L. Wang and Y. Qiao, UniFormerV2: Spatiotemporal Learning by Arming Image ViTs with Video UniFormer, 2022.
10. L. Wang, B. Huang, Z. Zhao, Z. Tong, Y. He, Y. Wang, Y. Wang, Y. Qiao, VideoMAE V2: Scaling Video Masked Autoencoders with Dual Masking, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023, pp.14549-14560.
11. T. Doan, An Efficient Patient Activity Recognition using LSTM Network and High-Fidelity Body Pose Tracking, International Journal of Advanced Computer Science and Applications (IJACSA), VoL 13, 2022.
12. L. Smaira, J. Carreira, E. Noland, E. Clancy, A. Wu and A. Zisserman, A Short Note on the Kinetics-700-2020 Human Action Dataset, 2019.

Hiromasa Yamaguchi, Akinobu Mizutani, Arie Rachmad Syulistyo, Yuichiro Tanaka, Hakaru Tamukoh

Authors Introduction

Mr. Hiromasa Yamaguchi

He is currently an undergraduate student at Kyushu Institute of Technology, Japan. His research interests include reservoir computing and home service robots.

Mr. Akinobu Mizutani

He received the B.Eng. and the M.Eng. degree from the Kyushu Institute of Technology, Japan, in 2021 and 2023, respectively. He is currently a Ph.D. student at the Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology, Japan. His research interests include brain-inspired artificial intelligence and its robot application.

Mr. Arie Rachmad Syulistyo

He received the bachelor degree in computer science from the University of Brawijaya, Indonesia, in 2011 and the master degree in computer science from the University of Indonesia, Indonesia, in 2015. He is currently pursuing a Ph.D. degree in life science and systems engineering at Kyushu Institute of Technology, Wakamatsu, Japan. His research interests include computer vision and machine learning.

Dr. Yuichiro Tanaka

He received the B.Eng., M.Eng., and Ph.D. degrees from the Kyushu Institute of Technology in 2016, 2018, and 2021, respectively. He has also been a research fellow at the Japan Society for the Promotion of Science (JSPS) from 2019 to 2021. He has been an assistant professor at the Research Center for Neuromorphic AI Hardware of the Kyushu Institute of Technology since 2021. His research interests include neural networks, home service robots, etc. He is a member of IEEE, IEICE, and JNNS.

Prof. Hakaru Tamukoh

He received the B.Eng. degree from Miyazaki University, Japan, in 2001. He received the M.Eng and the Ph.D. degree from Kyushu Institute of Technology, Japan, in 2003 and 2006, respectively. He was a postdoctoral research fellow of 21st century center of excellent program at Kyushu Institute of Technology, from April 2006 to September 2007. He was an assistant professor of Tokyo University of Agriculture and Technology, from October 2007 to January 2013. He had been an associate professor from February 2013 to March 2021 and is currrectly a professor in the graduate school of Life Science and System Engineering, Kyushu Institute of Technology, Japan. His research interest includes hardware/software complex system, digital hardware etc. He is a member of IEICE, SOFT, JNNS, IEEE, JSAI and RSJ.

A Rapidly Adjustable Object Recognition System
through Language Based Prompt Engineering

Naoki Yamaguchi
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu, 808-0196, Japan

Tomoya Shiba
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu, 808-0196, Japan

Kosei Isomoto
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu, 808-0196, Japan

Hakaru Tamukoh
Research Center for Neuromorphic AI Hardware, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu, 808-0196, Japan
E-mail: yamaguchi.naoki892@mail.kyutech.jp
https://www.lsse.kyutech.ac.jp/

Abstract

We propose the use of language-based prompt engineering to achieve rapidly adjustable object recognition in RoboCup@Home. The proposed prompt engineering involves humans adding features, such as the color and material of an object, into the text prompts inputted into Language Segment Anything. In this research, we evaluated the effectiveness of our proposed method in three benchmark tests for object recognition at RoboCup@Home 2023. The results show that the highest scores were obtained in specific tasks, indicating that the proposed method applies to various recognition tasks.

Keywords: Home service robot, Object recognition, Prompt engineering, RoboCup@Home

1. Introduction

In recent years, service robots have attracted increasing attention against the aging population background [1]. This trend has spurred active research in this domain [2], [3]. These robots operate in dynamic environments, frequently encountering new and unfamiliar objects, making the learning process of recognition technologies and operational efficiency crucial. Ono et al.'s research [4] automates dataset generation and annotation, reducing the time and cost involved in the learning process while maintaining real-time performance using You Only Look Once v4 (YOLOv4) [5]. However, their experiments [6] showed that YOLO learning took two hours to prepare 500,000 training images and 18 hours. We aim to achieve a faster learning process and improve operational efficiency.

This study proposes Language Segment Anything (Lang SAM) [7] as a language-based object recognition system, offering quick adjustability. To validate its effectiveness, we conducted experiments at RoboCup@Home [8], an international competition focused on developing practical home service robots. Additionally, we evaluated operational efficiency by measuring inference time and power consumption, comparing it with the traditional method.

Naoki Yamaguchi, Tomoya Shiba, Kosei Isomoto, Hakaru Tamukoh
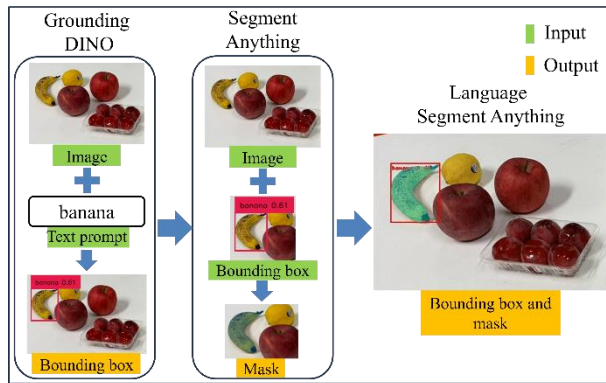
Fig. 1. Process flow of object recognition with language

2. Related Work

2.1. *Object recognition*

In object recognition, annotating datasets is an essential process [9]. Annotation involves labeling each object in an image with its name, location, shape, and other characteristics. Typically, manual processes accomplish this task. The performance of object recognition heavily depends on the quality and quantity of the dataset, as well as the precision of the annotations. Manually creating large volumes of high-quality annotations can incur substantial costs. Moreover, maintaining the accuracy of manual annotations is challenging.

2.2. *Dataset generation*

To address the issue of manual annotations in object recognition datasets, research on Sim2Real, which involves generating datasets using simulators, is progressing [10], [11]. Utilizing simulators eliminates human error and enables the rapid and high-quality generation of datasets. Additionally, this approach ensures consistency in annotations.

2.3. *Problems*

In addition to the challenges of dataset annotation, a significant issue in object recognition is the need to train recognition models with these datasets. Recognizing new objects requires generating datasets for these objects and conducting training, which can be very costly. This process involves substantial time and resources, mainly when introducing new objects to the system, thereby increasing the overall expense and complexity of developing effective object recognition models.

3. Proposed Method

3.1. *Object recognition with language*

For object recognition using language, our proposed method employs Lang SAM. Lang SAM combines Grounding DINO [12] and Segment Anything [13] into a single model, enabling language-based object recognition and segmentation. Fig. 1 shows a process flow of object recognition with language.

3.2. *Prompt engineering*

Adjust the text prompts that humans enter Lang SAM. Lang SAM's recognition accuracy depends on the text prompt. Below, we outline the procedure for adjusting text prompts:

(i)   Take pictures of multiple scenes containing the object of interest.
(ii) The initial text prompt shall be the name of the recognition target.
(iii) Input both the images and text prompts into the recognition system to check if the object is correctly recognized and assess the accuracy.
(iv) Improve the text prompts by human addition of details like color and material, aiming to recognize objects not initially identified or to increase accuracy.

4. Experimental Condition

First, we compared the preparation costs for recognizing new objects using the recognition models YOLOv8 [14] and Lang SAM. We generated the dataset for training YOLOv8 using conventional methods. Next, we implemented the proposed method on the Human Support Robot developed by TOYOTA MOTOR CORPORATION [15] to reveal its impact on enhancing object recognition capabilities. The effectiveness of this implementation was validated based on the scores obtained at RoboCup@Home. Finally, we compared the operational efficiency of YOLOv8 and Lang SAM, focusing on their inference time and power consumption, to determine which method was more efficient in real-world scenarios.

Table 1. Specifications of the
Computational System Used in Experiments

CPU	11th Gen Intel® Core™ i7-11700 @ 2.50GHz × 16
GPU	NVIDIA GeForce RTX 3080
Memory	32GB

Table 2. Time to prepare recognition model

Method	Dataset [images]	Prepare time [hours]
Dataset generator + YOLOv8	500,000	18
Proposed method	10	2

4.1. *Time to prepare recognition model*

We created a dataset to train YOLOv8 by generating 500,000 images using the method developed by the Ono et al. system. Following this, we conducted manual prompt engineering on these images. We then trained this dataset using YOLOv8. Next, images from 10 scenes were captured in real environments to serve as inputs for Lang SAM's prompt engineering. We then manually conducted prompt engineering on these images. The time taken to prepare these recognition models was measured.

4.2. *Evaluate the effectiveness of the proposed method*

The conditions of the second task performed at the RoboCup@Home 2023 are summarized:

(1) Storing Groceries (SG): The task involved organizing and shelving items scattered on a table, including categorizing objects and handling unknown items.
(2) Stickler for the Rules (SR): In this task, the robot acts as a party host, identifying guests not following house rules and enforcing compliance by explaining the rules to them. The rules are as follows:
    a. Remove shoes inside the house.
    b. Do not enter prohibited areas.
    c. Do not throw garbage on the floor.
    d. Always hold a drink in hand.
  This task primarily requires object recognition and human interaction technologies.

4.3. *Inference time and power consumption*

In section 4.1, we measured the inference time and power

Table 3. Results of RoboCup@Home2023
(proposed method: HMA)

		1st	2nd	3rd
SG	Team	**HMA**	Tech United Eindhoven	eR@sers +Pumas
	Score	**420**	410	220
SR	Team	**HMA**	TRAIL	Tech United Eindhoven
	Score	**1000**	1000	800

Table 4. Inference time and power consumption.

	Inference time [msec]	Power consumption [W]
YOLOv8	71.8	95.7
Lang SAM	577.7	233.0

consumption for each of the prepared recognition models. The inference time is the time recorded as the output of each recognition system, and this experiment uses a maximum wattage recorded by the NVIDIA System Management Interface [16] as a power consumption. Table 1 shows the specifications of the PC used in the experiments.

5. Experimental Result

5.1. *Time to prepare recognition model*

Table 2 shows the time taken for dataset generation, training with YOLOv8, and the preparation of the proposed method. YOLOv8 required 8 hours to create the training dataset and 10 hours for training, for a total of 18 hours. In contrast, the proposed method required approximately 2 hours for prompt tuning with 10 scene images and text prompts.

5.2. *The score of RoboCup@Home 2023*

Table 3 shows the top three teams and their scores in SG and SR categories at RoboCup@Home 2023. The results show that HMA achieved the highest SG and SR scores.

5.3. *Inference time and power consumption*

Table 4 shows the inference time and power consumption between the traditional YOLOv8 method and the proposed method. The results indicated that the proposed method had approximately 8 times later inference time and about 2.4 times higher power consumption than YOLOv8.

6. Conclusion and Discussion

In this study, we proposed using language-based prompt engineering to achieve rapidly adjustable object recognition in RoboCup@Home. We validated the effectiveness of this approach through participation in RoboCup@Home2023. We can adjust the proposed method in just 2 hours, which is only one-ninth of the time required by conventional methods. Furthermore, The results from RoboCup@Home 2023 support the feasibility and effectiveness of our approach in various tasks in terms of scoring. However, compared to traditional methods, it became clear that there are issues with operational efficiency, such as inference time and power consumption. We consider the need to select recognition models appropriately, considering the preparation cost and operational efficiency.

References

1. "Fuji Keizai Group," https://www.fuji-keizai.co.jp/report/detail.html?code=162208813, (Accessed 12/13/2023).
2. T. Shiba, T. Ono and H. Tamukoh, Object Search and Empty Space Detection System for Home Service Robot, Proceedings of the 2023 International Conference on Artificial Life and Robotics (ICAROB2023), 2023.
3. K. Isomoto, Y. Yano, Y. Tanaka, and H. Tamukoh, Robust trash can lid opening system, Proceedings of the 2023 International Workshop on Smart Info-Media Systems in Asia (SISA), 2023.
4. T. Ono, D. Kanaoka, T. Shiba, S. Tokuno, Y. Yano, A. Mizutani, I. Matsumoto, H. Amano and H. Tamukoh, Solution of World Robot Challenge 2020 Partner Robot Challenge (Real Space), Advanced Robotics, 2022, 36, pp.870-889
5. A. Bochkovskiy, C. Wang, H. M. Liao, YOLOv4: optimal speed and accuracy of object detection. 2020, arXiv:2004.10934.
6. T. Shiba, A. Mizutani, Y. Yano, T. Ono, S. Tokuno, D. Kanaoka, Y. Fukuda, H. Amano, M. Koresawa, Y. Sakai, R. Takemoto, K. Tamai, K. Nakahara, H. Hayashi, S. Fujimatsu, Y. Mizoguchi, M. Anraku, M. Suzuka, L. Shen, K. Maeda, F. Matsuzaki, I. Matsumoto, K. Murai, K. Isomoto, K. Minje, Y. Tanaka, T. Morie, and H. Tamukoh, Hibikino-Musashi@Home 2023 Team Description Paper, 2023, arXiv:2310.12650.
7. L. Medeiros, Lang Segment Anything, https://lightning.ai/pages/community/lang-segmentanything-object-detection-and-segmentation-with-text-prompt (Accessed 12/13/2023).
8. RoboCup@Home, https://athome.robocup.org/, (Accessed 12/13/2023).
9. Shu Yang, JingWang, Sheeraz Arif, Minli Jia, Shunan Zhong, SAL-Net: Self-Supervised Attribute Learning for Object Recognition and Segmentation, Wireless Communications and Mobile Computing, 2021.
10. M. Denninger, M. Sundermeyer, D. Winkelbauer, D. Olefir, T. Hodan, Y. Zidan, M. Elbadrawy, M. Knauer, H. T. Katam and A. Lodhi, BlenderProc: Reducing the Reality Gap with Photorealistic Rendering, Proceedings of International Conference on Robotics: Science and Systems (RSS), 2020.
11. S. Max and B. Sven, Stillleben: Realistic Scene Synthesis for Deep Learning in Robotics, Proceedings of IEEE International Conference on Robotics and Automation (ICRA), 2020.
12. S. Liu, Z. Zeng, T. Ren, F. Li, H. Zhang, J. Yang, C. Li, J. Yang, H. Su, J. Zhu and L. Zhang, Grounding DINO: Marrying DINO with Grounded PreTraining for Open-Set Object Detection, 2023, arXiv:2303.05499
13. A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W. Lo, P, Dollár and R. Girshick, Segment Anything, 2023, arXiv:2304.02643
14. Ultralytics, Ultralytics YOLOv8 Docs, https://docs.ultralytics.com/ (Accessed on 12/13/2023).
15. T. Yamamoto, K. Terada, A. Ochiai, F. Saito, Y. Asahara and K. Murase. Development of human support robot as the research platform of a domestic mobile manipulator, ROBOMECH journal, Vol. 6(1), 2019, pp. 1-15.
16. NVIDIA.DEVELOPER, System Management Interface SMI, https://developer.nvidia.com/nvidia-system-management-interface (Accessed 12/13/2023).

Authors Introduction

Mr. Naoki Yamaguchi

He received the B.Eng. degree from the National Institute of Technology, Ube College, Japan, in 2023. He is a Master's degree student at the Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology. His research interests include dataset creation and Visualization.

Mr. Tomoya Shiba
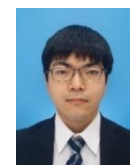
He received the B.Eng. degree from National Institute of Technology, Kagoshima College, Japan, in 2021. He received the M.Eng. from Kyushu Institute of Technology, Japan, in 2023. He is currently in a Ph.D. student in the graduate school of Life Science and Systems Engineering, Kyushu Institute of Technology. His research interest includes image processing, motion planning, and domestic service robots.

Mr. Kosei Isomoto

He received the B.Eng. degree from Kyushu Institute of Technology, Japan, in 2023, respectively. He is a Master's degree student at the Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology. His research interests include brain-inspired artificial intelligence models and home service robots.

Prof. Hakaru Tamukoh

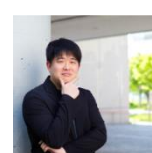He received the B.Eng. degree from Miyazaki University, Japan, in 2001. He received the M.Eng and the Ph.D. degree from Kyushu Institute of Technology, Japan, in 2003 and 2006, respectively. He was a postdoctoral research fellow of 21st century center of excellent program at Kyushu Institute of Technology, from April 2006 to September 2007. He was an assistant professor of Tokyo University of Agriculture and Technology, from October 2007 to January 2013. He had been an associate professor from February 2013 to March 2021 and is currrectly a professor in the graduate school of Life Science and System Engineering, Kyushu Institute of Technology, Japan. His research interest includes hardware/software complex system, digital hardware design, neural networks, soft-computing and home service robots. He is a member of IEICE, SOFT, JNNS, IEEE, JSAI and RSJ.

Development of A SayCan-based Task Planning System Capable of Handling Abstract Nouns

Kosei Yamao
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu, 808-0196, Japan

Daiju Kanaoka
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu, 808-0196, Japan

Kosei Isomoto
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu, 808-0196, Japan

Akinobu Mizutani
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu, 808-0196, Japan

Yuichiro Tanaka
Research Center for Neuromorphic AI Hardware, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu 808-0196, Japan

Hakaru Tamukoh
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu, 808-0196, Japan
Email: yamao.kosei665@mail.kyutech.jp

Abstract

The task planning system is required to accomplish various requests from a human in real-world environments. SayCan, one of the task planning systems, has high accuracy. However, its accuracy decreases for requests that include abstract nouns of the ambiguous word/phrase. We propose a novel task planning system based on SayCan that introduces a function for checking concrete names of abstract nouns and a rule-based skill extraction, enhancing accuracy. The proposed system facilitates the interpretation of requests and enables appropriate task planning. The effectiveness of the proposed system was demonstrated at RoboCup@Home, where it achieved high performance.

Keywords: Task planning, Home service robot, Large language model, RoboCup@Home

1. Introduction

Recently, the demand for home service robots has been increasing due to a low birth rate and an aging population, and research on such robots has been active [1], [2], [3], [4], [5], [6]. Highly accurate task planning is necessary to realize a general-purpose service robot that performs appropriate actions in response to human requests.

SayCan, one of the task planning systems [7]. consists of two modules: a Say module and a Can module. The Say module outputs the likelihood of each skill based on the language instruction, and the Can module outputs the likelihood of each skill based on the robot's current state. It decides which skill to execute next based on the likelihood output from two modules. Although SayCan is highly accurate, the accuracy decreases for commands containing abstract nouns such as "fruit" and "drink."

In this study, we propose a SayCan-based task planning system that introduces a function for checking concrete names of abstract nouns and a rule-based skill extraction. The contributions of this study are as follows.

- More straightforward interpretation of human requests and more accurate task planning
- Demonstrated high performance in a competition to evaluate the performance of home service robots

Kosei Yamao, Daiju Kanaoka, Kosei Isomoto, Akinobu Mizutani, Yuichiro Tanaka, Hakaru Tamukoh
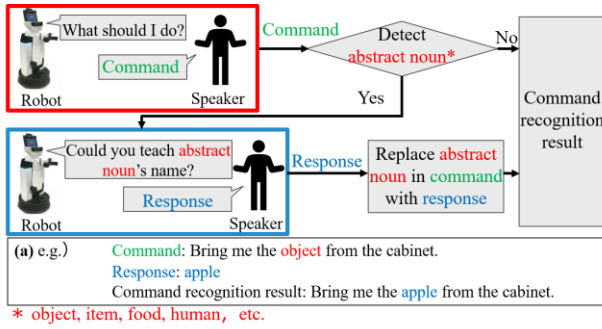
Fig. 1. Schematic of command recognition

2. Preliminaries and Related Works

2.1. *RoboCup@Home*

RoboCup@Home is an international competition for the technical development of home service robots [8]. RoboCup@Home includes two tests: General Purpose Service Robot (GPSR) and the more challenging Enhanced GPSR (EGPSR), in which robots listen to various commands and perform appropriate actions in the home environment. In competitions, commands generated from a command generator [9] are used. The command generator may output commands that include abstract nouns such as "fruit," as in "Please find the fruit in the dishwasher." The commands used in EGPSR are more complex than those used in GPSR and require more complex actions.

2.2. *SayCan*

SayCan has a predefined skill set that the robot can perform, and it determines which skill to execute from a given command. SayCan consists of the Say module and the Can module. The Say module uses a large language model (LLM) function to output the likelihood of each word in a sentence. The Say module inputs the given command and skill descriptions to the LLM and outputs the likelihood of each skill. The Can module acquires information about the external world from cameras and sensors and predicts which skills will most likely be executed. The results of these two modules are combined to determine the next skill to perform.

We think the Can module is difficult to implement on other robots because it is realized by reinforcement learning. SayCan is shown in the paper to be highly accurate task planning even with only a Say module without a Can module. The experimental results in the paper confirm that commands containing abstract nouns, such as "Bring me a fruit," reduce accuracy.
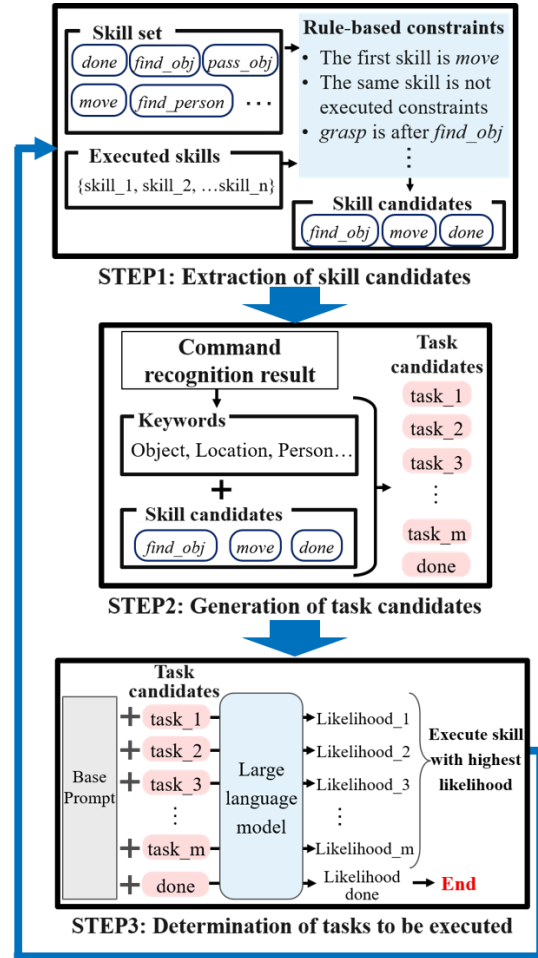
Fig. 2. Schematic of task planning

3. Proposed System

In this chapter, we describe the proposed system. The proposed system consists of command recognition and task planning modules.

3.1. *Command recognition*

Fig. 1 shows a schematic diagram of the function of command recognition in the proposed system. If an abstract noun such as an object, item, or food registered in the database is included in a given command, the proposed system asks for the concrete name of the abstract noun. This process aims to obtain a concrete name that the robot can recognize since it may not recognize the object of the abstract noun. As an example, as shown in Fig. 1(a), the robot asks for the concrete name of the abstract noun "object" and replaces "object" with "apple" depending on the result of the response.

Table 1. Skill set of the proposed system

Skill	Task
move	go to the {PLACE}
follow	follow the target
find_obj	find the {OBJECT} on the {PLACE}
find_person	find {PERSON}
observe_obj	look at the {PLACE} to check objects
observe_person	look at the {PLACE} to check people
grasp_obj	grasp the {OBJECT}
put	take the {OBJECT} to the {PLACE}
pass_obj	pass the {OBJECT}
answer_question	answer a question
say	say {LLM's output}
done	done

3.2. *Task Planning*

Fig. 2 shows a schematic diagram of the task planning function in the proposed system. The task planning consists of three steps.

STEP1: Extraction of skill candidates

Extract skill candidates from the skill set using rule-based constraints. The constraints are based on the order of skills, such as "The first skill is the move," "The same skill is not executed consecutively," and so on. This process reduces the number of LLM inputs. Eliminating skills that are very unlikely to be used is also expected to improve the accuracy of task planning. Table 1 shows the skill set of the proposed system. This skill set includes the basic skills the home service robot needs to perform various tasks. The say skill's speech content is complemented by LLM's output.

STEP2: Generation of task candidates

First, this process extracts object, place, and person names registered in the database in advance as keywords from the command recognition results. Then, generate the task candidates by combining the keywords and skill candidates. Table 1 shows the format of task candidates.

STEP3: Determination of tasks to be executed

A combination of base prompts and candidate tasks are input to the LLM, which outputs the likelihood of each candidate task. The robot then performs the task with the highest likelihood. Fig. 3 shows the base prompt. The base prompt gives the LLM information, such as the date, time, and role the command should follow.
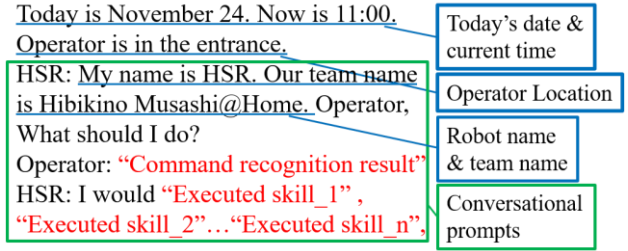
Fig. 3. Base prompt for the proposed system

4. Experiments

We conducted the following experiments to evaluate the proposed system. We use OpenAI's text-davinci-003 [10] as the LLM in the proposed system as a condition in experiments.

Experiment 1. Evaluation experiment with command output using the command generator [9]
Experiment 2. Evaluation experiment using actual competitions at RoboCup@Home 2023

In Experiment 1, we randomly generated 100 GPSR and 100 EGPSR commands, each using the command generator. In this experiment, the object and location names of the generated commands are those used in RoboCup@Home 2023. We input the generated commands into the proposed system and evaluate whether the task plan obtained as the output of the generated commands is feasible to realize the given commands. Specifically, we determined whether the sequence and targets of the skills included in the task plans can accomplish the corresponding commands.

In Experiment 2, we participated in RoboCup@Home 2023 held in Bordeaux, France, in July 2023 as Hibikino-Musashi@Home (HMA) team and evaluated the proposed system in the GPSR and EGPSR.

5. Experimental Results

5.1. *Experiment 1: Use the command generator*

As a result, we confirmed that task plans for 61 commands in GPSR and 27 commands in EGPSR were successful. In particular, we confirmed that the proposed system can successfully perform task planning by checking the concrete names of the abstract nouns in the commands to obtain the concrete names of the target objects. For example, in the command "Please bring me the fruit on the desk," the robot was checking the concrete name of the abstract noun "fruit," and the response "apple" was obtained. Table 2 shows examples of commands for which task planning was inaccurate and the possible causes. We found in many cases that task planning failures were due to incomplete skill sets and commands that did not include the required actions.

Kosei Yamao, Daiju Kanaoka, Kosei Isomoto, Akinobu Mizutani, Yuichiro Tanaka, Hakaru Tamukoh

Table 2. Examples of commands that failed in task planning and possible causes

Command	Causality
Go to the desk, look for the mug, and place it on the entrance door	The command does not specify *grasp*.
Could you please close the entrance door	No skill to open/close doors.
Hand me some coke in a mug	No skill to pour liquids.

Table 3. Results of RoboCup@Home 2023

	GPSR		EGPSR	
	Team	Score	Team	Score
1st	Tidyboy	400	**HMA**	700
2nd	TRAIL	300	TRAIL	400
3rd	**HMA**	200	Tidyboy	300

5.2. *Experiment 2: RoboCup@Home 2023*

In GPSR, the command given to the robot was, "Bring me the object behind the lemon from the cabinet." In response to this command, the robot asked for the concrete name of the abstract noun "object." The operator responded, "Tropical juice," however, the speech recognition failed, and the system did not replace the abstract noun. It scored partial points because it brought the object to the operator.

In EGPSR, in response to the command, "Get acquainted with Morgan at the exit, then find him in the living room, please," the robot found Morgan at the exit. After that, the robot said to Morgan, "Hello, Morgan. I'm HSR from Hibikino Musashi@Home". Finally, the robot moved to the living room and found Morgan. This command was judged to be successful and received a score.

Table 3 shows the top three teams and their scores in GPSR and EGPSR at RoboCup@Home 2023. As a result, HMA was ranked 3rd in GPSR and 1st in EGPSR, showing that the robot works well under real-world conditions.

6. Discussion and Conclusion

In this study, we proposed a SayCan-based task planning system, which introduces a function to check the concrete name of abstract nouns and a rule-based skill extraction. We demonstrated the system's effectiveness through experiments using the proposed system and evaluations in competitions. However, the proposed system has some issues, such as the incompleteness of the skill set and the fact that task planning sometimes fails for commands that do not contain the required actions. To realize a more accurate task planning system, we think it is necessary to develop skills such as those for opening/closing doors and to devise prompts to optimize the effectiveness of LLM [11].

In addition, for the system to operate in a real-world environment, it is necessary to acquire environment-specific knowledge, such as the preferences and habits of each family member. For environment-specific knowledge acquisition and utilization, studies of episodic memory acquisition models inspired by the hippocampus are undergoing [12], [13], [14]. In the future, we will integrate the episodic memory acquisition model into the robot to operate based on previously acquired environment-specific knowledge instead of checking concrete names of abstract nouns each time.

Acknowledgments

References

1. "Fuji Keizai Group," https://www.fuji-keizai.co.jp/report/detail.html?code=162208813, (Accessed 12/12/2023).
2. T. Ono, D. Kanaoka, T. Shiba, S. Tokuno, Y. Yano, A. Mizutani, I. Matsumoto, H. Amano and H. Tamukoh, Solution of World Robot Challenge 2020 Partner Robot Challenge (Real Space), Advanced Robotics, 2022, pp. 870-889.
3. Y. Yano, K. Isomoto, T. Ono, and H. Tamukoh, Autonomous waiter robot system for recognizing customers, taking orders, and serving food, Proceedings of the 26th RoboCup International Symposium, 2023.
4. K. Isomoto, Y. Yano, Y. Tanaka, and H. Tamukoh, Robust trash can lid opening system, Proceedings of the 2023 International Workshop on Smart Info-Media Systems in Asia (SISA), 2023.
5. T. Shiba, T. Ono and H. Tamukoh, Object Search and Empty Space Detection System for Home Service Robot, Proceedings of the 2023 International Conference on Artificial Life and Robotics (ICAROB2023), 2023.
6. D. Kanaoka, Y. Tanaka and H. Tamukoh, Applying Center Loss to Multidimensional Feature Space in Deep Neural Networks for Open-set Recognition, Proceedings of 17th International Conference on Computer Vision Theory and Applications (VISAPP2022), Vol. 5, 2022, pp. 359-365.
7. M. Ahn, A. Brohan, N. Brown, Y. Chebotar, O. Cortes, B. David, C. Finn, C. Fu, K. Gopalakrishnan, K. Hausman, A. Herzog, D. Ho, J. Hsu, J. Ibarz, B. Ichter, A. Irpan, E. Jang, R. J. Ruano, K. Jeffrey, S. Jesmonth, N. J. Joshi, R. Julian, D. Kalashnikov, Y. Kuang, K.H. Lee, S. Levine, Y. Lu, L. Luu, C. Parada, P. Pastor, J. Quiambao, K. Rao, J. Rettinghouse, D. Reyes, P. Sermanet, N. Sievers, C. Tan, A. Toshev, V. Vanhoucke, F. Xia, T" Xia', P. Xu, S. Xu,

M. Yan, and A. Zeng, Do As I Can, Not As I Say: Grounding Language in Robotic Affordances, arXiv preprint arXiv:2204.01691, 2022.

8. "RoboCup@Home." https://www.robocup.org/domains/3. (Accessed 12/12/2023).
9. "RoboCup@Home Command Generator." https://github.com/kyordhel/GPSRCmdGen. (Accessed 12/12/2023).
10. "OpenAI GPT-3.5 API [text-davinci-003]." https://platform.openai.com/docs/models/gpt-3-5. (Accessed 12/12/2023).
11. J. Wei, X. Wang, D. Schuurmans, M. Bosma, B. Ichter, F. Xia, E. Chi, Q. V. Le, and D. Zhou, Chain-of-thought prompting elicits reasoning in large language models, Proceedings of the Advances in Neural Information Processing Systems (NeurIPS), vol. 35, 2022, pp. 24824–24837.
12. Y. Tanaka, H. Tamukoh, K. Tateno, Y. Katori, and T. Morie, A brain-inspired artificial intelligence model of hippocampus, amygdala, and prefrontal cortex on home service robots, Proceedings of the 2020 International Symposium on Nonlinear Theory and Its Applications (NOLTA), 2020, pp. 138–141.
13. A. Mizutani, Y. Tanaka, H. Tamukoh, Y. Katori, K. Tateno and T. Morie, Brain-inspired neural network navigation system with hippocampus, prefrontal cortex, and amygdala functions, Proceedings of 2021 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS), 2021.
14. A. Mizutani, Y. Tanaka, H. Tamukoh, K. Tateno, O. Nomura and T. Morie, A knowledge acquisition system with a large language model and a hippocampus model for home service robots, Proceeding of the Institute of Electronics, Information and Communication Engineers (IEICE) Rep. vol. 123, no. 208, 2023, pp. 13–18.

Authors Introduction

Mr. Kosei Yamao

He received the B.Eng. degree from the National Institute of Technology, Ube College, Japan, in 2023. He is a Master's degree student at the Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology. His research interests include task planning and home service robots.

Mr. Daiju Kanaoka

He received the B.Eng. and the M.Eng. degree from Kyushu Institute of Technology, Japan, in 2020 and 2022, respectively. He is currently a Ph.D. student at the Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology. His research interests include object recognition, multimodal learning, and home service robots. He is a student member of IEEE.

Mr. Kosei Isomoto

He received the B.Eng. degree from Kyushu Institute of Technology, Japan, in 2023, respectively. He is currently a Master's degree student at the Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology, Japan. His research interests include brain-inspired artificial intelligence and home service robots.

Mr. Akinobu Mizutani

He received the B.Eng. and the M.Eng. degree from the Kyushu Institute of Technology, Japan, in 2021 and 2023, respectively. He is currently a Ph.D. student at the Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology, Japan. His research interests include brain-inspired artificial intelligence and its robot application.

Dr. Yuichiro Tanaka

He received the B.Eng. M.Eng., and Ph.D. degrees from the Kyushu Institute of Technology in 2016, 2018, and 2021, respectively. He has also been a research fellow at the Japan Society for the Promotion of Science (JSPS) from 2019 to 2021. He has been an assistant professor at the Research Center for Neuromorphic AI Hardware of the Kyushu Institute of Technology since 2021. His research interests include neural networks, digital hardware implementation, and home service robots. He is a member of IEEE, IEICE, and JNNS.

Prof. Hakaru Tamukoh

He received the B.Eng. degree from Miyazaki University, Japan, in 2001. He received the M.Eng and the Ph.D. degree from Kyushu Institute of Technology, Japan, in 2003 and 2006, respectively. He was a postdoctoral research fellow of 21st century center of excellent program at Kyushu Institute of Technology, from April 2006 to September 2007. He was an assistant professor of Tokyo University of Agriculture and Technology, from October 2007 to January 2013. He had been an associate professor from February 2013 to March 2021 and is currrectly a professor in the graduate school of Life Science and System Engineering, Kyushu Institute of Technology, Japan. His research interest includes hardware/software complex system, digital hardware design, neural networks, soft-computing and home service robots. He is a member of IEICE, SOFT, JNNS, IEEE, JSAI and RSJ.

RoboCup@Home 2023: Stickler for the Rules Task Solutions

Tomoya Shiba
Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku, Kitakyushu, 808-0196, Japan

Hakaru Tamukoh
Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku, Kitakyushu, 808-0196, Japan

Email: shiba.tomoya627@mail.kyutech.jp, tamukoh@brain.kyutech.ac.jp
https://www.brain.kyutech.ac.jp/~tamukoh/

Abstract

This paper proposes a robot vision system that detects illegal persons who violate house rules such as wearing shoes in a home environment. Such complex vision systems often require multiple AI systems including person detection, object detection, and more. Our approach simplifies this by leveraging the combined capabilities of Grounding DINO and SAM to detect rule violations effectively. The success of our method was proven at RoboCup@Home 2023, where it secured the highest score among all participating teams.

Keywords: Service robot, Object recognition, Mobile manipulator, RoboCup@Home,

1. Introduction

Home service robots are gaining increased attention due to the needs of aging societies with declining birth rates [1], [2], [3]. Their key functions include object recognition, picking and placing, identifying people and environments, and interacting with humans. Our research has involved using TOYOTA's home service robot HSR [4] and our Exi@ robot [5]. Various studies focusing on the social implementation of service robots for home use have been explored, often presented in competitions [6], [7], [8], [9]. These studies have mainly been proposed using dedicated recognition models. However, home environments vary greatly, and even within the same household, the objects to be recognized change over time and with the seasons. In RoboCup@Home competitions, the objects are usually announced in advance, leading most teams to develop recognition systems tailored to these specific objects [10]. However, some competitions require the detection of items not disclosed beforehand, such as people, their clothing, and common household objects. Therefore, it has been challenging to prepare recognition systems that can distinguish a wide variety of shoes to enforce a no-shoes rule or ensure adherence to a dress code at a party.

In this study, we propose a method for detecting rule violations in the competition task "Stickler for the Rules Task" using the detection models Grounding DINO [11] and Segment Anything [12], which employ prompt tuning for general object recognition. We implement these methods on TOYOTA's home service robot HSR. As a validation, we used them in RoboCup@Home [13], a competition for home service robots, to verify their effectiveness. (Fig. 1)

Fig. 1. Detection of rule violations

2. Related works
2.1. Object recognition

We have been researching methods for home service robots to detect the location of people and objects and then perform some action in response. Specifically, we have studied methods for tidying up and detecting people within a home. In our previous research, we developed a method for estimating the position of objects and available space, utilizing YolactEdge [14] and point cloud information [15].

2.2. *Problem*

The home environment contains many objects, including furniture such as desks and shelves. Additionally, the variety in human clothing is vast, requiring recognition of entirely different items depending on the season. Therefore, preparing a model that accounts for all these variations in a short period is challenging, and repurposing existing DNN models may lead to unstable recognition outcomes.

Tomoya Shiba, Hakaru Tamukoh

3. Proposed Method

3.1. *RoboCup@Home: Stickler for the Rules Task*

Here, I will explain the "Stickler for the Rules" task in RoboCup@Home. This task is designed for home service robots to identify party guests who have broken house rules and politely ask them to stop. The following four house rules are set:

1. All guests must take off their shoes at the entrance.
2. No guests are allowed in the Black Room.
3. Guests are not allowed to leave garbage on the floor.
4. All guests must always have a drink in hand.

3.2. *Proposed Method*

We propose a method for detecting violations using two Transformer-based models capable of prompt tuning. Grounding DINO is used for detecting bounding boxes, and Segment Anything is employed for segmentation tasks within the bounding box area. This allows for the identification of a wide variety of clothing and objects held in hands (Fig. 2). Additionally, to obtain the three-dimensional position, we use the pixel coordinates of the centroid of the bounding box and depth images.
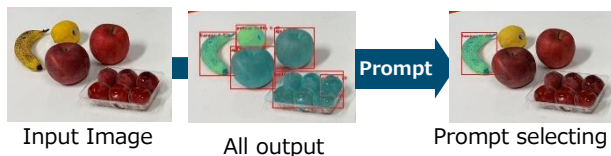
Fig. 2. SAM (Segmentation) & Grounding DINO (BBox)

4. Experiments and Results

4.1. *Partial Recognition*

To detect rule violations within a home, high recognition accuracy, as well as ease of surveillance, are required. To evaluate these performances, we tested the recognition accuracy of partially visible individuals. In this experiment, we prepared data with humans partially visible in images and conducted a comparison between YOLO v8 [16], MMDetection [17], and our proposed method. The images used were from the open-source human dataset DensePose-COCO [18], and we created images divided into quarters, reducing the image size to half, to test recognition accuracy. The results are summarized in (Table 1) As a result, our proposed method showed the highest recognition performance in the same task. However, it was also shown to be the most time-consuming method.

Table. 1. Results of Partial Recognition
(Stickler for the Rules)

Method	Detection rate [%]	Speed[s]
Proposed Method	**43.2**	1.122
Yolo v8 (Detect)	23.1	**0.103**
MMDetection	24.2	0.239

4.2. *RoboCup@Home task*

At RoboCup@Home 2023, we tested the effectiveness of this task by featuring more than 5 individuals committing rule violations, arranged in 3 different patterns of person placement. The results are summarized in (Table 2). As a result, we were able to perfectly address two types of rule violations and achieve the highest score.

Table. 2. Results of RoboCup@Home2023
(Stickler for the Rules)

Team	Score
HMA (Our team)	**1000**
Tidyboy	700
TRAIL	**1000**
Tech United	800
eR@sers+ Pumas	300
RoboCanes-VISAGE	200

5. Discussion

The proposed method has demonstrated higher recognition accuracy for partial visibility compared to existing methods. However, since it is based on a Transformer model, it tends to have longer processed times. When implementing this method as a function of home service robots, it may be effective to use the existing models for fast recognition in some scenarios, and our method for its strong recognition accuracy in cases of partial visibility. Additionally, the recognition method using prompt tuning is often influenced by the skill of the prompt setter. In the future, it will be necessary to establish objective criteria for prompt tuning.

6. Conclusion

In this study, we proposed a method for utilizing recognition models capable of prompt tuning to address the "Stickler for the Rules" task in RoboCup@Home, and conducted experiments in RoboCup tasks. The results showed that our method was effective, achieving the highest score in the same task. In the future, we aim to reclassify similar objects based on recognition scores and other information to reduce misrecognition.

References

1. T. Yamamoto, K. Terada, A. Ochiai, F. Saito, Y. Asahara, and K. Murase, ROBOMECH Journal, 2019.
2. L. Iocchi, D. Holz, J. Ruiz-del-Solar, K. Sugiura, T. van der Zant, Artificial Intelligence, pp. 258-281, 2015.
3. H. Okada, T. Inamura, and K.Wada, Advanced Robotics, 2019.
4. Yamamoto, T., Terada, K., Ochiai, A., Saito, F., Asahara, Y., & Murase, K. (2019). Development of human support robot as the research platform of a domestic mobile manipulator. ROBOMECH journal, 6(1), 1-15.
5. Hori S, Yutaro I, Kiyama Y, et al. Hibikino-Musashi@Home 2017 team description paper. Preprint. 2017. Avail-able from: arXiv:1711.05457
6. Savage J, Rosenblueth DA, Matamoros M, et al. Semantic reasoning in service robots using expert systems. Robot Auton Syst.2019;114:77–92..
7. Ishida Y, Morie T, Tamukoh H. A hardware intelligent processing accelerator for domestic service robots. Adv Robot. 2020 June;34(14):947–957.
8. Yoshimoto Y, Tamukoh H. FPGA implementation of a binarized dual stream convolutional neural network for service robots. J Robot Mechatron. 2021;33(2):386–399
9. Taniguchi, A., Isobe, S., El Hafi, L., Hagiwara, Y., & Taniguchi, T. (2021). Autonomous planning based on spatial concepts to tidy up home environments with service robots. Advanced Robotics, 35(8), 471-489.
10. T. Shiba, A. Mizutani, Y. Yano, et al,. "Hibikino-Musashi@Home 2023 Team Description Paper," arXiv:2310.12650, 2023.
11. Liu, S., Zeng, Z., Ren, T., et al.: 'Grounding DINO: Marrying DINO with Grounded PreTraining for Open-Set Object Detection', 2023, arXiv:2303.05499
12. Kirillov, A., Mintun, E., Ravi, N., et al.: 'Segment Anything', 2023, arXiv:2304.02643
13. "RoboCup Federation website" (2023 12/15 accessed)
14. Liu, H., Soto, R. A. R., Xiao, F., & Lee, Y. J. (2021, May). Yolactedge: Real-time instance segmentation on the edge. In 2021 IEEE International Conference on Robotics and Automation (ICRA) (pp. 9579-9585). IEEE.
15. T. Shiba, T. Ono, and H. Tamukoh, Object Search and Empty Space Detection System for Home Service Robot, The 2023 International Conference on Artificial Life and Robotics (ICAROB2023) OS17-2, Online, February 9-12(10), 2023.
16. Ultralytics.: 'Ultralytics YOLOv8 Docs', https://docs.ultralytics.com/, (Accessed on 12/14/2023).
17. Chen, K, et al. "MMDetection: Open mmlab detection toolbox and benchmark." arXiv preprint arXiv:1906.07155 (2019).
18. Güler, Rıza Alp, Natalia Neverova, and Iasonas Kokkinos. "Densepose: Dense human pose estimation in the wild." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

Authors Introduction

Mr. Tomoya Shiba

He received the B.Eng. degree from National Institute of Technology, Kagoshima College, Japan, in 2021. He received the M.Eng. from Kyushu Institute of Technology, Japan, in 2023. He is currently in a Ph.D. student in the graduate school of Life Science and Systems Engineering, Kyushu Institute of Technology. His research interest includes image processing, motion planning, and domestic service robots.

Dr. Hakaru Tamukoh

He received the B.Eng. degree from Miyazaki University, Japan, in 2001. He received the M.Eng and the Ph.D. degree from Kyushu Institute of Technology, Japan, in 2003 and 2006, respectively. He was a postdoctoral research fellow of 21st century center of excellent program at Kyushu Institute of Technology, from April 2006 to September 2007. He was an assistant professor of Tokyo University of Agriculture and Technology, from October 2007 to January 2013. He is currently an associate professor in the graduate school of Life Science and System Engineering, Kyushu Institute of Technology, Japan. His research interest includes hardware/software complex system, digital hardware design, neural networks, soft-computing and home service robots. He is a member of IEICE, SOFT, JNNS, IEEE, JSAI and RSJ.

Offloading Intellectual Processing from Home Service Robots to Edge Devices

Yuma Yoshimoto
NIT, Kitakyushu College, 5-20-1 Shii, Kokuraminamiku, Kitakyushu, Fukuoka, 802-0985 Japan

Mizuki Kawashima
NIT, Kitakyushu College, 5-20-1 Shii, Kokuraminamiku, Kitakyushu, Fukuoka, 802-0985 Japan

Shun Yonehara
NIT, Kitakyushu College, 5-20-1 Shii, Kokuraminamiku, Kitakyushu, Fukuoka, 802-0985 Japan
Email: yoshimoto@kct.ac.jp, k19053mk@apps.kct.ac.jp, k19206sy@apps.kct.ac.jp

Abstract

In this study, we focus on extending the operational time of home service robots by offloading intellectual processing to circuit devices such as Field Programmable Gate Arrays (FPGAs), which in turn reduces power consumption. The core of our approach involves developing a method for implementing intellectual processing on FPGAs, coupled with a dynamic circuit reconfiguration technique. This enables the FPGA to adaptively respond to frequent task changes. We present: (a) methods for transitioning circuits from a robot's control computer to FPGA in response to varying tasks, and (b) an evaluation of the effectiveness of using FPGA to extend operational time under these rapidly changing task conditions.

Keywords: Intelligent Processing, Edge Device, Service Robots, FPGA, Neural Networks

1. Introduction

In recent years, service robots have gained attention against the backdrop of a society experiencing declining birthrates and an aging population. Service robots are designed to perform a variety of tasks in homes and stores, such as tidying up and waitering, thereby reducing the workload on humans. These robots are equipped with various intelligent processing capabilities such as object recognition and voice recognition, allowing them to understand their surroundings and operate flexibly. Recently, Neural Networks (NNs) have been increasingly used for these intelligent processing tasks, and it is known that NNs require a substantial amount of computational power. The intelligent processing in service robots demands real-time operation, but achieving this through software alone is challenging. Therefore, hardware architecture acceleration is necessary.

Today, GPUs are primarily used as the hardware architecture for implementing NNs. However, running numerous NNs simultaneously leads to issues with insufficient hardware resources. Additionally, GPUs are generally known for their high-power consumption, which limits the operational time of robots.

This research aims to extend the operating time of robots by offloading the intelligent processing to circuit devices like Field Programmable Gate Arrays (FPGAs), thereby reducing power consumption. To achieve this, we propose the implementation of intelligent processing on FPGAs and a method to rewrite the circuits on the FPGA depending on the situation. This paper reports on the initial investigation, including (a) implementing state machines in robots and verifying the ability to construct appropriate state machines in response to user commands, (b) exploring methods to switch circuits from the robot control computer to the FPGA, and (c) examining the effectiveness of offloading the intelligent processing to the FPGA.

2. System Configuration of Service Robots

This section describes the configuration of the service robots targeted in this study.

2.1. *Behavior Planning with State Machines*

In service robots, processing is carried out through the construction and execution of state machines, as shown in Fig. 1. The process flow is as follows:

(1) Listen to the Commands

The robot listens to the user's voice and understands the commands as text.

(2) Construction of the State Machine

Yuma Yoshimoto, Mizuki Kawashima, Shun Yonehara

The robot interprets the heard commands and decides in what order and which intelligent processes to utilize.

(3) Execution of the State Machine

The robot executes the constructed state machine in sequence.

At this stage of executing the state machine (3), the state machine has already been constructed, and it is known in what order and which intelligent processes are needed.
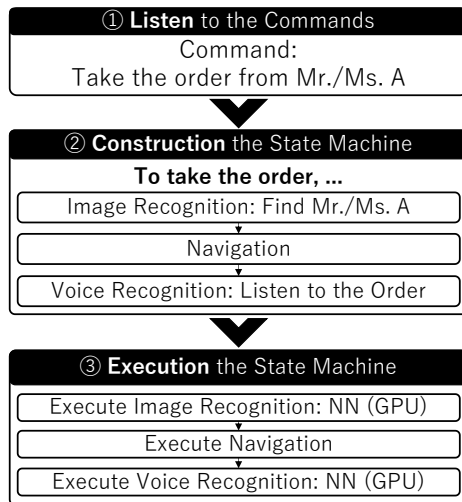
Fig.1 Constructing and Executing the State Machine

2.2. *Required Intelligent Processing*

Various intelligent processes are necessary for service robots. For example, Tsuji et al. have implemented the Neural Networks (NN) shown in Table 1 for the realization of versatile robots. In conventional methods, responding to human commands requires the simultaneous activation of numerous NNs.

Table 1. AI Implemented in the Robot [1]

AI	Model
Object Detection	Mask R-CNN, UOLS
Object Classification	CLIP (ViT-B/32)
Speech Recognition	Whisper
NLP and NLU	GPT-3

2.3. *Pathways Language Model (PaLM* [2]*), SayCan* [3]

There are developed by Google, this robot-specific AI allows robots to autonomously plan actions when given ambiguous verbal commands. For instance, if input like 'I spilled my drink. Help me,' is given, PaLM-SayCan can make decisions such as bringing a sponge or disposing of an empty can. The number of parameters required for PaLM is said to be 540B [2]. If these parameters are managed as float64, the memory size required for PaLM amounts to 216 GB. However, the memory in GPUs used

for AI processing in service robots ranges from 12 GB to 24 GB. Therefore, it is not feasible to implement an AI of size 216 GB in the GPU of a robot control computer. Thus, integrating this method into the edge side is challenging. Additionally, incorporating this method into robots would rely on external computing devices like cloud services, which introduces instability due to network connections, making it difficult to achieve a stable system.

3. Field Programmable Gate Array

An FPGA is a semi-custom LSI composed of a grid of reconfigurable logic gates. FPGAs can achieve high computational performance by reconfiguring their internal logic circuits. Unlike GPUs, which are used for AI processing and consist of block-based arithmetic architectures like multiplication and division, FPGAs construct their architecture with logic operation units, enabling lower power consumption and faster processing. Therefore, it is conceivable that using FPGAs allows for real-time processing.

3.1. *Comparison of FPGA with GPU and CPU*

Nakahara et al. implemented CNNs on FPGA and compared it with CPU and GPU [4]. The comparison table is shown in Table 2. They used the NVIDIA Jetson TX1 board for comparison, which includes an embedded CPU (ARM Cortex-A57) and an embedded GPU (Maxwell GPU). The CPU and GPU used Caffe (version 0.14) and were compared when running VGG-11 with a batch size of 1 for latency measurement. As shown in Table 2, FPGA is faster and consumes less power than the embedded CPU and GPU. Thus, a binary CNN on FPGA is more suitable for embedded systems than CPU and GPU.

Table 2. Comparison with Embedded Platforms [5]

Platform	Embedded CPU	Embedded GPU	FPGA
Device	ARM Cortex-A57	Maxwell GPU	Zynq 7020
FPS [s⁻¹]	0.23	36.7	421.9
Power Consumption [W]	7	17	2.3

3.2. *FPGA Hardware Resource Utilization*

Masatomo et al. implemented a model based on YOLOv3-tiny with Depth wise Separable Convolution on FPGA (Zynq-7020). They reported a BRAM utilization rate of 85.71%, indicating that implementing just one CNN on FPGA already exceeds 50% of the necessary resources. Therefore, it is difficult to implement other NNs simultaneously.

4. Proposal

We propose a method to control FPGA with a state machine. Fig. 2 shows the diagram of the proposed system. The operating procedure of the system is as follows:

① A control computer's state machine sends an instruction for the AI mode to be executed to the System on a Chip Manager (SoC Manager), which then instructs the FPGA Manager to rewrite.

② The FPGA Manager calls pre-defined circuit modules and rewrites the intelligent processing circuits on the FPGA.

③ The FPGA Manager sends input data for inference to the intelligent processing circuit on the FPGA.

④ The intelligent processing circuit on the FPGA completes the inference and sends the results back to the FPGA Manager.

⑤ The FPGA Manager sends the inference results to the SoC Manager, which then sends them to the state machine.

By repeating these steps, intelligent processing is executed on the FPGA while time-sharing its resources. Executing intelligent processing on the FPGA resolves the issue of power consumption, and time-sharing the FPGA resolves the resource issue.
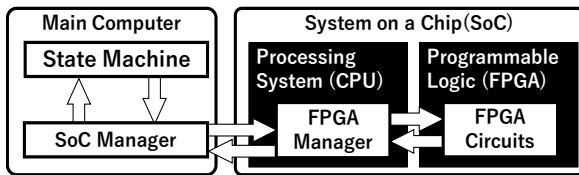
Fig.2 Overview of the Proposed System

5. Experimental

5.1. *Construction of the State Machine*

A system as shown in Fig. 1 was constructed, and it was verified that it could (1) listen to commands, (2) construct a state machine, and (3) execute the state machine. In this verification, the following technologies were implemented as elemental technologies:

- Recognition of opening and closing doors through Depth image processing
- SLAM for creating environmental maps and estimating self-position
- Navigation allowing the robot to autonomously move to a target location
- Person and posture detection using skeleton estimation AI
- Object detection with Faster-RCNN

- Depth image processing to estimate the three-dimensional position of objects
- Speaking function using gTTS and HSR-API
- Voice recognition to understand human speech using Whisper
- Context understanding that interprets the meaning from recognized text strings

The system's operation was also confirmed. The operational verification was conducted at GPSR task of RoboCup JapanOpen 2022 @Home League, a benchmark competition for service robots. This task competes on how well robots can respond to input from people, evaluating the constructed state machine. The robot was given instructions in English, and its ability to go to a designated location and perform specific actions was tested.

As a result of the experiment, the constructed state machine was given the input 'Go to the Living room, find Jennifer and answer her question.' First, the command was correctly processed using the voice recognition function, and then the robot moved to the living room using the autonomous movement algorithm. Next, it used person recognition to find Jennifer and moved to her location. After moving, it answered questions using the voice recognition AI. It was confirmed that the state machine was appropriately constructed for these states up to the movement, and the states changed according to the situational changes.

5.2. *Consideration of FPGA Circuit Switching Method*

As a preliminary step to constructing the system shown in Fig. 3, a method to switch the FPGA circuit from the control computer was realized. The specific configuration is shown in Fig. 3. This was implemented on the KV260 evaluation board equipped with FPGA-SoC, and its operation was verified. The operation flow is the same as described in Chapter 4. In the experiment, the inference process was confirmed to be possible by switching between two models of YOLO v3-Tiny [6], each trained on different datasets.

Fig.3 System Configuration

Yuma Yoshimoto, Mizuki Kawashima, Shun Yonehara
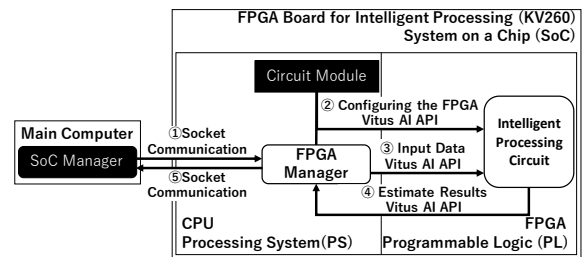
5.3. *Consideration of the Effectiveness When Implemented on FPGA Circuit*

The object recognition AI was implemented on AMD's KV260 FPGA evaluation board and an MSI-manufactured notebook PC equipped with RTX3060. The power consumption of each was measured both in a steady state and during the operation of the object recognition AI. The results are shown in Table 3. The difference between the power consumption during the operation of the object recognition AI and in the steady state was 2 W for the FPGA and 96 W for the GPU, which is a reduction to 1/48th. From this, it can be understood that the implementation of intelligent processing on FPGA is advantageous in terms of power consumption compared to GPU implementation.

Table 3. Difference in Power Consumption between FPGA and GPU

	Power Consumption [W]		
	Steady State	During Object Recognition AI Operation	Difference
KV260	10	12	2
MSI Stealth-15M-B12UE-012JP	44	140	96

5.4. *Verification Using Images Captured by the Robot*

Images were captured using an RGB-D camera (Xtion Pro) connected to the robot, and these images were input into the system. As a result, an output was obtained that shows what appears where in the images.

6. Conclusion

In this paper, we proposed a method for offloading power-intensive intelligent processing to FPGA circuit devices, aiming for extended operational times of service robots. In our verification, as a first step towards realizing the system, we (a) implemented a state machine in the robot and confirmed the ability to construct an appropriate state machine in response to user commands, (b) explored a method for switching circuits implemented on FPGA from the robot control computer, and (c) investigated the effectiveness of offloading intelligent processing to FPGA. The results confirmed that (a) the construction and execution of a state machine are possible, as demonstrated in the RoboCup@Home DSPL's GPSR, (b) it is feasible to switch actual circuits from the robot control computer to FPGA, (c) significant power reduction can be achieved when running the object recognition AI on FPGA compared to GPU. In addition (d) the edge device (FPGA) estimated the images from robot camera (Xtion Pro). Future work will focus on integrating these developments with the state machine and extending support to other intelligent processes.

References

1. C. Tsuji, D. Komukai, et al, "TRAIL 2023 Team Description Paper," https://trail.t.u-tokyo.ac.jp/project/robocup2023/.
2. A. Chowdhery, S. Narang, et al., "PaLM: Scaling Language Modeling with Pathways," arXiv:2204.02311.
3. M. Ahn, A. Brohan, et al., "Do As I Can, Not As I Say: Grounding Language in Robotic Affordances," arXiv:2204.01691.
4. H. Nakahara, H. Yonekawa, T. Fujii, M. Shimoda, S. Sato, "GUINNESS: A GUI Based Binarized Deep Neural Network Framework for Software Programmers," IEICE Transactions on Information and Systems," E102.D, 5, pp.1003-1011, 2019.
5. M. Matsuda, Y. Araki, et al., "An FPGA Implementation of Object Recognition System with Low Power Consumption using YOLOv3-tiny-based CNN," 2022 "Hinokuni - Land of Fire" Information Processing Symposium (HINOKUNI2022), 2022.
6. J. Redmon, A. Farhadi, "YOLOv3: An Incremental Improvement," IEEE Computer Vision and Pattern Recognition, pp. 1-8, 2018.

Authors Introduction

Dr. Yuma Yoshimoto

He received his B.Eng. degree from National Institute of Technology (KOSEN), Maizuru College, Japan, in 2016. He received his M.Eng. and D. Eng. degrees from Kyushu Institute of Technology, Japan, in 2018 and 2021, respectively. And he was JSPS researcher, in 2019 - 2021. He was a post-doctoral researcher at the Kyushu Institute of Technology, Japan in 2021-2022. Currently, he is assistant professor at the National Institute of Technology (KOSEN), Kitakyushu College, Japan. His research interests include deep learning, robot vision and digital hardware design. He is a member of IEICE, IEEE.

Mr. Mizuki Kawashima

He enrolled at National Institute of Technology (KOSEN), Kitakyushu College, Japan, in 2019. In 2021, he pursued the Information Systems course, focusing on algorithms and control. He commenced his research in robotics in 2022.

Mr. Shun Yonehara

He enrolled at National Institute of Technology (KOSEN), Kitakyushu College, Japan, in 2019. In 2021, he pursued the Information Systems course, focusing on algorithms and control. He commenced his research in robotics in 2022.

A Research on Performance Information Editing Support System for Automatic Piano

Yoshiki Horii

Kyushu Institute of Technology, 680-4 Kawazu, Iizuka-shi, Fukuoka, 820-8502, Japan

Eiji Hayashi

Kyushu Institute of Technology, 680-4 Kawazu, Iizuka-shi, Fukuoka, 820-8502, Japan
Email: hori.yoshiki311@mail.kyutech.jp, haya@mse.kyutech.ac.jp

Abstract

In 1996, Hayashi et al. from our laboratory developed an automatic piano playing device that requires data with nuances for each note to perform in a human-like piano playing. However, this device lacks the function to infer such nuanced data. Therefore, prior research focused on developing a system to infer data with nuances for each note. Yet, this system required manual inference, consuming a significant amount of time. In this study, we have constructed a system capable of automatically performing inference using deep learning. This system not only improves the accuracy of piano playing inference but also contributes to the efficiency of the inference process.

Keywords: Automatic Piano, Computer Music, Deep Learning

1. Introduction

The automatic piano playing device (Fig. 1) developed by Hayashi and others, equips a grand piano's keyboard with striking mechanisms and its pedals with driving mechanisms, allowing for precise keystrokes and pedal operations through computer control. During its development, the behavior of the piano action was analyzed using a unique mechanical model, examining its dynamic properties. Additionally, to create optimal drive waveforms according to each key, the results of the key-wise behavior analysis were stored in a database for stable sound reproduction. The waveforms are generated by referencing the database in accordance with the music data. This technology enables stable performance, even replicating delicate and rapid repetitive strikes that are challenging for professional pianists.

Fig. 1 Automatic piano playing device [1]

The device plays as instructed by inputting MIDI (Musical Instrument Digital Interface) standard data, a common standard for transferring and sharing performance data among electronic instruments, containing information like pitch and volume. By recording a professional pianist's performance as MIDI data and inputting it into the device, it's possible to replicate their performance. However, due to the system's design, inputting only the score data does not result in a human-like performance. To address this issue, efforts are being made to develop a system that generates human-like performances from score data. As a first step, efforts are being made to replicate the performances of world-renowned pianists.

Previous research developed a system that infers data with nuances for each note, but this system required manual inference. Since even short musical scores contain over a thousand notes, the inference process was time-consuming and labor-intensive. Therefore, in this study, we constructed a system that automatically performs inference using deep learning.

2. Inference system using Deep Learning
2.1. *Dataset*

In this research, we used performance data of the pianist Vladimir Davidovich Ashkenazy, which were recorded in accordance with the MIDI standard.

2.2. *Performance data*

In performance data, the raw state is often too complex for direct machine learning application. Therefore, this research involved separating the performance data into performance information and score information and setting parameters to make it manageable for machine learning. This process essentially simplifies and organizes the data, making it suitable for the algorithms to process and learn from, thereby enhancing the effectiveness of the machine learning application in the context of piano performance interpretation.

In Sections 2.3 and 2.4, the research details the parameters set for both performance information and score information.

2.3. *Performance information*

Performance information includes data that encapsulates the pianist's expression within the performance data. In this study, four parameters were set for the performance information. These parameters of the performance information are shown in Table 1.

Table 1 The format of the performance information

Parameter	Units	About
Velo	None	Sound intensity
Gate	ms	Length of note
Step	ms	Interval between the next note
Time	ms	Time of sound

In this research, Time is not a target of inference because it can be calculated from the Gate and Step values.

2.4. *Score information*

Score information is performance data that includes musical notation such as notes and musical symbols. In this research, five parameters were set from the score information. Table 2 shows the parameters of the score information.

Table 2 The format of the score information

Parameter	Units	About
Key	None	Sound height
Bar	None	Bar number
Dyn	None	Dynamics mark
Tgate	ms	Length of note on the score
Tstep	ms	Interval between the next note in the score

2.5. *Neural network configuration*

The inference systems developed to date have not been built using deep learning. However, in developing an inference system using deep learning, it is necessary to find the optimal network structure for inference of performance information. For this reason, we decided to construct and verify a system with a simple network structure. Fig. 2 shows a schematic of the constructed network.
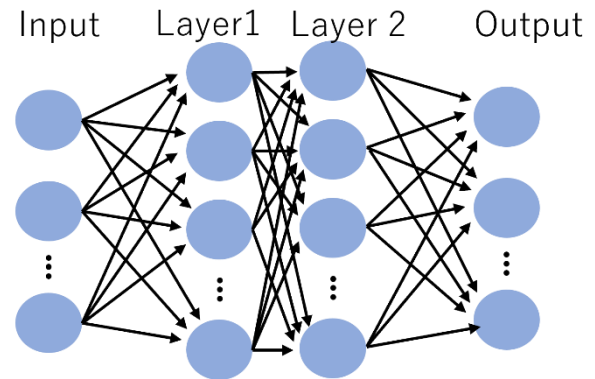
Fig. 2 Network overview

3. Inference Experiment and Results

3.1. *Data splitting*

In this experiment, the performance data was divided into training data, validation data, and test data. K-fold cross-validation is a statistical method for evaluating generalization performance in which the data is partitioned into K pieces, one of which is the validation data, and the remaining K-1 pieces are used as training data and evaluated with the test data. This is done by dividing the data into K pieces, one of which is the validation data. Fig. 3 shows a schematic of K-fold cross-validation. In this experiment, Prelude Op.28-7 by Fryderyk Francszek Chopin was used as the test data.
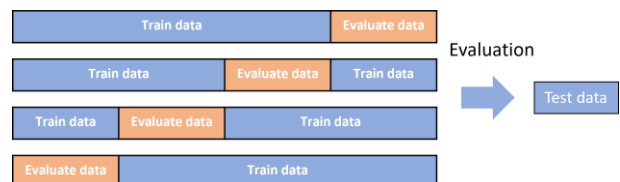
Fig. 3 K-fold Cross-Validation overview

3.2. *Validated network structure*

In this experiment, the layers of neurons in intermediate layer 1 and intermediate layer 2 were varied according to the patterns shown in Table 3.

Table 3 Combination of neurons in intermediate layers 1 and 2

Pattern number	Layer 1	Layer 2
1	50	50
2	50	100
3	100	100
4	100	50

3.3. *Results*

The performance information inferred by the inference system and the pianist's performance information were output as graphs for each Velo, Gate, and Step. The graphs of the patterns with the highest correlation coefficients are shown in Fig. 4, Fig. 5 and Fig. 6. The correlation coefficients between the inferred performance information and the pianist's performance information are shown in Table 4.
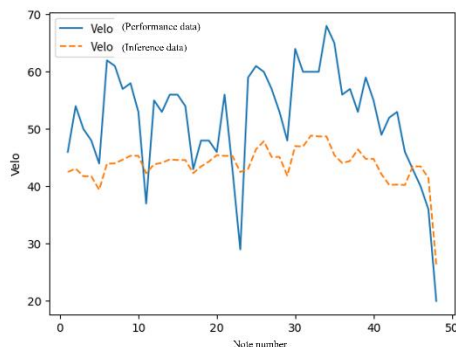
Fig. 4 Velo of pattern 2
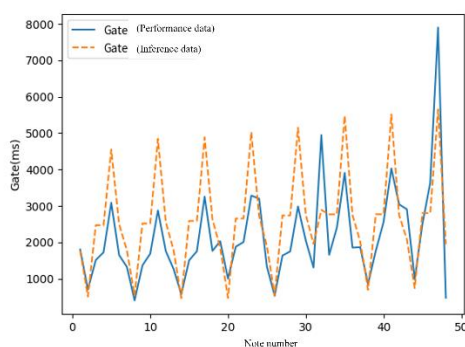
Fig. 5 Gate of pattern 3
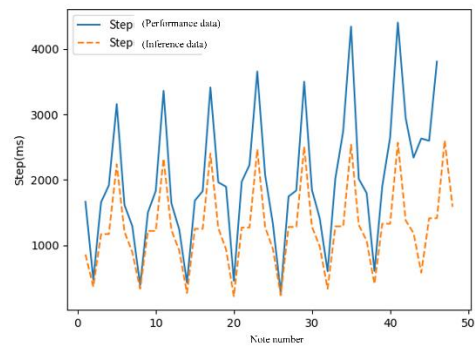
Fig. 6 Step of pattern 3

Table 4 Correlation coefficient of performance information

Pattern number	Layer 1	Layer 2	Velo	Gate	Step
1	50	50	0.763	0.720	0.876
2	50	100	0.776	0.741	0.880
3	100	100	0.770	0.743	0.885
4	100	50	0.763	0.691	0.832

4. Consideration

Table 4 shows that Velo's correlation coefficient does not change significantly for any of the patterns, and although Velo's correlation coefficient is high, the waveforms of the graphs are not similar, as can be seen in Fig. 4. This is because the points with large value fluctuations were learned as outliers, and the fluctuations in the inferred values became smaller.

The correlation coefficient of Step is larger than that of Velo and Gate. This is because the variation of the Step waveform is more regular than the other values, making it easier to learn, according to Fig. 6.

Furthermore, looking at the correlation coefficients for patterns 2 and 3, we see that the correlation coefficients for Velo, Step, and Gate are all large. This suggests that a network structure with a large number of neurons in the layers close to the output layer is suitable for inference.

5. Conclusion

In this study, we developed an inference system with a simple network structure using deep learning. Using the developed system, we verified the optimal network structure for inference of performance information. The results showed that a network structure with a large number of neurons in the layers close to the output layer is suitable for inference.

The future goal is to develop a more accurate inference system. To this end, we will search for the optimal

network structure by testing a system with an additional number of neurons, layers, and a gating mechanism that takes previous inputs into account.

References

1. E. Hayashi, M. Yamane, H. Mori, Development of a moving coil actuator for an automatic piano, Int. J. Japan Soc. Prec. Eng. 28 (1994), 164–169.

Authors Introduction

Mr. Yoshiki Hori

He received bachelor degree in Engineering in 2023 from mechanical system engineering, Kyushu Institute of Technology in Japan. He is currently a Master student at Kyushu Institute of Technology and conducts research at Hayashi Laboratory.

Prof. Eiji Hayashi

He is a professor in the Department of Intelligent and Control Systems at Kyushu Institute of Technology. He received the Ph.D. (Dr. Eng.) degree from Waseda University in 1996. His research interests include Intelligent mechanics, Mechanical systems and Perceptual information processing. He is a member of The Institute of Electrical and Electronics Engineers (IEEE) and The Japan Society of Mechanical Engineers (JSME).

Development of Autonomous Mobile Field Robot
– Accuracy Verification of Self-Localization through Simulation -

Takamasa Hayashi

Graduate School of Computer Science and Systems Engineering, Kyushu Institute of Technology,
680-4 Kawazu, Iizuka-city, Fukuoka, 820-8502, Japan

Shintaro Ogawa[1], Yuto Okawachi[1], Tan Chi Jie[1], Janthori Titan[1],
Ayumu Tominaga[2], Eiji Hayashi[1], Satoko Seino[3]
[1]Graduate School of Computer Science and Systems Engineering, Kyushu Institute of Technology,
680-4 Kawazu, Iizuka-city, Fukuoka, 820-8502, Japan
[2]Department of Creative Engineering Ribitucs and Mechatronics Course, National Institute of Technology,
Kitakyushu College, 5-20-1 Shii, Kokuraminamiku, Kitakyushu, Fukuoka, 802-0985, Japan
[3]Graduate School of Engineering, Kyushu University,
744 Motooka, Nishi-ku, Fukuoka, 819-0395, Japan

Email: hayashi.takamasa784@mail.kyutech.jp, tominaga@kct.ac.jp,
haya@ics.kyutech.ac.jp, seino@civil.kyushu-u.ac.jp,
http://www.kyutech.ac.jp/

Abstract

In recent years, the increase in marine debris has become a significant challenge in terms of its collection. Costal debris, a type of marine debris, can be collected by human hands, but the variety in shapes, and sizes presents limitations to human-only collection efforts. To address this, I focused on developing an autonomous mobile robot, establishing a simulation environment was considered crucial for facilitating smooth progress. This paper focuses on self-localization, an essential aspect for autonomous movement. We replicated an actual coastal cleaning site within the simulation environment and evaluated the accuracy of self-localization using an EKF (Extended Kalman Filter) with multiple sensors.

Keywords: Field Robot, Self-Localization, Gazebo, Simulation, Extended Kalman Filter (EKF)

1. Introduction

In recent years, the increase in marine debris has become a significant challenge in terms of its collection. Costal debris, a type of marine debris, can be collected by human hands, but the variety in shapes, and sizes presents limitations to human-only collection efforts. To address this, we focused on developing an autonomous mobile robot, establishing a simulation environment was considered crucial for facilitating smooth progress. This paper focuses on self-localization, an essential aspect for autonomous movement. We replicated an actual costal cleaning site within the simulation environment and evaluated the accuracy of self-localization.

2. Autonomous Mobile Field Robot "BUNKER"

In our previous study, we developed an autonomous mobile field robot [1] based on Kawasaki Heavy Industries' KFX®90, an all-terrain vehicle (ATV) powered by a gasoline engine. However, due to the vehicle's structure, the turning radius was large and stable traveling on steep slopes and rocky terrain was difficult, among other problems. To solve these problems, the platform was changed. The new platform uses Agilex's BUNKER [2] as shown in Figure 1. The main changes to the platform are that the traveling mechanism has two opposing wheels and the drive wheels are crawlers. This enables the platform to make super-clever turns, improving maneuverability composed to conventional platforms. In addition, the robot can climb hills with a slope angle of 36°, which is expected to enable it to run at higher speeds. The robot is equipped with an RGB-D sensor and 3D LiDAR as a visual system for autonomy. An encoder is mounted inside the vehicle body to measure the rotation speed of the wheels.

Takamasa Hayashi, Shintaro Ogawa, Yuto Okawachi, Tan Chi Jie, Janthori Titan, Ayumu Tominaga, Eiji Hayashi, Satoko Seino

a. KFX®90   b.BUNKER
Fig. 1. Platform Overview

2.1. *Simulation Construction*

In this study, we not only represented the Agilex BUNKER in the simulation environment but also recreated the coast of Hokuto Mizukumi Park in Munakata City, Fukuoka Prefecture, where actual beach cleaning is performed. We used 3D mapping technology developed by prior research using drones. For the simulation environment, we employed Gazebo, a 3D dynamic simulator capable of efficiently and accurately simulating groups of robots in complex indoor and outdoor environments. Gazebo can also be integrated with ROS (Robot Operating System) [3]. The projected BUNKER and the coast on Gazebo are shown in Figure 2.
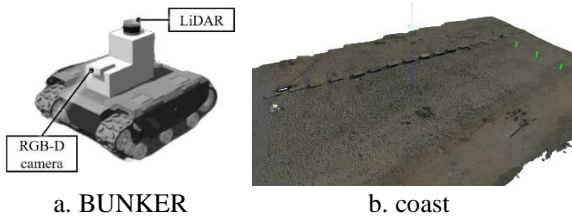
a. BUNKER   b. coast
Fig. 2. Built simulation environment

2.2. *EKF-Based Self-Localization System*

The posture measurement of mobile robots is achieved through various sensors and methods. However, the measurements obtained from these sensors are not true values but are considered to contain errors. In this study, we implemented the integration of odometry using the Extended Kalman Filter (EKF) shown in Figure 3, to achieve robust and low-error self-localization. We used wheel odometry, RGB-D odometry, and LiDAR odometry. The system was designed to accept any of these, individually or in combination, and produce an integrated odometry output.
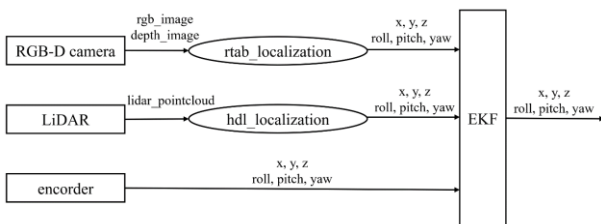
Fig. 3. Input/output data to/from EKF

3. Experiment

In the coastal environments shown in Figure 4 and Figure 5, the initial position of the robot was set as the origin of spatial coordinates, and from there, three destinations were set at (40, -5), (40, 0), (40, 5). The robot was remotely operated to each destination five times to extract the necessary odometry information. For self-localization, both individual odometries and multiple odometries integrated via EKF were implemented, resulting in seven estimation patterns. Mean Absolute Error (MAE) was used for error calculation.
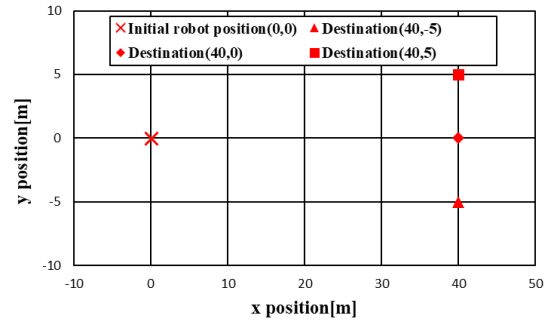
Fig. 4. Test environment

a. Coastal cross section
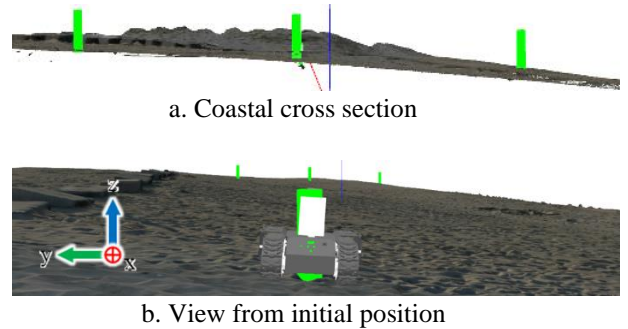
b. View from initial position
Fig. 5. Experiment in a simulation environment

3.1. *Result*

Table 1 summarizes the errors in self-localization results in the x-axis direction. In Pattern 4, the error was minimal, with the odometry integrating wheel odometry and RGB-D odometry via EKF achieving the highest accuracy in self-localization. Additionally, focusing on Pattern 2, the RGB-D odometry alone also showed generally similar values. From Pattern 3, it was observed that the LiDAR had an error of about 20 m, indicating it was not very useful in the simulated environment replicating an actual beach.

Table 1. Error in Self-Localization along the X-axis [m]

	Wheel	RGB-D	LiDAR	(40,5)	(40,0)	(40,-5)
1	○			11.06	1.22	0.35
2		○		0.08	1.06	1.44
3			○	21.66	25.25	19.50
4	○	○		0.85	0.47	0.86
5	○		○	13.73	5.42	4.47
6		○	○	10.36	10.07	8.27
7	○	○	○	3.20	4.33	5.54

Table 2 summarizes the errors in self-localization results in the y-axis direction. Compared to the results in the x-axis direction, there were significantly larger discrepancies overall. In every pattern, high-accuracy self-localization was not achieved. While some input destinations yielded accurate results, a stable trend was not observed. Among them, the odometry that integrated wheel odometry and RGB-D odometry via EKF had the smallest error.

Table 2. Error in Self-Localization along the X-axis [m]

	Wheel	RGB-D	LiDAR	(40,5)	(40,0)	(40,-5)
1	○			17.90	3.71	1.17
2		○		2.35	12.93	6.09
3			○	15.11	19.68	13.11
4	○	○		2.49	12.68	5.44
5	○		○	14.60	3.14	4.13
6		○	○	7.30	13.64	3.94
7	○	○	○	2.96	12.50	5.28

3.2. *Consideration*

Throughout this experiment, RGB-D odometry was found to have an error and greatly contributed to the improvement of self-localization accuracy. However, the accuracy of LiDAR odometry was considerably low in coastal environments like those in this study, lacking distinct landmarks for reference. The experimental environment featured a slope descending to the right from the travel direction, as depicted in Figure 5, leading to vehicular lateral skidding, which mainly affected y-axis self-localization. The introduction of an Inertial Measurement Unit (IMU) is proposed as a solution for more accurate vehicle posture determination, with prior research suggesting its substantial utility in enhancing results.

4. Conclusion

In this study, we developed a simulation environment for an autonomous mobile field robot and conducted accuracy verification of self-localization. As a result, odometry that integrated wheel odometry and RGB-D odometry using RGB-D and EKF was found to achieve more accurate self-localization. In the future, we aim to improve the accuracy of self-localization by adding an

IMU. we will also conduct verification on sandy beaches different from those in this study.

References

1. Ayumu Tominaga, Akihiro Koubara, Ryusuke Fujisawa, Eiji Hayashi, Abbe Mowshowitz. Development of lidar based navigation system for automation of tree harvesting process, Proceedings of International Conference on Artificial Life and Robotics 2021, pp. 469-471
2. BUNKER-Agilex Robotics
3. ROS, Open Robotics

Authors Introduction

Mr. Takamasa Hayashi

He received his Bachelor's degree in Engineering in 2023 from intelligent and Control Systems, Kyushu Institute of technology in Japan. He is currently a Master student in Kyushu Institute of Technology and conducts research at Hayashi Laboratory.

Mr. Shintaro Ogawa
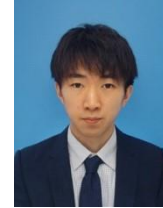
He received his Bachelor's degree in Engineering in 2022 from intelligent and Control Systems, Kyushu Institute of technology in Japan. He is currently a Master student in Kyushu Institute of Technology and conducts research at Hayashi Laboratory.

Mr. Yuto Okawachi

He received his Bachelor's degree in Engineering in 2023 from intelligent and Control Systems, Kyushu Institute of technology in Japan. He is currently a Master student in Kyushu Institute of Technology and conducts research at Hayashi Laboratory.

Mr. Tan Chi Jie

He received his Master of Creative Informatics from the Department of Computer Science and Systems Engineering, Kyushu Institute of Technology, Japan in 2023. He is currently a Doctoral student at Kyushu Institute of Technology and conducts research at Hayashi Laboratory.

Mr. Janthori Titan

He received his Bachelor of Engineering in Robotic Engineering and Automation System from the Department of Production Engineering, Faculty of Engineering, King Mongkut's University of Technology North Bangkok in 2022. He is currently a Master student at Kyushu Institute of Technology and conducts research at Hayashi Laboratory.

Assist. Prof. Ayumu Tominaga

He is a professor in Department of Creative Engineering Robotics and Mechatronics Course at National Institute of Technology Kitakyushu College. He received the Ph.D. (Dr. Eng.) degree from Kyushu Institute of Technology in 2021. His research interests include Intelligent mechanics, Mechanical systems and Perceptual information processing.

Prof. Eiji Hayashi

He is a professor in the Department of Intelligent and Control Systems at Kyushu Institute of Technology. He received the Ph.D. (Dr. Eng.) degree from Waseda University in 1996. His research interests include Intelligent mechanics, Mechanical systems and Percentual information processing. He is a member of The Institute of Electrical and Electronics Engineers (IEEE) and The Japan Society of Mechanical Engineers (JSME).

Associate. Prof. Satoko Seino

She is an associate professor in Graduate School of Engineering, Kyushu university. She received the Ph.D. (Eng.) from Kyushu University. Her research interests include coastal and riverine environmental conservation.

Development of AR System for Grasping String Foods on Introduction of Industrial Robot

Yoshihiro Koyama
*Graduate School of Computer Science and Systems Engineering, Kyushu Institute of Technology,
680-4 Kawazu, Iizuka-city, Fukuoka, 820-8502, Japan*

Eiji Hayashi[1], Akira Kawaguchi[2]
[1]Kyushu Institute of Technology, 680-4 Kawazu, Iizuka-city, Fukuoka, 820-8502, Japan
[2]Department of Computer Science, The City College of New York of The City University of New York
Email: koyama.yoshihiro272@mail.kyutech.jp, haya@ics.kyutech.ac.jp, akawaguchi@ccny.cuny.edu

Abstract

In recent years, the food service industry has been facing a labor shortage. However, the introduction of industrial robots is not easy due to the high cost of equipment and system integration. Therefore, we are developing an Augmented Reality (AR) application for the purpose of introducing robots to small and medium-sized companies. By using this application to perform tasks necessary for introducing robots, such as teaching, cost reductions can be expected when introducing industrial robots. In a previous study, an AR-based grasping and serving simulation system was developed for solidified foods such as fried chicken and rice balls. In this study, we focused on string-shaped food items such as spaghetti and attempted to develop an AR system for grasping and serving a string-shaped object by controlling a gripper.

Keywords: AR, Factory Automation Robots, Unity, ROS, String Foods

1. Introduction

In recent years, the food service industry has been facing a labor shortage. However, the introduction of industrial robots is not easy due to the high cost of equipment and system integration. Therefore, we are developing an Augmented Reality (AR) application for the purpose of introducing robots to small and medium-sized companies. By using this application to perform tasks necessary for introducing robots, such as teaching, cost reductions can be expected when introducing industrial robots.

In a previous study, an AR-based grasping and serving simulation system was developed for solidified foods such as fried chicken and rice balls. However, since there are a variety of foods that can be served in lunchboxes, it is necessary to develop a system that can simulate the grasping and serving of not only solidified foods but also other food items.

In this study, we focused on string-shaped food items such as spaghetti and attempted to develop an AR system for grasping and serving a string-shaped object by controlling a gripper.

2. Methodology
2.1. Robot Overview

Fig. 1 shows the appearance of the robot used in this study. The robot arm UTRA 6-550 and the gripper FLXI E-Type 2Finger manufactured by UmbraTek were used in this study. Since this robot is required to perform the same work as a human in a food factory, a 6-axis vertically articulated robot with a high degree of freedom of movement was used.

a.   UTRA 6-550   b.   FLXI E-Type 2 Finger
Fig. 1. Appearance of the robot

2.2. System Configurations

The system in this research is a simulation environment that reproduces the control of a gripper on a computer. The system consists of two main platforms, ROS (Robot Operating System) and Unity. ROS# [1], a Unity package, was used for communication between Unity and ROS.

Yoshihiro Koyama, Eiji Hayashi, Akira Kawaguchi

In addition, AR Foundation was used for AR application development. AR Foundation is a framework developed by Unity for AR application development, and by using this platform and software, it is possible to develop applications for different operating systems without worrying about program differences. By using the above platform and software, it is possible to develop applications for different operating systems without worrying about program differences. In this study, we used Google ARCore XR Plug-in, an asset for AR application development, assuming operation on Android devices.

Obi Rope is a particle-based physics engine based on the XPBD method that can simulate the physics of various objects.

The system configuration is shown below in Fig. 2. A comparison of software versions with previous studies is shown in Table. 1.
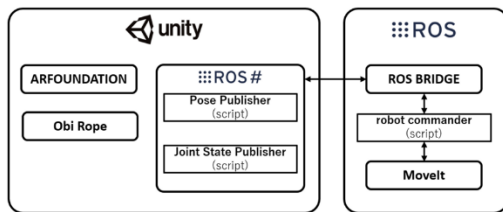
Fig. 2. System configuration

Table 1. Comparison of software versions

	Previous Research	This research
Unity	2019.4.28f	2021.3.7f
ROS	Melodic	Noetic
ROS#	1.7.0	1.7
AR Foundation		4.2.7
ARCore	ARCore SDK for Unity 1.24.0	ARCore XR Plugin 4.2.7

2.3. *Robot Accuracy of numerical integration*

This section describes the operation of the system in this study. Fig. 3 shows the motion planning of the robot. Each posture of the robot is described below. First, the coordinates of the grasping target and serving position are sent to the robot commander of ROS by Unity's Pose Publisher. Next, the trajectory of the robot arm is calculated using MoveIt based on the sent coordinates.

The calculation results are then sent to the Joint State Subscriber, which draws the robot arm and gripper on the Unity side. In the robot commander, the robot arm and gripper operate sequentially according to the motion planning shown in Fig. 3.
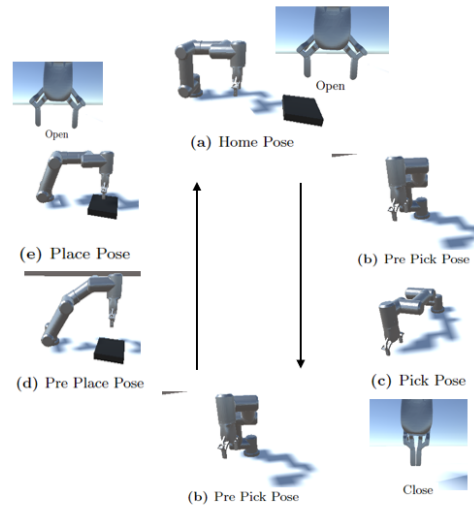
Fig. 3. Robot Planning
(a) Home Pose: Initial pose before motion is started
(b) Pre Pick Pose: Preliminary pose before grasping the object to be grasped
(c) Pick Pose; Pose when grasping the object to be grasped
(d) Pre Place Pose: Preliminary posture before placing the object to be grasped
(e) Place Pose: Pose when placing the grasped object

3. Results and Discussion

3.1. *Results of implementation*

The results of the implementation of this system are described below. First, we confirmed that the robot arm and gripper can be controlled by Unity, since the robot operates in the same way in Gazebo and Unity. Next, the grasping of the string object is shown in Fig. 4. Fig. 4 shows that the robot was able to lift the string object. However, it did not reach the point where it was heaped up, and the string object behaved in an unnatural manner. Next, communication between Unity and ROS was confirmed, as the robot control was confirmed, so communication was possible during Unity execution.
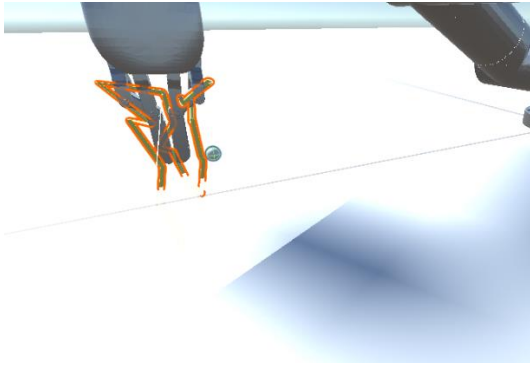
Fig. 4. Grasping

3.2. *Considerations*

First, we describe the grasping of string-like objects. The Obi Rope asset used to reproduce the string-like object is a chain of spheres, and the XPBD [2] method is used to reproduce the string. We believe that the gripper's pinching motion and the complicated structure of the gripper caused the string object to move in an unnatural manner. As a solution to this problem, we consider designing and introducing a gripper with a simple structure.

Next, we discuss communication between ROS and Unity in an AR environment. In this study, we use an updated version of the software used in the system in the previous study. We believe that this update has caused a problem in the conversion to JSON format for communication using ROS#. To solve this problem, it is necessary to check compatibility and find a new communication method.

4. Conclusion

In this study, we attempted to develop an AR system for grasping and serving string-shaped food items, with the main objective of extending the simulation environment for the application of AR applications for the introduction of industrial robots.

As a result of the implementation of the system in this study, it was confirmed that the gripper can be controlled and the Obi Rope can draw string-like objects in the AR environment, in addition to the system in the previous study. However, the communication between ROS and Unity in the AR environment and the grasping and placing of string-like objects could not be confirmed to work properly. In addition, we will continue to develop an application that combines the system with previous research.

References

1. ros-sharp, https://github.com/siemens/ros-sharp

2. Position Based Dynamics , Matthias Müller Bruno Heidelberger Marcus Hennix John Ratcliff, 2006.

Authors Introduction

Mr. Yoshihiro Koyama

He received his Bachelor's Degree in Engineering in 2023 from intelligent and Control Systems, Kyushu Institute of technology in Japan. He is currently a Master student in Kyushu Institute of Technology and conducts research at Hayashi Laboratory.

Prof. Eiji Hayashi

He is a professor in the Department of Intelligent and Control Systems at Kyushu Institute of Technology. He received the Ph.D. (Dr. Eng.) degree from Waseda University in 1996. His research interests include Intelligent mechanics, Mechanical systems and Percentual information processing. He is a member of The Institute of Electrical and Electronics Engineers (IEEE) and The Japan Society of Mechanical Engineers (JSME).

Prof. Akira Kawaguchi

He is a Professor and Head of the Computer Science Department of The City College of New York. He has been working together with Kyushu Institute of Technology on a unique project to make the future internet more secure, resilient, and efficient under the support of the U.S. National Science Foundation. He received B.S. and M.S. degrees in Administration Engineering from Keio University and M.S. and Ph.D. degrees in Computer Science from Columbia University.

An Image Analysis of Coastal Debris Detection
-Detection of microplastics using deep learning-

Yuto Okawachi

Department of Intelligent and Control Systems, Kyushu Institute of Technology, 680-4, Kawazu
Iizuka-city, Fukuoka, 820-8502, Japan

Shintaro Ogawa[1], Takamasa Hayashi[1], Tan Chi Jie[1], Janthori Titan[1], Eiji Hayashi[1], Ayumu Tominaga[2], Satoko Seino[3]

[1]Department of Mechanical Information Science and Technology, Kyushu Institute of Technology
680-4, Kawazu, Iizuka-City, Fukuoka, 820-8502, Japan
[2]Department of Creative Engineering Robotics and Mechatronics Course, National Institute of Technology Kitakyushu
Colllege, 5-20-1 Shii, Kokuraminamiku, Kitakyushu, Fukuoka, 802-0985, Japan
[3]Graduate School of Engineering, Kyushu University 744 Motooka, Nishiku, Fukuokacity, Fukuoka, 819-0395, Japan

E-mail: ookawachi.yuuto693@mail.kyutech.jp,
tominaga@kct.ac.jp, haya@mse.kyutech.ac.jp,
http://www.kyutech.ac.jp/

Abstract

To address the issue of litter drifting ashore, this study developed a deep learning-based microplastic detection system. The system employed yolov7 [1] as its deep learning network, complemented by SAHI (Slicing Aided Hyper Inference) [2] as an additional vision library. yolov7 is renowned for its efficacy in real-time object detection. Our experimental framework involved four tests, utilizing two variations of yolov7 - the standard model and yolov7-e6e - in conjunction with SAHI. The effectiveness of each test was quantified using metrics such as Intersection over Union (IoU), Precision, Recall, F-measure, and Detection Time in seconds. For our dataset, we gathered images from actual cleanup locations, such as Hokuto Mizukumi Park. The model's discriminator underwent 700 training iterations, with a learning rate set at 0.001. Experimental results showed that it detects fairly small microplastics.

Keywords: Deep learning, Object detection, YOLOv7

1. Introduction

Drifted litter has become widespread throughout Japan. In many regions, the composition of this litter, based on volume, weight, and count, is predominantly artificial rather than natural. Plastic litter does not decompose naturally and, due to the influence of ultraviolet rays and waves, breaks down into microplastics that drift in the ocean. They are then washed up on the sandy beaches. Microplastics, which may contain harmful substances from the time of product manufacturing or accumulate other toxicants, such as polychlorinated biphenyls, during their drift, can be ingested by marine life, potentially causing adverse effects. These microplastics, with a size of less than 5mm, make them difficult to visually detect on the sand.

A significant amount of waste washes up on the coasts of Japan. Among this, small items like microplastics, often buried in the sand, are difficult to detect by the current survey methods. Furthermore, some coastal areas, which are undeveloped or have wave-dissipating blocks, are difficult to access, making it impossible to undertake surveys there. To solve the problem of beached waste, a wide range of data is needed. However, the regular and systematic collection of data on the waste in various places is challenging with the present means of survey.

In this study, we aimed to realize labor-saving in beached waste surveys by using deep learning for image-based waste detection. Particularly, we focused on the challenge of detecting microplastics, extremely small items, previously undetected in other studies. To achieve this, we used deep learning to create a dataset with coastal images and utilized it for both training the detection model and validating its effectiveness. Fig.1 presents an image sample from the dataset. This research was conducted with a specific focus on a single-detection model, solely for microplastics.

Yuto Okawachi , Shintaro Ogawa , Takamasa Hayashi , Tan Chi Jie , Janthori Titan , Eiji Hayashi , Ayumu Tominaga , Satoko Seino

Fig.1 Examples Dataset

2. Methodology

2.1. *Composition of YOLOv7*

In this study, we employed the deep learning model known as YOLOv7. YOLOv7 is an object detection algorithm that operates at a higher speed than the existing YOLO series. As shown in Fig.2 on the MS COCO dataset, it achieves an AP value that significantly surpasses the existing YOLO series. There are several models within YOLOv7, and for our research, we used the base model, YOLOv7 (plain), and YOLOv7-E6E. Each model has a three-stage pyramid structure, generating multiple feature maps through repeated upsampling and downsampling. Ultimately, three of these feature maps are used for estimation. The base model, YOLOv7 (plain), uses a network composed of multiple convolutional layers known as ELAN. Additionally, YOLOv7-E6E employs an expanded version of the ELAN network, called E-ELAN. The structures of ELAN and E-ELAN are shown in Fig.3
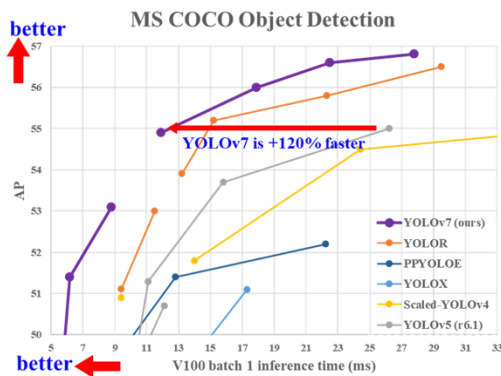
Fig.2 Comparison of YOLOv7 with other real-time object detectors
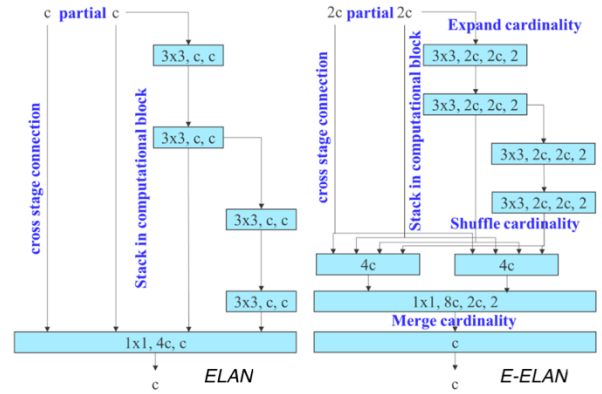
Fig.3 ELAN, E-ELAN Structure

2.2. *SAHI(Slicing Aided Hyper Inference)*

SAHI is a lightweight vision library for large-scale object detection and instance segmentation. As illustrated in Fig.4, it divides the input image into small rectangles. The model then performs an evaluation on each segmented image and finally merges and outputs the detection results. While SAHI can be applied to any deep learning model in principle, it has the limitation that the estimation time increases linearly, as it conducts the task on each of the segmented images.

Fig.4 Image segmentation image using SAHI

2.3. *Dataset creation*

The dataset used in this study was created using photographs taken at real-world beach cleanup locations, including Hokuto Miukumi Park. The main subject of the photographs was microplastics found on the sand. To enable the identification of these microplastics from other beach debris like shells and stones, we also included images of only shells and stones. The only class created was for microplastics. A total of 534 photographs were taken, and the number of images for training and final test datasets was tripled to 1,376 by using a brightness filter. These were then divided into 977 for training, 291 for final test dataset, and 108 for the general test dataset.

2.4. *Identifier creation and evaluation methods*

The identifier was trained 700 times using the created dataset. The study used models that were optimized to

achieve the highest weighted sum of Average Precision (AP) and mean Average Precision (mAP) at a 1:9 ratio during the training. Experiments were conducted using two different types of YOLOv7 (plain) and YOLOv7-E6E, as well as SAHI. The evaluation of the detector's effectiveness was based on the range of the IoU value being over 0.65, and the results were compared in terms of precision, recall, F1-score, and time of estimation. The F1-score, representing the harmonic mean of precision and recall, and the time of estimation, which is the time taken for the model to read the input image and produce an output, were the main points of performance analysis. The formulas for precision, recall, and F-score are shown in equations (1) to (4). Additionally, the confusion matrix is presented in Table1

$$Precision = \frac{TruePositive}{TruePositive + FalseNegative} \quad (1)$$

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative} \quad (2)$$

$$F - score = \frac{2 * Precision * Recall}{Precision + Recall} \quad (3)$$

$$IoU = \frac{Area\ of\ Intersection}{Area\ of\ Union} \quad (4)$$

Table1. Confusion Matrix

		Predictions	
		Positive	Negative
Actual Results	Positive	True Positive	False Negative
	Negative	False Positive	True Negative

3. Results and Discussion

Fig.5 Example of detection results

Table2. Score per model combination

Model	Precision	Recall	F-score	Time of estimation
Yolov7	0.771	0.679	0.722	0.157
Yolov7 + SAHI	0.801	0.671	0.730	1.63
YOLOv7-E6E	0.836	0.571	0.679	0.216
YOLOv7-E6E + SAHI	0.832	0.765	0.753	2.92

As observed in Fig.5, the detector can identify even very small microplastics. It also effectively prevents false detections of stones with similar shapes. Using SAHI with both models resulted in improved accuracy across all metrics except for estimation time. Specifically focusing on the YOLOv7-E6E model, the use of SAHI significantly increased the recall rate by nearly 0.2, demonstrating that the model becomes more robust in detecting microplastics with SAHI. However, the use of SAHI resulted in an estimation time that was approximately ten times longer.

The detection results demonstrate that the detector can accurately identify microplastics, indicating its practical viability. However, as indicated in Table2, high precision and recall rates were not achieved. A potential reason for this could be the dataset's quality. The annotations for the dataset were manually done, and objects not visually identifiable as trash were not labeled. The results show that the detector can identify microplastics that were too small to be labeled, leading to a lower precision rate. Regarding recall, the failure to detect microplastics resembling stones or shells in color resulted in lower accuracy. The dataset images were captured with an iPhone at resolutions like 3024×4032. To minimize the PC's load in the experiment, the input images for the model were resized to 1280×1280, making smaller objects even harder to recognize. Addressing these issues could involve labeling even smaller objects and increasing training data for items like stones and shells. Adjusting hyperparameters such as batch size might also enhance accuracy. The increased estimation time due to SAHI's use could be mitigated with software like TensorRT [3], which accelerates deep learning inference.

4. Conclusion

In a validation experiment, we developed a detector using YOLOv7 to identify microplastics and evaluated its accuracy. Accuracy was assessed using four metrics: precision, recall, F1-score, and estimation time. The results suggest that the detector achieves sufficiently accurate detection for practical use. It is suggested that modifying the annotation method and adjusting hyperparameters could lead to the development of a more accurate model.

Acknowledgements

Yuto Okawachi , Shintaro Ogawa , Takamasa Hayashi , Tan Chi Jie , Janthori Titan , Eiji Hayashi , Ayumu Tominaga , Satoko Seino

References

1. Chien-Yao Wang, Alexey Bochkovskiy, and Hong-Yuan Mark Liao, YOLOv7:Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors arXiv:2206.14651,2022.3,9,10
2. Akyon, Fatih Cagatay and Altinuc, Sinan Onur and Temizel, and Alptekin，Slicing Aided Hyper Inference and Fine-tuning for Small Object Detection 2022 IEEE International Conference on Image Processing (ICIP),2022.966- 970
3. NVIDIA Tensort.

Authors Introduction

Yuto Okawachi

He received his Bachelor's degree in Engineering in 2023 from intelligent and Control Systems, Kyushu Institute of technology in Japan. He is currently a Master student in Kyushu Institute of Technology and conducts research at Hayashi Laboratory.

Mr. Shintaro Ogawa

He received his Bachelor's degree in Engineering in 2022 from intelligent and Control Systems, Kyushu Institute of technology in Japan. He is currently a Master student in Kyushu Institute of Technology and conducts research at Hayashi Laboratory.

Mr. Takamasa Hayashi

He received his Bachelor's degree in Engineering in 2023 from intelligent and Control Systems, Kyushu Institute of technology in Japan. He is currently a Master student in Kyushu Institute of Technology and conducts research at Hayashi Laboratory.

Mr. Tan Chi Jie

He received his Master of Creative Informatics from the Department of Computer Science and Systems Engineering, Kyushu Institute of Technology, Japan in 2023. He is currently a Doctoral student at Kyushu Institute of Technology and conducts research at Hayashi Laboratory.

Mr. Janthori Titan

He received his Bachelor of Engineering in Robotic Engineering and Automation System from the Department of Production Engineering, Faculty of Engineering, King Mongkut's University of Technology North Bangkok in 2022. He is currently a Master student at Kyushu Institute of Technology and conducts research at Hayashi Laboratory.

Assist Prof. Ayumu Tominaga

He is a professor in Department of Creative Engineering Robotics and Mechatronics Course at National Institute of Technology Kitakyushu College. He received the Ph.D. (Dr. Eng.) degree from Kyushu Institute of Technology in 2021. His research interests include Intelligent mechanics, Mechanical systems and Perceptual information processing.

Prof. Eiji Hayashi

He is a professor in the Department of Intelligent and Control Systems at Kyushu Institute of Technology. He received the Ph.D. (Dr. Eng.) degree from Waseda University in 1996. His research interests include Intelligent mechanics, Mechanical systems and Percentual information processing. He is a member of The Institute of Electrical and Electronics Engineers (IEEE) and The Japan Society of Mechanical Engineers (JSME).

Associate. Prof. Satoko Seino

She is an associate professor in Graduate School of Engineering, Kyushu university. She received the Ph.D. (Eng.) from Kyushu University. Her research interests include coastal and riverine environmental conservation.

A High-Speed Estimation Method of Parameters in Impulse Response

Toshiki Tanaka, Ivan Tanev
Doshisha University, 1-3 Tatara Miyakodani, Kyotanabe, Kyoto 610-0394, Japan

Tetsuo Hattori,
Kagawa University, 1-1 Saiwai-cho, Takamatsu, Kagawa 760-0016, Japan
Email: hattori@pe.kagawa-u.ac.jp

Abstract

This paper proposes a high-speed parameters estimation method for compartment model where output function is described by the convolution between input function and impulse response, which is like a time-invariant linear system. The proposed method uses linear regression analysis based on the equivalently transformed equation that can be obtained using the processing of Differentiation of Convolution with Exponential function (DCE). In this paper, taking the parameters estimation problem of PET (Positron Emission Tomography) inspection system and RLC series electrical circuit for examples, we show that the method can estimate parameters of those impulse responses in high speed.

Keywords: Parameter estimation, Compartment model, Cumulative function, Linear regression analysis, PET Inspection

1. Introduction

Compartment model has been used in the various fields such as pharmacokinetics, chemical reaction system, environmental diffusion, and electrical circuit, etc. Especially, the estimation of parameters in PET (Positron Emission Tomography) inspection is very important in practice ([1], [2], [3], [4], [5], [6], [7], [8]).

In this paper, we propose a high-speed parameters estimation method for compartment model where output function is described by the convolution between input function and impulse response, which is like a time-invariant linear system ([9], [10], [11], [12]). The proposed method uses linear regression analysis based on an equivalent equation that can be obtained from DCE (Differentiation of Convolution with Exponential function).

Moreover, taking the parameters estimation problem of PET inspection system and RLC series electrical circuit for examples, we concretely describe that the method can estimate parameters of those impulse responses in high-speed.

2. Proposed Estimation Method

2.1 Impulse Response of PET Compartment Model

The compartment model and parameters used in the PET inspection is shown in Fig. 1. And the simultaneous differential equations of the 3-comparement model is shown in Eq. (1). The $C_p(t)$ means the tracer's radioactivity concentration in plasma that can be directly observed by arterial blood sampling from the patient after intravenous injection of FDG (^{18}F-fluorodeoxyglucose) tracer. The $C_e(t)$ means the tracer concentration before metabolism in tissue, and $C_m(t)$ is the tracer concentration after metabolism. $C_i(t)$ is measured by PET camera as the sum of $C_e(t)$ and $C_m(t)$. The flow parameters K_1, k_2, k_3, k_4 ($[ml \cdot g^{-1} \cdot min^{-1}]$ or $[min^{-1}]$) show the transport and binding rates of the tracer between two compartments.
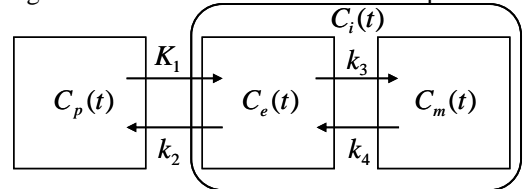
Fig.1. 3-Compartment model of tissue in FDG PET inspection [2], [4] with four tracer's flow parameters K_1, k_2, k_3, k_4.

$$
\begin{cases}
C_i(t) = C_e(t) + C_m(t) \\
\dfrac{dC_e(t)}{dt} = K_1 C_p(t) - (k_2 + k_3) C_e(t) + k_4 C_m(t) \\
\dfrac{dC_m(t)}{dt} = k_3 C_e(t) - k_4 C_m(t)
\end{cases}
\tag{1}
$$

In this model, the input function is $C_p(t)$, and the output is $C_i(t)$. Let $\otimes$ denote the convolution, and let g(t) be the impulse response of the compartment model described by Fig. 1 and Eq. (1). Then the output function $C_i(t)$

including the flow parameters is described by the following Eq. (2) and Eq. (3) ([2], [4]).

$$C_i(t) = g(t) \otimes C_p(t) = \int_0^t g(t-\tau) C_p(\tau) d\tau \qquad (2)$$

where

$$\begin{cases} g(t) = \dfrac{K_1}{\beta-\alpha}\left[(k_3+k_4-\alpha)e^{-\alpha t} + (\beta-k_3-k_4)e^{-\beta t}\right] \\ \alpha, \beta = \dfrac{1}{2}\left\{(k_2+k_3+k_4) \mp \sqrt{(k_2+k_3+k_4)^2 - 4k_2k_4}\right\} \end{cases} \qquad (3)$$

In this PET inspection, the problem is to estimate the values of flow parameters K_1, k_2, k_3, k_4 in the impulse response $g(t)$ from the two observed time series data, $C_p(t)$ and $C_i(t)$, where $C_i(t)$ is normally measured as a noisy time series data $r(t)$ $(=C_i(t)+noise)$ by PET Camera.

2.2 Estimation Using Linear Regression Analysis and DCE.

We equivalently transform the Eq. (1) to a linear regression model (LRM) equation using cumulative function, based on the following Theorem 1 ([9]) and Theorem 2 of DCE [12] that is derived from Theorem 1.

Theorem 1 [9]:

Let $f(x,t)$ and $\partial f(x,t)/\partial t$ be a continuous function and its continuous partial derivative function over the closed section $[a,b]$, respectively. And let the a and b are the functions $a(t)$ and $b(t)$ with respect to t, respectively. Then, Eq. (4) holds.

$$\frac{d}{dt}\int_{a(t)}^{b(t)} f(x,t)dx = \int_{a(t)}^{b(t)} \frac{\partial}{\partial t} f(x,t)dx$$
$$+ f(b(t),t)\frac{db(t)}{dt} - f(a(t),t)\frac{da(t)}{dt} \qquad (4)$$

Theorem 2 (DCE: Differentiation of Convolution with Exponential function) [12]:

Let $J(t)$ be the convolution between the functions $exp(-kt)$ and differentiable continuous function $C(t)$ as shown in Eq. (5), then from Theorem 1, Eq. (6) holds.

$$\frac{dJ(t)}{dt} \equiv J'(t) = (-k)J(t) + C(t) \qquad (5)$$
$$J(t) \Box e^{-kt} \otimes C(t) = \int_0^t e^{-k(t-\tau)} C(\tau)d\tau \qquad (6)$$

Let $r(t)$ be the measured value by PET camera at time t. The r(t) normally includes the additive noise $n(t)$, so $r(t)=C_i(t) + n(t)$. However, for the simplicity, we simply assume the case where the noise is not included in the measured value, then $r(t)=C_i(t)$.

Then, from Theorem 2, we obtain the Eq. (7).

$$r''(t) = (k_3+k_4)K_1 C_p(t) + K_1 C_p'(t)$$

$$r'(t) = dr(t) + \{(k_2+k_3+k_4) r'(t) C_p(\tau) \cdots \} / dt. \qquad (7)$$
where

Integrating both sides of Eq. (7) twice with respect to time t, we have the following Eq. (8) using cumulative functions such as

$$r(t) = A C_p^{II}(t) + B C_p^{I}(t) + C r^{I}(t) + D r^{II}(t) \qquad (8)$$
where

$$C_p^{I}(t) \Box \int_0^t C_p(\tau)d\tau, \quad C_p^{II}(t) \Box \int_0^t C_p^{I}(\tau)d\tau$$
$$r^{I}(t) \Box \int_0^t r(\tau)d\tau, \quad r^{II}(t) \Box \int_0^t r^{I}(\tau)d\tau \qquad (9)$$

$$\begin{cases} A = K_1(k_3+k_4), \quad B = K_1 \\ C = -(k_2+k_3+k_4), \quad D = -k_2 k_4 \end{cases} \qquad (10)$$

Eq. (8) can be regarded as a linear regression equation with objective variable $r(t)$ and four explanatory variables such as $C_p^{II}(t), C_p^{I}(t), r^{I}(t), r^{II}(t)$. Since the time series data $r(t)$ and $C_p(t)$ are given by PET Camera and blood sampling respectively, it is possible to obtain the values of parameters A, B, C, and D of Eq. (8), by linear regression analysis ([10], [11]). Then we can estimate the parameters K_1, k_2, k_3, and k_4 from Eq. (11).

This is the main idea of the proposed method, that is, we transform the simultaneous differential equations that describes the system operating characteristics into an equivalent equation that is convenient for linear regression analysis.

2.3 Weighted LRM-DCE Method.

Generally, in many cases, it is expected that noise will be mixed into the parameter estimation. For such occasion, we propose the Weighted LRM-DCE method that multiplies a known function $w(t)$ to the both sides of Eq. (8) as weight function, as shown in Eq. (11).

$$r(t)w(t) = A C_p^{II}(t)w(t) + B C_p^{I}(t)w(t) + C r^{I}(t)w(t) + D r^{II}(t)w(t) \qquad (11)$$

To determine the parameters A, B, C, and D in Eq. (9) is equivalent to obtain the same parameters in Eq. (8). This weighting function $w(t)$ has to be appropriately selected according to the noise characteristics.

For example, as the weighting function $w(t)$, we propose $r(t)^{-1}$ that is the inverse of the observed $r(t)$ in the region where $r(t)>0$. Then we have the following Eq. (12).

$$1 = A C_p^{II}(t)r(t)^{-1} + B C_p^{I}(t)r(t)^{-1} + C r^{I}(t)r(t)^{-1} + D r^{II}(t)r(t)^{-1} \qquad (12)$$

3. Applied Experimentation

3.1 Programming Environment for Experiments

The programming environment for numerical computation used in the experimentation are as follows.
CPU: Intel Core i7-3770 (3.40GHz), RAM: 16GB,
OS: Windows 10 Professional (64bit),
Programming Language: Borland C++ 5.51 (32bit).

3.2 Experiments for PET compartment Model

We have experimented for an input data $C_p(t)$ (Fig. 2 (a)) that is generated based on the literature [6] and noiseless output data $r(t)$ (Fig. 2 (b)) generated from the parameters $(K_1, k_2, k_3, k_4) = (0.200, 0.130, 0.060, 0.007)$ using Eq. (2) and Eq. (3). Moreover, we also experimented for $r(t)$ (Fig. 3 (a)) where the noise $n(t) \sim N(0,100^2)$ is added to the aforementioned noiseless $r(t)$. In this noisy case, as a matter of fact, we used the smoothed $r(t)$ (Fig.3 (b)) where moving average processing is done twice over the range of ±4 points before and after each time point of the noisy $r(t)$.

For the evaluation of estimation accuracy, we define the relative error as (|(Estimated Value)-(True Value)| / (True Value))× 100[%].

A. *Noiseless Case*
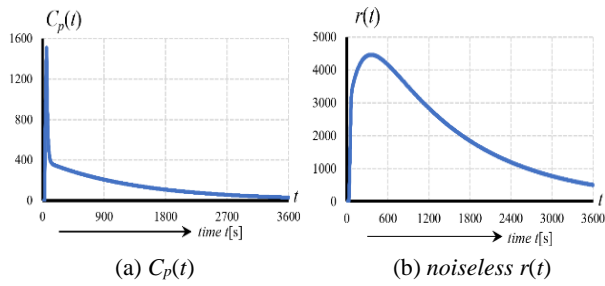
(a) $C_p(t)$ (b) *noiseless* $r(t)$

Fig.2. Generated input $C_p(t)$ and noiseless $r(t)$.

B. *Noisy Case*

(a) noisy $r(t)$ (b) smoothed $r(t)$
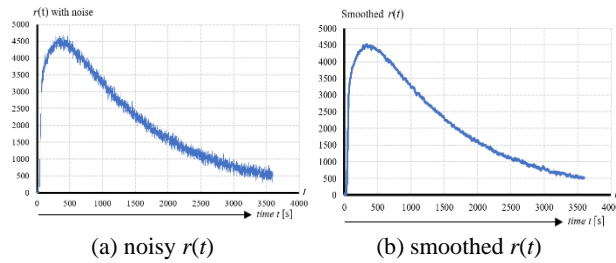
Fig.3. Generated noisy $r(t)$ and smoothed $r(t)$.

C. *Results*

We briefly call the LRM-DCE method and Weighted LRM-DCE method as LRM and WLRM, respectively.

The estimated results of (K_1, k_2, k_3, k_4) by both LRM and WLRM are shown in Fig. 4 for each parameter in each time interval.

Fig.4 (a) shows that the relative error rate (RER) of parameters estimation by LRM becomes less than 0.4 [%] for noiseless $r(t)$, if we take the time series data of 660 seconds (11 minutes) or longer for estimation computing. Fig. 4(b) shows that the resultant (RER) by WLRM is less than 0.46[%] until 1800[sec]. Even in the whole-time interval (=3600[sec]), the RER is less than 0.55[%]. Fig. 4(c) illustrates the resultant RER for noisy $r(t)$ by both WLRM and LRM, where WLRM shows that the rate is less than 3.3 [%] for all of four parameters in the whole-time interval, however the RER of LRM is worse than WLRM.

As for the processing time for both noisy and noiseless $r(t)$ by WLRM and LRM in the whole-time interval, the computing time of both method is less than 80 [ms].

(a) LRM for noiseless $r(t)$.

(b) WLRM for noiseless $r(t)$.

(c) WLRM and LRM for nosy $r(t)$.
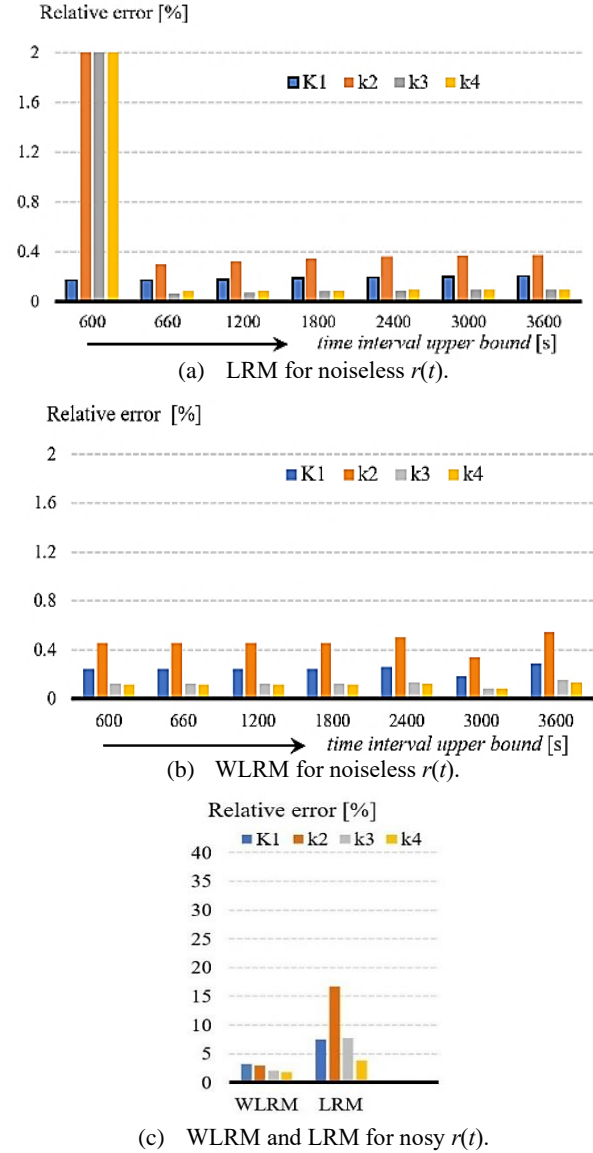
Fig. 4. Accuracy of the parameter estimation for K_1, k_2, k_3, and k_4 by the methods of LRM and WLRM ([12]).

3.3 Experiments for RLC Series Electrical Circuit
A. *Equation and the Impulse Response*

RLC series electrical circuit (RLC-EC) is one of most simple compartment model, that consists of three basic passive elements connected in series: a resistor R, an inductor L, and a capacitor C. Eq. (13) shows the system equation of RLC-EC with DC power supply (DC-PS) E (Fig. 5(a)), where we can regard it as the form of LRM itself without DCE processing.

$$L\frac{di(t)}{dt} + Ri(t) + \frac{1}{C}\int_0^t i(\tau)d\tau = E \qquad (13)$$

Then, we have the following Eq. (14) for parameters estimation, using cumulative functions.

Toshiki Tanaka, Ivan Tanev, Tetsuo Hattori

$$Li(t) + Ri^I(t) + \frac{1}{C}i^{II}(t) = \int_0^t E dt$$

where

$$\begin{cases} i^I(t) \square \int_0^t i(\tau)d\tau \\ i^{II}(t) \square \int_0^t i^I(\tau)d\tau = \int_0^t \int_0^\tau i(\xi)d\xi d\tau \end{cases} \quad (14)$$

Considering the RLC series circuit as a system of compartment model, the parameters of impulse response are those values of $R[\Omega]$, $L[H]$, $C[F]$, themselves. And, there are three types of impulse responses, depending on the magnitude relationship between $(R/2L)^2$ and $1/LC$.

Let those impulse responses be $g_1(t)$, $g_2(t)$, $g_3(t)$, depending on the cases, $(R/2L)^2 < 1/LC$, $(R/2L)^2 = 1/LC$, $(R/2L)^2 > 1/LC$, respectively.

(i) Oscillatory case: $(R/2L)^2 < 1/LC$

$$g_1(t) = \frac{1}{\sqrt{\left(\frac{L}{C}\right) - \left(\frac{R}{2}\right)^2}} \exp\left(-\frac{R}{2L}t\right) \cdot \sin\left(\sqrt{\left(\frac{1}{LC}\right) - \left(\frac{R}{2L}\right)^2}\, t\right)$$

(ii) Critical damping case: $(R/2L)^2 = 1/LC$

$$g_2(t) = \frac{1}{L}t \cdot \exp\left(-\frac{R}{2L}t\right)$$

(iii) Over-damping case: $(R/2L)^2 > 1/LC$

$$g_3(t) = \frac{1}{\sqrt{\left(\frac{R}{2}\right)^2 - \left(\frac{L}{C}\right)}} \exp\left(-\frac{R}{2L}t\right) \cdot \sinh\left(\sqrt{\left(\frac{R}{2L}\right)^2 - \left(\frac{1}{LC}\right)}\, t\right)$$

Fig. 5 shows the graphs of RLC circuit with DC-PS E and cumulative functions of i(t) in the case of oscillatory case.
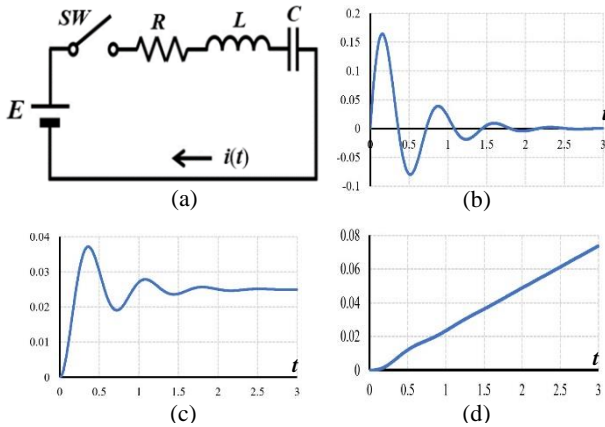
Fig.5. (a) RLC series electrical circuit with DC-PS E. (b) The $i(t)$ in the oscillatory case. (c) $i^I(t)$. (d) $i^{II}(t)$.

B. Results of RLC parameters estimation

We have experimented for the parameters estimation, from generated time series data of 3 seconds current $i(t)$ by taking 0.003[s] increments, that consists of 1001 pieces from time t=0 to t=3.
Estimated results (ER) and RER for each of above three cases with E=5[V] are as in the followings.
(i) Oscillatory case: $(R, L, C) = (10, 2.5, 0.005)$
  ER: (10.00120, 2.49988, 0.00499), RER < 0.2%.

(ii) Critical damping case: $(R, L, C) = (10, 2.5, 0.1)$
  ER: (10.00010, 2.50001, 0.10000), RER < 0.01%.
(iii) Over-damping case: $(R, L, C) = (10, 2.5, 0.2)$
  ER: (10.00000, 2.50003, 0.19999), RER <0.01%.

4. Conclusion

In this paper, we have proposed a method for parameters estimation of impulse response in compartment model systems. The proposed idea is based on the equivalent transformation from the simultaneous differential equations of the system behavior into an equation that is suitable for the estimation by linear regression analysis.

For further study, we are going to experiment for more complicated circuits and pursue the noise cancelation method.

References

1. Godfrey, K.: "Compartmental Models and Their Application", Academic Press (1983).
2. Hideo Onishi, Masanori Matsumoto, and Kazutaka Masuda ed.: "Nuclear Medicine Testing Technology", supervised by The Japanese Society of Radiological Technology, Ohmsha (2002). (in Japanese)
3. Kinichi Hisada supervised, Norihisa Tonami, Atsushi Kubo ed.: "Latest Clinical Nuclear Medicine", A New Textbook of Clinical Nuclear Medicine (3rd edition), Kanehara Shuppan Co., Ltd. (2005). (in Japanese)
4. R. P. Maguire and K. L. Leenders, ed.: "PET Pharmacokinetic Course Manual", International Society of Cerebral Blood Flow and Metabolism, Summer School 2007. (2007)
5. Hiroshi WATABE, Yoko IKOMA, Yuichi KIMURA, Mika NAGANAWA, and Miho SHIDAHARA: "PET kinetic analysis—compartmental model", Annals of Nuclear Medicine, Vol. 20, No. 9, 583–588, (2006).
6. Kudomi N, Choi E, Yamamoto S, Watabe H, Kim K, Shidahara M, et al.:" Development of a GSO detector assembly for a continuous blood sampling system", IEEE Trans. Nuclear Science, Vol. 50, No.1, pp.70–73 (2003).
7. Naganawa M, Kimura Y, Ishii K, Oda K, Ishiwata K, Matani A, "Extraction of a plasma time-activity curve from dynamic brain PET images based on independent component analysis", IEEE Trans. Biomedical Engineering, Vol. 52, No.2, pp.201–210. 2005.
8. Hiroshi Hirayama, Tatsuo Otsuki: "Electrical Circuit Theory (3rd Edition Revised)", Institute of Electrical Engineers of Japan (2008) (in Japanese)
9. Shigeichi Moriguchi, Kanehisa Udagawa, Shin Hitotsumatsu: "Iwanami Mathematics Formula (I)", Iwanami Shoten (1987). (in Japanese)
10. Hirokazu Anai: "Practical Guide of Mathematical Optimization", Kodansha, ISBN978-4-06-156510-4 (2016). (in Japanese)
11. William H. Press, Brian P. Flannery, Saul A. Teukolsky, William T. Vetterling: "Numerical Recipes in C", Cambridge University Press (1988).
12. Toshiki Tanaka, Tetsuo Hattori, Yusuke Kawakami, Yo Horikawa, Yoshiro Imai: "Parameters Estimation of Impulse Response in Compartment Model Using Cumulative Function and Linear Regression Analysis", IEEJ Transactions on Electronics, Information and Systems, Vol.143 No.2, pp.185-191, 2023.

Authors Introduction

Mr. Toshiki Tanaka

He received the Bachelor's degree in Electronics & Information Engineering from Kagawa University, Japan, in 2015. In the same year, he joined Kinkei System Co., Ltd. Currently, he is a doctoral student at Doshisha University, Japan, while working for the company. His research interests include circuit design, numerical simulation, and parameter estimation, etc. He is a member of IEEJ.

Dr. Ivan Tanev

He received the Ph.D. degree in Computer Engineering from Saint-Petersburg State Electrotechnical University, Russia, in 1993. He also received the degree of Dr. Eng. from Muroran Institute of Technology, Japan, in 2001. Currently, he is a Professor at Doshisha University, Japan. His research interests include Evolutionary Computations, Evolutionary Robotics, etc.

Dr. Tetsuo Hattori

He received the B.E. and M.E. degrees in Electrical Engineering from Doshisha University, and the Ph.D. degree in Electronics & Information Engineering from Osaka University, Japan. After he worked for Toshiba Eng. Co., Ltd., he had been with Kagawa University from 1988 to 2015. Currently, he is a Professor Emeritus at Kagawa University. Member of IEEJ and IEEE.

A Consideration on Amplification Function in BJT Evers-Moll Model and PTT (I)
---- V-I Characteristics ----

Shimon Hattori, Osamu Matoba
Kobe University, 1-1 Rokkodai-cho, Nada-ku, Kobe City 675-8061, Japan

Toshiki Tanaka
Kinkei System Corporation, 8-2-61 Nanko-Higashi, Suminoe-ku, Osaka 559-0031, Japan

Yusuke Kawakami
NIT Kagawa College, 355 Chokushi-cho, Takamatsu, Kagawa 761-8058, Japan

Tetsuo Hattori
Kagawa University, 1-1 Saiwai-cho, Takamatsu, Kagawa 760-0016, Japan
hattori@pe.kagawa-u.ac.jp

Abstract

In 1989, the notion of Photon Transport Transistor (PTT) has been proposed by B.J. Van Zeghbroeck et al., at IBM Research Center at that time. PTT is an optical coupling device of Light Emitting Diode (LED) and light receiving diode (Photo Diode, PD) where the carrier of the base layer is light (Photon) only. Later in 1996, W. N. Cheung and Paul J. Edwards have shown that PTT can be a very low noise amplifier in a positive feedback circuit, based on theoretical calculations of the noise figure by Quantum Mechanics. In this paper we consider the amplification principle noticing the similarity of the VI characteristics between the Bipolar Junction Transistor (BJT) and PTT.

Keywords, PTT, LED, PD, BJT, positive feedback circuit, Ebers-Moll model, VI characteristics

1. Introduction

As is well known, the Bipolar Junction Transistor (BJT) has been invented in 1948. The transistor has a structure in which two diodes, each made of a P-type and an N-type semiconductor, are further joined at the base layer.

On the other hand, B. J. Van Zeghbroeck et al. at IBM Research Laboratories presented the notion of a Photon Transport Transistor (PTT) based on optical coupling between a Light Emitting Diode (LED) and a Photo Diode (PD) [1]. Later in 1996, W. N. Cheung and Paul J. Edwards have shown that PTT can be a very low noise amplifier in a positive feedback circuit, based on theoretical calculations of the noise figure by Quantum Mechanics [2].

After that, it has been reported that, using the PTT that consists of high-brightness LED and PD ([3], [4], [5]). audio amplifier can be developed and that the PTT expresses not only an amplification function but also a switching function similar to a thyristor ([6], [7], [8]).

In this paper, we discuss the amplification function of PTT in comparison with the conventional BJT.

For this discussion, firstly, we show the similarity of the VI characteristics equation of BJT Ebers-Moll Model (BJT EMM) and PTT in the emitter common (or grounded) circuit as a positive feedback circuit. This is because, when using the BJT transistor for its amplification function, a common emitter circuit is generally used.

And secondary, we show the circuit equations in order to analyze how a fixed bias PTT circuit achieves DC current amplification function.

2. BJT EMM and PTT

2.1. Circuit and Characteristic Equation

The BJT emitter common (or grounded) circuit and its BJT EMM are illustrated in Fig. 1, Fig. 2, respectively. Furthermore, a PTT emitter common circuit for small signal amplification is shown in Fig. 3. In this section, we show substantial correspondence between the two devices of BJT EMM and PTT in their emitter common circuits, by representing the VI characteristics of their devices.

Shimon Hattori, Osamu Matoba, Toshiki Tanaka, Yusuke Kawakami, Tetsuo Hattori
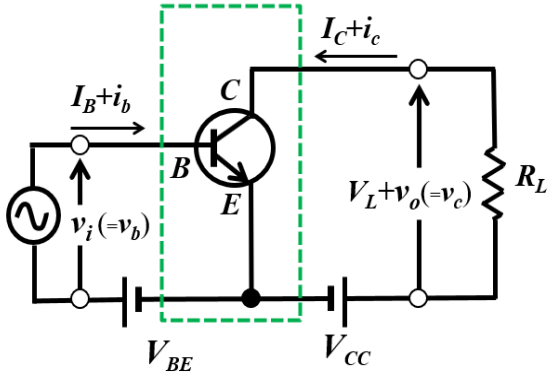
Fig. 1 BJT emitter common circuit for small signal voltage amplification.
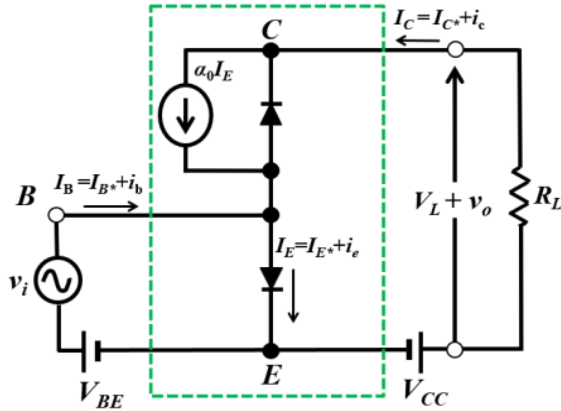
Fig. 2 BJT emitter common circuit that corresponds to the where $I_E = I_B + I_C$, based on Ebers-Moll model.

The VI characteristics equations of two diode parts in Fig. 2 are shown as Eq. (1) and Eq. (2). The each meaning of symbols such as I_{C0}, I_{E0}, q, k, T is shown in Table 1

$$I_C = \alpha_0 I_E - I_{C0}\left(\exp\left(\frac{qV_{BC}}{kT}\right) - 1\right) \qquad (1)$$

$$I_E = I_{E0}\left(\exp\left(\frac{qV_{BE}}{kT}\right) - 1\right) \qquad (2)$$

Table 1 The symbol list of the VI characteristic equation of the two diode parts in BJT Ebers-Moll Model.

I_{E0}	Reverse saturation current in Emitter	I_{C0}	Reverse saturation current in Collector
q	Charge	k	Boltzmann's constant
T	Absolute temperature	α_0	Current amplification factor
V_{BE}	Base-Emitter voltage	V_{BC}	Base-Collector voltage
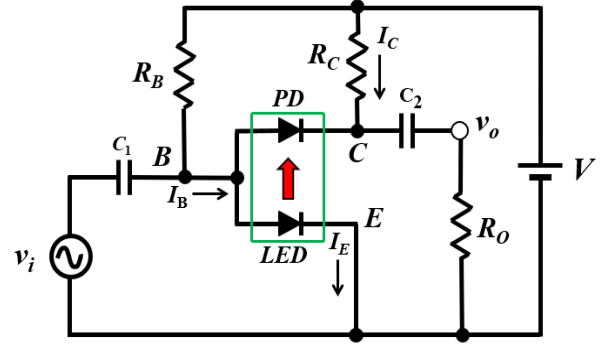
Fig. 3 Fixed bias PTT emitter common circuit for small signal voltage amplification.

The approximate VI characteristic equations for PD part and LED part of the PTT circuit in Fig. 3 are shown in Eq. (3) and Eq. (4), respectively. The meaning of each symbol is shown in Table 2.

We can point out that the pair of Eq. (3) and Eq. (4) is corresponding to that of Eq. (1) and Eq. (2) as the similarity between the BJT EMM and PTT. The difference of the two sets of pair equations is as follows.

$$I_C = \gamma^* I_E - I_{C0}\left(\exp\left(c_{PD}V_{BC}\right) - 1\right), \quad c_{PD} = \frac{q}{m_2 k T_{PD}} \qquad (3)$$

$$I_E = I_{E0}\left(\exp\left(c_{LED}V_{BE}\right) - 1\right), \quad c_{LED} = \frac{q}{m_1 k T_{LED}} \qquad (4)$$

Table 2 The symbol list of the VI characteristic equation of PTT circuit.

I_{E0}	Reverse saturation current in Emitter (LED)	I_{C0}	Reverse saturation current in Collector (PD)
m_1	Ideality factor (LED)	m_2	Ideality factor (PD)
q	Charge	k	Boltzmann's constant
T_{LED}	Absolute temperature of LED	T_{PD}	Absolute temperature of PD
V_{BE}	Base-Emitter voltage	V_{BC}	Base-Collector voltage
γ^*	Proportional constant	---	----

The Eq. (3) is an approximated VI characteristic equation of PD, based on the Eq. (5) that can be obtained from the PD equivalent circuit as shown in Fig. 4 ([4], [5]).

$$\begin{cases} I_O = I_L - I_S\left(\exp\frac{qV_D}{kT} - 1\right) - I' \\ V_D = V_{BC} - I_O R_S \\ I' = \frac{V_{BC} - I_O R_S}{R_{sh}} \end{cases} \qquad (5)$$
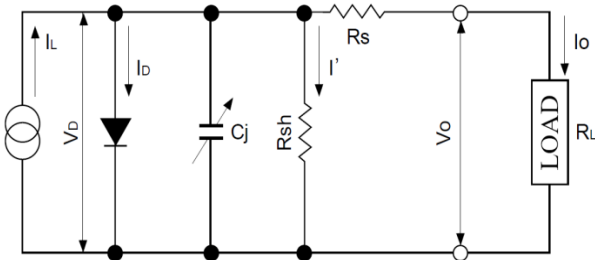
I_L : Generated current by incident light,
I_D : Current of diode, Cj : Junction capacitance,
Rsh : Shunt resistance, R_S : Series resistance,
I' : Current of Rsh, V_D : Voltage of diode,
I_O : Output current, V_O : Output voltage.

Fig. 4. Equivalent circuit of PD ([4], [5]), where the upper part corresponds to Base side and the lower Collector side.

The internal resistance Rs in Eq. (5) and in Fig. 4 is approximated as 0 because it is a small resistance, and $V_D = V_{BC} = -V_{CB}$. The resistance Rsh is a large resistance, so that I' is approximated as zero. Furthermore, I_L corresponds to Isc in Fig. 5 and represents the photocurrent proportional to the amount of light incident on the PD (amount of light received by the PD).
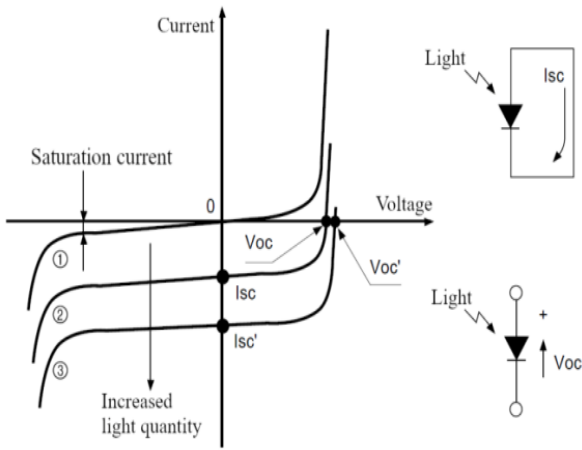
V_{OC} : open circuit voltage, Isc : Photocurrent

Fig.5. V-I characteristic of the equivalent circuit as shown in Fig. 4

In the PTT of Fig. 3, the LED and the PD are physically facing each other at a certain distance, so the amount of light received by the PD is considered to be proportional to the amount of light emitted from the LED. Since the amount of light emitted from the LED is proportional to the current I_E flowing in the LED, I_L can

be expressed as $I_L = \gamma^* I_E$, where γ^* is the proportionality constant. Then, we obtain the Eq. (3) as the result of introducing a correction term into the approximate equation derived from Eq. (5).

2.2. Similarity and Difference in Diode Equation

The similarity points between BJT and PTT: Both devices are composed of two diode parts that are collector-side diode and emitter-side one, and while the collector-side diode is in a state where the current IC flows in the opposite direction of the diode and includes the generated current as proportional to the current I_E of the emitter-side diode.

The differences in the equations for the two diodes in BJT and PTT: Absolute temperature T of two diodes is same in BJT EMM, while it is different in PTT. And the Ideality factors are 1 in BJT EMM, while they are different in PTT.

2.3. Direct Current Amplification from Equation of PTT Circuit

Under the assumption that the input signal voltage $v_i=0$ in Fig. 3, applying the Kirchhoff's law to the closed circuit 1 (V→A→B→E→V) and closed circuit 2 (A→B→C→A) in the PTT circuit of Fig. 3, we obtain the equation on direct current circuit as shown in Eq. (6).

$$\begin{cases} R_B I_B + V_{BE} = V \\ R_B I_B + V_{BC} - R_C I_C = 0 \end{cases} \quad (6)$$

On the other hand, if the ratio of current I_C and I_E is α, then $I_C = \alpha I_E$. This is a variable commonly used in BJT, and from $I_E = I_B + I_C$, $I_B = (1-\alpha) I_E$, and $I_C/I_B = \alpha/(1-\alpha)$ (current amplification factor: hFE, β). By rewriting Eq. (6) using this α, we have the following Eq. (7).

$$\begin{cases} R_B(1-\alpha)I_E + V_{BE} = V \\ \left(R_B - \alpha\left(R_B + R_C\right)\right)I_E + V_{BC} = 0 \end{cases} \quad (7)$$

On the other hand, by transforming the Eq. (3) and Eq. (4), we have another form of VI characteristic equation as in the following ([13], [14]).

$$\begin{cases} V_{BC} = \left(\dfrac{1}{c_{PD}}\right)\ln\left(\dfrac{(\gamma^* I_E - I_C)}{I_{C0}}+1\right) = \left(\dfrac{1}{c_{PD}}\right)\ln\left(\dfrac{(\gamma^* - \alpha)I_E}{I_{C0}}+1\right) \\ V_{BE} = \left(\dfrac{1}{c_{LED}}\right)\ln\left(\dfrac{I_E}{I_{E0}}+1\right) \end{cases}$$

$$(8)$$

Shimon Hattori, Osamu Matoba, Toshiki Tanaka, Yusuke Kawakami, Tetsuo Hattori

Assume that, when the power supply V increases to $V+\Delta V$, α and I_E change to $\alpha+\Delta\alpha$, $I_E+\Delta I_E$. Then. from the Eq. (7) and Eq. (8), we have the following Eq. (9) [13].

$$
\begin{cases}
\Delta\alpha = \left(\dfrac{K_2}{K_1}\right)\Delta I_E \\[2ex]
\Delta I_E = \dfrac{\Delta V}{\left(R_B(1-\alpha)+\dfrac{1}{c_{LED}I_E}-\left(R_B I_E\right)\left(\dfrac{K_2}{K_1}\right)\right)}
\end{cases}
$$

where

$$
\begin{cases}
K_1 = \left((R_B+R_C)+\left(\dfrac{1}{c_{PD}}\right)\dfrac{1}{(\gamma^*-\alpha)I_E+I_{C0}}\right)I_E \\[2ex]
K_2 = \left(R_B-(R_B+R_C)\alpha+\left(\dfrac{1}{c_{PD}}\right)\dfrac{(\gamma^*-\alpha)}{(\gamma^*-\alpha)I_E+I_{C0}}\right)
\end{cases}
$$

$$(9)$$

Eq. (9) shows that the amplification factor α and emitter direct current I_E will increase according to the increase of the DC voltage V. However, we find out that when the relation Eq. (10) holds, $\Delta\alpha=0$, even if $\Delta V>0$ and $\Delta I_E>0$.

$$
\alpha = \gamma^* = \frac{R_B}{R_B+R_C} \tag{10}
$$

At that time, from Eq. (8), $V_{BC} = 0$. When $V_{BC} = 0$, from our experiments for PTT with positive feedback circuit as shown in Fig. 3, we have obtained the direct current amplification factor $\beta = 999$ ($\alpha=0.999$) [13], [14].

We consider that the same situation can happen in the fixed bias emitter common circuit of BJT EMM, because of the similarity of VI characteristic equation.

3. Conclusion

In this paper, from the similarity of the VI characteristics between the BJT EMM and PTT, we have discussed the expression of direct current amplification function of PTT. As a further study, we will analyze the voltage amplification function for input AC signal of the PTT circuit based on the VI characteristics.

References

1. B.J. Van Zeghbroeck, Ch. Herder, H. P. Meiner and W. Walter, Photon transport transistor, Proc. of 1989 IEEE International Electron Devices Meeting (IEDM), pp.543-546 (1989-12)
2. W. N. Cheung and Paul J. Edwards, Characteristics of a Class of New Opto-Coupler Amplifiers with Positive Feedback, IEEE Journal of QUANTUM ELECTRONICS, Vol.32, No.3, pp.502-506 (1996-3)
3. OSRAM Opto-Semiconductors GmbH, OSLON® SSL 150, LH CPDP Datasheet (https://dammedia.osram.info /media/resource/hires/osram-dam-2493913/LH%20CPD P.pdf)
4. Hamamatsu Photonics K.K., Si PIN photodiode S5107 Datasheet (https://www.hamamatsu.com/resources/pdf/ ssd/s5106_etc_kpin1033e.pdf)
5. Hamamatsu Photonics K.K., Characteristic and Usage (https://www.hamamatsu.com/resources/pdf/ssd/si_pd_te chinfo.pdf)
6. Junichi Fujita, Kensho Okamoto, and Tetsuo Hattori, Development of Novel Transistor and Thyristor composed of LED and photodiode, Japanese Journal of Applied Physics, ISSN 1347-4065, Vol.53, 05FEB18, p.1-6 (2014-4)
7. Junichi Fujita, Daisuke Sato, Kensho Okamoto, and Tetsuo Hattori, Novel Opt-coupling Transistor by LED and PD and its Application, Journal of Robotics, Networking and Artificial Life, ISSN 2352-6386, Vol.1, No.3 pp.244-248 (2014-12)
8. Junichi FUJITA, Yusuke KAWAKAMI, Katsushi TAKAHASHI, Kensho OKAMOTO, and Tetsuo HATTORI, Novel Audio Amplifier and its Kansei Evaluation, International Journal of Affective Engineering, ISSN 2187-5413, Vol.13, No.4 pp.245-251 (2014-12)
9. Katsuari Kamei, Fundamentals of Electronic Circuits for Students, ISBN 978-4-320-08638-8, Kyoritsu Shuppan (2010-3). (in Japanese)
10. Nobuo Fujii ed., Handy-book Electronics, ISBN4-274-03471-2, Ohmsha (2002-2). (in Japanese)
11. Akira Matsuzawa, Fundamental Electronics Circuits Engineering---Focusing on Analog Circuits ---, IEEJ University Course, ISBN978-4-88686-276-1, IEEJ, Ohmsha (2014-2). (In japanese)
12. Isao Dendo, Basic Electronic Circuits, http://www.mu-sic.tohoku.ac.jp/attach/ kiso-denshikairo.pdf (2020-11).
13. Shimon Hattori, Osamu Matoba, Yusuke Kawakami, Toshiki Tanaka, Tetsuo Hattori, Consideration on the Expression of Amplification Function of PTT with Positive Feedback Circuit Based on Approximated Characteristic Equations of LED and PD, IEEJ Trans. on Electronics, Information and Systems, Vol.143, No.2, pp.165-171 (2023-2) (in Japanese)
14. Shimon Hattori, Osamu Matoba, Yusuke Kawakami, Toshiki Tanaka, Tetsuo Hattori, Current Amplification Function of PTT in Positive Feedback Circuit, Proc. of

Authors Introduction

Mr. Shimon Hattori

He received the B.E. and M.E. degrees in Electrical and Electronics Engineering from Oita University, Japan, in 2010 and 2012, respectively. He is presently a doctoral student at Kobe University, while he is working for a company. He is interested in information processing and computing mechanism using optoelectronic devices. Member of IEEJ.

Dr. Osamu Matoba

He received the Ph.D. (Eng.) degree, from Osaka University, Japan, in 1996. From April 1996 to August 2002, he worked as an Assistant at the Institute of Industrial Science, the University of Tokyo, Japan. Since September 2002, he has been with Kobe University. He is currently a Professor at the Graduate School of Systems and Informatics, also at Kobe University Center for Advanced Interdisciplinary Research. His research interests include measurement optics, information optics, and optoelectronics. Member of IEEE, JSAP, OSJ, Senior Member of LSJ, and also a Fellow of SPIE and OSA.

Mr. Toshiki Tanaka

He received the B.E. degree in Electronics and Information Engineering from Kagawa University, Japan, in 2015. In the same year, he joined Kinkei System Co., Ltd. Currently, while he is engaged in measurement and data management of electrical and electronic circuit systems at the company, he is studying as a doctoral student of Doshisha University, Japan. His research interests include numerical simulation and parameter estimation. Member of IEEJ.

Dr. Yusuke Kawakami

He received the B.E., M.E., and Ph.D. degrees in Information System Engineering from Kagawa University, Japan, in 2009, 2011, and 2014, respectively. Since 2020, he has been with National Institute of Technology (NIT) Kagawa College, Japan. Currently, he is an Assistant Lecturer in NIT Kagawa College. His research interests include Kansei Engineering, electronics and machine control, image processing and sound signal processing. Member of IEEJ.

Dr. Tetsuo Hattori

He received the B.E. and M.E. degrees in Electrical Engineering from Doshisha University, Japan, and the Ph.D. degree in Electronics and Information Engineering from Osaka University, Japan. After he worked for Toshiba Eng. Co., Ltd., he had been with Kagawa University from 1988 to 2015. Currently, he is a Professor Emeritus at Kagawa University. His research interests include STEM education, signal & image analysis, and hybrid computer, etc. Member of IEEJ and IEEE.

A Consideration on Amplification Function in BJT Evers-Moll Model and PTT (II) ---- H Parameters in the Small Signal Amplifier Circuit----

Shimon Hattori, Osamu Matoba
Kobe University, 1-1 Rokkodai-cho, Nada-ku, Kobe City 675-8061, Japan

Toshiki Tanaka
Kinkei System Corporation, 8-2-61 Nanko-Higashi, Suminoe-ku, Osaka 559-0031, Japan

Yusuke Kawakami
NIT Kagawa College, 355 Chokushi-cho, Takamatsu, Kagawa 761-8058, Japan

Tetsuo Hattori
Kagawa University, 1-1 Saiwai-cho, Takamatsu, Kagawa 760-0016, Japan
hattori@pe.kagawa-u.ac.jp

Abstract

The notion of PTT (Photon Transport Transistor) has been proposed in 1989 as an optical coupling device of light emitting diode (LED) and light receiving diode (Photo Diode, PD), where the carrier of the base layer is light (Photon) only. In this paper, in order to deal with the various applications of PTT circuit more theoretically, based on the approximate VI characteristic equation of LED and PD, we newly derive the h parameters in PTT emitter common circuit, while referring to the h parameter of Bipolar Junction Transistor (BJT).

Keywords: PTT, LED, PD, positive feedback circuit, amplification function, equivalent circuit, h parameters

1. Introduction

The notion of Photon Transport Transistor (PTT) has been presented by B. J. Van Zeghbroeck et al. at IBM Research Laboratories presented in 1989 {1}. The PTT consists of the optical coupling between Light Emitting Diode (LED) or Laser Diode and Photo Diode (PD). Moreover in 1996, it has been theoretically shown that the PTT can be a very low-noise transistor-like device with an amplification function in a positive feedback circuit [2].

Later, an audio amplifier prototype using PTT positive feedback circuit with optical coupling between high-brightness LED and PD was developed, and it was reported that the PTT even exhibited functions similar to those of a thyristor ([3], [4], [5], [6], [7], [8]). Also, while referencing the conventional studies on BJT ([9], [10], [11], [12]), authors have analyzed the current amplification function of the PTT based on experimental results ([13], [14]).

In this paper, in order to deal with the various applications of PTT circuit more theoretically, based on the approximate VI characteristic equation of LED and PD, we derive the h parameters in PTT emitter common circuit, while referring to the h parameter of Bipolar Junction Transistor (BJT) ([9], [10], [12]).

Moreover, from the similarities between PTT and BJT Ebers-Moll Model (EMM), we discuss the essential factors of amplification function in PTT and BJT.

2. PTT and BJT EMM in Emitter Common Circuit
2.1. VI Characteristic of PTT and BJT EMM

The PTT with fixed bias in emitter common circuit is illustrated in Fig. 1. Generally, the emitter common circuits are used for small signal voltage amplification. The VI characteristic equations for each of the PD and LED that make up the PTT are shown in Eq. (1) and Eq. (2), respectively ([13], [14]). Each meaning of the symbols in those equations is shown in Table 1. As for the BJT and BJT EMM, those circuits without fixed bias in emitter common are illustrated in Fig. 2(a) and Fig. 2(b). respectively. And, the VI characteristic equations of two diode parts of the BJT EMM are shown in Eq. (3)

Shimon Hattori, Osamu Matoba, Toshiki Tanaka, Yusuke Kawakami, Tetsuo Hattori

and Eq. (4), respectively, where α_0 of Eq. (3) means the current amplification constant.
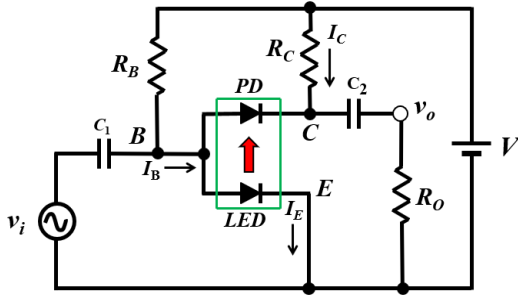
Fig.1. PTT emitter common circuit.

$$I_C = \gamma^* I_E - I_{C0}\left(\exp(c_{PD} V_{BC}) - 1\right), \quad c_{PD} = \frac{q}{m_2 k T_{PD}} \quad (1)$$

$$I_E = I_{E0}\left(\exp(c_{LED} V_{BE}) - 1\right), \quad c_{LED} = \frac{q}{m_1 k T_{LED}} \quad (2)$$

Table 1. The symbol list of PTT characteristic formula.

I_{E0}	Reverse saturation current in Emitter (LED)	I_{C0}	Reverse saturation current in Collector (PD)
m_1	Ideality factor (LED)	m_2	Ideality factor (PD)
q	Charge	k	Boltzmann's constant
T_{LED}	Absolute temperature of LED	T_{PD}	Absolute temperature of PD
V_{BE}	Base-Emitter voltage	V_{BC}	Base-Collector voltage
γ^*	Proportional constant	---	----

(a)

(b)

Fig.2. (a) BJT emitter common circuit for small signal voltage amplification. (b) BJT EMM without fixed bias where α_0 means the current amplification constant, and $I_E = I_B + I_C$.

$$I_C = \alpha_0 I_E - I_{C0}\left(\exp\left(\frac{qV_{BC}}{kT}\right) - 1\right) \quad (3)$$

$$I_E = I_{E0}\left(\exp\left(\frac{qV_{BE}}{kT}\right) - 1\right) \quad (4)$$

Comparing the pair of Eq. (3) and Eq. (4) with that of Eq. (1) and Eq. (2). we can expect that, the h parameters of PTT circuit using Eq. (3) and Eq. (4) will become the h parameters of BJT EMM by setting $\gamma^* = \alpha$, $m_1 = 1$, $m_2 = 1$. Then, we only have to find out the h parameters of PTT.

2.2. H Parameter for Small Signal Equivalent Equivalent Circuit

As is well known, for small signal amplification circuit of BJT in emitter common circuit, the small signal equivalent circuit with a set of h parameters (h_{ie}, h_{re}, h_{fe}, and h_{oe}) is used as shown in Fig. 3 ([9], [10], [12]).
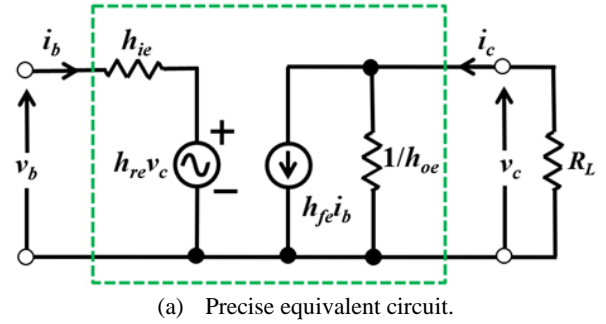
(a) Precise equivalent circuit.
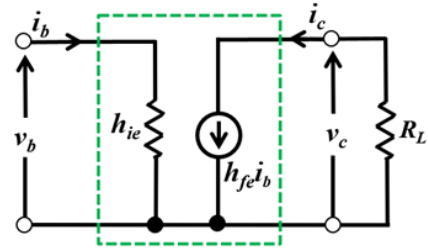
(b) Simplified equivalent circuit.

Fig. 3. Equivalent circuit of BJT emitter common with h parameters.

The equivalent relation equations in Fig. 3 are as follows.

$$\begin{bmatrix} v_b \\ i_c \end{bmatrix} = \begin{bmatrix} h_{ie} i_b + h_{re} v_c \\ h_{fe} i_b + h_{oe} v_c \end{bmatrix} = \begin{bmatrix} h_{ie}, & h_{re} \\ h_{fe}, & h_{oe} \end{bmatrix} \cdot \begin{bmatrix} i_b \\ v_c \end{bmatrix} \quad (5)$$

These h parameters are obtained from the total differential equation for V_{BE} and I_C, assuming the following functional relationship between f_1 and f_2.

Since we assume that $V_{BE} = f_1(I_B, V_{CE})$, $I_C = f_2(I_B, V_{CE})$,

$$
\begin{cases}
\Delta V_{BE}(\equiv v_b) = \left(\dfrac{\partial V_{BE}}{\partial I_B}\right)\Delta I_B + \left(\dfrac{\partial V_{BE}}{\partial V_{CE}}\right)\Delta V_{CE} = h_{ie} i_b + h_{re} v_c \\[2ex]
\Delta I_C(\equiv i_c) = \left(\dfrac{\partial I_C}{\partial I_B}\right)\Delta I_B + \left(\dfrac{\partial I_C}{\partial V_{CE}}\right)\Delta V_{CE} = h_{fe} i_b + h_{oe} v_c
\end{cases}
$$

$$(6)$$

First of all, as for h_{re}, we find the following relationship between the other h parameter, h_{oe}.

$$
h_{re} = \left.\frac{\partial V_{BE}}{\partial V_{CE}}\right|_{I_B = const} = \left(\frac{\partial V_{BE}}{\partial I_C}\right)\left(\frac{\partial I_C}{\partial V_{CE}}\right) = \left(\frac{\partial V_{BE}}{\partial I_E}\right)\left(\frac{\partial I_E}{\partial I_C}\right)\left(\frac{\partial I_C}{\partial V_{CE}}\right)
$$
$$
= \left(\frac{\partial I_E}{\partial V_{BE}}\right)^{-1} \cdot \left(\frac{\partial I_E}{\partial I_C}\right) \cdot h_{oe}
$$

$$(7)$$

Since

$$
\frac{\partial I_E}{\partial V_{BE}} = \frac{\partial I_{E0}\left(\exp(C_{LED}V_{BE}) - 1\right)}{\partial V_{BE}} = I_{E0} \cdot C_{LED} \cdot \exp(C_{LED}V_{BE})
$$
$$
= C_{LED} \cdot (I_E + I_{E0}) \approx C_{LED} \cdot I_E
$$

$$(8)$$

and

$$
\left(\frac{\partial I_E}{\partial I_C}\right) = \left(\frac{\partial(I_B + I_C)}{\partial I_C}\right) = 1
$$

$$(9)$$

then we have

$$
h_{re} = \frac{h_{oe}}{C_{LED}(I_E + I_{E0})} \approx \frac{h_{oe}}{C_{LED} \cdot I_E}
$$

$$(10)$$

Since $I_E = I_B + I_C$ in Eq. (1),

$$
I_C = \gamma^*(I_B + I_C) - I_{C0}\left(\exp(c_{PD}V_{BC}) - 1\right), \quad c_{PD} = \frac{q}{m_2 k T_{PD}}
$$

$$(11)$$

then

$$
I_C = \frac{\gamma^* I_B}{1 - \gamma^*} - \frac{I_{C0}\left(\exp(c_{PD}V_{BC}) - 1\right)}{1 - \gamma^*}
$$
$$
= \frac{\gamma^* I_B}{1 - \gamma^*} - \frac{I_{C0}\left(\exp(-c_{PD}(V_{CE} - V_{BE})) - 1\right)}{1 - \gamma^*}
$$

$$(12)$$

Therefore, we have the h_{fe} of PTT as in the following.

$$
h_{fe} = \frac{\partial I_C}{\partial I_B} = \frac{\gamma^*}{1 - \gamma^*} - \frac{I_{C0}}{1 - \gamma^*}\left(\frac{\partial \exp(C_{PD}V_{BC})}{\partial I_B}\right)
$$
$$
= \frac{\gamma^*}{1 - \gamma^*} - \frac{I_{C0}}{1 - \gamma^*}\left(\exp(-C_{PD}V_{CE}) \cdot \frac{\partial \exp(C_{PD}V_{BE})}{\partial V_{BE}} \cdot \frac{\partial V_{BE}}{\partial I_B}\right)
$$
$$
= \frac{\gamma^*}{1 - \gamma^*} - \frac{I_{C0}}{1 - \gamma^*} C_{PD} \cdot \exp(C_{PD}V_{BC}) \cdot h_{ie}
$$

$$(13)$$

Similarly, we have h_{oe} as shown in the following.

$$
h_{oe} = \frac{\partial I_C}{\partial V_{CE}} = \frac{\partial\left(\dfrac{\gamma^*}{1-\gamma^*}I_B - \dfrac{I_{C0}}{1-\gamma^*}\left(\exp(C_{PD}V_{BC}) - 1\right)\right)}{\partial V_{CE}}
$$
$$
= \frac{\left(\dfrac{-I_{C0}}{1-\gamma^*}\right)\partial \exp(C_{PD}(V_{BE} - V_{CE}))}{\partial V_{CE}}
$$
$$
= \frac{\left(\dfrac{-I_{C0}}{1-\gamma^*}\right) \cdot \partial\left(\exp(C_{PD}V_{BE}) \cdot \exp(-C_{PD}V_{CE})\right)}{\partial V_{CE}}
$$
$$
= \left(\frac{-I_{C0}}{1-\gamma^*}\right) \cdot \left\{ \frac{\partial(\exp(C_{PD}V_{BE}))}{\partial V_{CE}} \cdot \exp(-C_{PD}V_{CE}) + \exp(C_{PD}V_{BE}) \cdot \frac{\partial(\exp(-C_{PD}V_{CE}))}{\partial V_{CE}} \right\}
$$
$$
= \left(\frac{C_{PD} \cdot I_{C0}}{1-\gamma^*}\right) \cdot \left\{ 1 - \left(\frac{\partial V_{BE}}{\partial V_{CE}}\right) \right\} \cdot \exp(C_{PD}V_{BC})
$$
$$
= \left(\frac{C_{PD} \cdot I_{C0}}{1-\gamma^*}\right) \cdot \{1 - h_{re}\} \cdot \exp(C_{PD}V_{BC})
$$

$$(14)$$

As for h_{ie}, from Eq. (6), we have

$$
h_{ie} = \frac{\partial V_{BE}}{\partial I_B} = \left(\frac{\partial V_{BE}}{\partial I_E}\right)\left(\frac{\partial I_E}{\partial I_B}\right) = \left(\frac{\partial I_E}{\partial V_{BE}}\right)^{-1}\left(\frac{\partial(I_C + I_B)}{\partial I_B}\right)
$$
$$
= \left(\frac{\partial I_E}{\partial V_{BE}}\right)^{-1}\left(h_{fe} + 1\right) = \frac{\left(h_{fe} + 1\right)}{C_{LED}(I_E + I_{E0})} \approx \frac{\left(h_{fe} + 1\right)}{C_{LED} \cdot I_E}
$$

$$(15)$$

So far, we have obtained the following equations for the four parameters, h_{ie}, h_{fe}, h_{re} and h_{oe}.

Shimon Hattori, Osamu Matoba, Toshiki Tanaka, Yusuke Kawakami, Tetsuo Hattori

$$
\begin{cases}
h_{ie} = \dfrac{\partial V_{BE}}{\partial I_B} = \dfrac{(h_{fe}+1)}{C_{LED}(I_E+I_{E0})} \\[12pt]
h_{fe} = \dfrac{\partial I_C}{\partial I_B} = \dfrac{\gamma^*}{1-\gamma^*} - \dfrac{I_{C0}}{1-\gamma^*} C_{PD} \cdot \exp(C_{PD}V_{BC}) \cdot h_{ie} \\[12pt]
h_{oe} = \dfrac{\partial I_C}{\partial V_{CE}} = \left(\dfrac{C_{PD} \cdot I_{C0}}{1-\gamma^*}\right) \cdot \{1 - h_{re}\} \cdot \exp(C_{PD}V_{BC}) \\[12pt]
h_{re} = \dfrac{\partial V_{BE}}{\partial V_{CE}} = \dfrac{h_{oe}}{C_{LED}(I_E+I_{E0})}
\end{cases}
\tag{16}
$$

From Eq. (16), we can see the recursive relation between h_{ie} and h_{fe}, and another between h_{re} and h_{oe}. So, we derive those parameters in the details, as follows.

$$
\begin{aligned}
h_{fe} &= \frac{\gamma^*}{1-\gamma^*} - \frac{I_{C0}}{1-\gamma^*} C_{PD} \cdot \exp(C_{PD}V_{BC}) \cdot h_{ie} \\
&= \frac{\gamma^*}{1-\gamma^*} - \frac{I_{C0}}{1-\gamma^*} C_{PD} \cdot \exp(C_{PD}V_{BC}) \cdot \frac{(h_{fe}+1)}{C_{LED}(I_E+I_{E0})}
\end{aligned}
$$

So, we define the following K.

$$
K \; \square \; \frac{I_{C0}}{1-\gamma^*} C_{PD} \cdot \exp(C_{PD}V_{BC}) \cdot C_{LED}(I_E+I_{E0})^{-1}
$$

Then we have

$$
\begin{aligned}
h_{fe} &= \left(\frac{\gamma^*}{1-\gamma^*} - K\right) \cdot (1+K)^{-1} \\
&= \frac{\dfrac{\gamma^*}{1-\gamma^*} - \dfrac{I_{C0}}{1-\gamma^*} C_{PD} \cdot \exp(C_{PD}V_{BC}) \cdot (C_{LED}(I_E+I_{E0}))^{-1}}{1 + \dfrac{I_{C0}}{1-\gamma^*} C_{PD} \cdot \exp(C_{PD}V_{BC}) \cdot (C_{LED}(I_E+I_{E0}))^{-1}} \\
&= \frac{\dfrac{\gamma^* C_{LED}(I_E+I_{E0})}{1-\gamma^*} - \dfrac{I_{C0}}{1-\gamma^*} C_{PD} \cdot \exp(C_{PD}V_{BC})}{C_{LED}(I_E+I_{E0}) + \dfrac{I_{C0}}{1-\gamma^*} C_{PD} \cdot \exp(C_{PD}V_{BC})}
\end{aligned}
\tag{17}
$$

Similarly, as for h_{re}, we have the following equation.

$$
h_{re} = \frac{\left(\dfrac{C_{PD} \cdot I_{C0}}{1-\gamma^*}\right) \cdot \exp(C_{PD}V_{BC})}{C_{LED} \cdot I_{E0} \cdot \exp(C_{LED}V_{BE}) + \left(\dfrac{C_{PD} \cdot I_{C0}}{1-\gamma^*}\right) \cdot \exp(C_{PD}V_{BC})}
\tag{18}
$$

If we consider the case when V_{BC} approximately equals

0, then the term $\exp(C_{PD}V_{BC})$ nearly becomes 1.0. In this case we have the more simplified four parameters as follows that are corresponding to Fig. 3(b).

$$
h_{ie} = \frac{\left(\dfrac{1}{1-\gamma^*}\right)}{C_{LED}(I_E+I_{E0})} = \frac{\left(\dfrac{1}{1-\gamma^*}\right)}{C_{LED} \cdot I_{E0} \exp(C_{LED}V_{BE})},
$$

$$
h_{fe} \approx \frac{\gamma^*}{1-\gamma^*}, \; h_{re} \approx 0, \; h_{oe} \approx 0
\tag{19}
$$

3. Conclusion

We have newly derived the four h-parameters in the PTT small signal circuit, based on the VI characteristic equation of LED and PD that composes the PTT. From those computed h parameters, we find out that PPT functions the amplifier, because approximated and simplified h parameters are very similar to the case of BJT.

Acknowledgements

References

1. B.J. Van Zeghbroeck, Ch. Herder, H. P. Meiner and W. Walter, Photon transport transistor, Proc. of 1989 IEEE International Electron Devices Meeting (IEDM), pp.543-546 (1989-12)
2. W. N. Cheung and Paul J. Edwards, Characteristics of a Class of New Opto-Coupler Amplifiers with Positive Feedback, IEEE Journal of QUANTUM ELECTRONICS, Vol.32, No.3, pp.502-506 (1996-3)
3. Junichi Fujita, Kensho Okamoto, and Tetsuo Hattori, Development of Novel Transistor and Thyristor composed of LED and photodiode, Japanese Journal of Applied Physics, ISSN 1347-4065, Vol.53, 05FEB18, p.1-6 (2014-4)
4. Junichi Fujita, Daisuke Sato, Kensho Okamoto, and Tetsuo Hattori, Novel Opt-coupling Transistor by LED and PD and its Application, Journal of Robotics, Networking and Artificial Life, ISSN 2352-6386, Vol.1, No.3 pp.244-248 (2014-12)
5. Junichi FUJITA, Yusuke KAWAKAMI, Katsushi TAKAHASHI, Kensho OKAMOTO, and Tetsuo HATTORI, Novel Audio Amplifier and its Kansei Evaluation, International Journal of Affective Engineering, ISSN 2187-5413, Vol.13, No.4 pp.245-251 (2014-12).
6. OSRAM Opto-Semiconductors GmbH, OSLON® SSL 150, LH CPDP Datasheet (https://dammedia.osram.info/media/resource/hires/osram-dam-2493913/LH%20CPDP.pdf)

7. Hamamatsu Photonics K.K., Si PIN photodiode S5107 Datasheet (https://www.hamamatsu.com/resources/pdf/ssd/s5106_etc_kpin1033e.pdf)

8. Hamamatsu Photonics K.K., Characteristic and Usage (https://www.hamamatsu.com/resources/pdf/ssd/si_pd_techinfo.pdf)

9. Nobuo Fujii ed., Handy-book Electronics, ISBN4-274-03471-2, Ohmsha (2002-2). (in Japanese)

10. Katsuari Kamei, Fundamentals of Electronic Circuits for Students, ISBN 978-4-320-08638-8, Kyoritsu Shuppan (2010-3). (in Japanese)

11. Akira Matsuzawa, Fundamental Electronics Circuits Engineering---Focusing on Analog Circuits ---, IEEJ University Course, ISBN978-4-88686-276-1, IEEJ, Ohmsha (2014-2). (In japanese)

12. Isao Dendo, Basic Electronic Circuits, http://www.music.tohoku.ac.jp/attach/ kiso-denshikairo.pdf (2020-11).

13. Shimon Hattori, Osamu Matoba, Yusuke Kawakami, Toshiki Tanaka, Tetsuo Hattori, Consideration on the Expression of Amplification Function of PTT with Positive Feedback Circuit Based on Approximated Characteristic Equations of LED and PD, IEEJ Trans. on Electronics, Information and Systems, Vol.143, No.2, pp.165-171 (2023-2) (in Japanese)

14. Shimon Hattori, Osamu Matoba, Yusuke Kawakami, Toshiki Tanaka, Tetsuo Hattori, Current Amplification Function of PTT in Positive Feedback Circuit, Proc. of 2023 IEEE 12th Global Conference on Consumer Electronics (GCCE), pp.84-88, Oct., 2023.

Authors Introduction

Mr. Shimon Hattori

He received the B.E. and M.E. degrees in Electrical and Electronics Engineering from Oita University, Japan, in 2010 and 2012, respectively. He is presently a doctoral student at Kobe University, while he is working for a company. He is interested in information processing and computing mechanism using optoelectronic devices. Member of IEEJ.

Dr. Osamu Matoba

He received the Ph.D. (Eng.) degree, from Osaka University, Japan, in 1996. From April 1996 to August 2002, he worked as an Assistant at the Institute of Industrial Science, the University of Tokyo, Japan. Since September 2002, he has been with Kobe University. He is currently a Professor at the Graduate School of Systems and Informatics, also at Kobe University Center for Advanced Interdisciplinary Research. His research interests include measurement optics, information optics, and optoelectronics. Member of IEEE, JSAP, OSJ, Senior Member of LSJ, and also a Fellow of SPIE and OSA.

Mr. Toshiki Tanaka

He received the B.E. degree in Electronics and Information Engineering from Kagawa University, Japan, in 2015. In the same year, he joined Kinkei System Co., Ltd. Currently, while he is engaged in measurement and data management of electrical and electronic circuit systems at the company, he is studying as a doctoral student of Doshisha University, Japan. His research interests include numerical simulation and parameter estimation. He is a member of IEEJ.

Dr. Yusuke Kawakami

He received the B.E., M.E., and Ph.D. degrees in Information System Engineering from Kagawa University, Japan, in 2009, 2011, and 2014, respectively. Since 2020, he has been with National Institute of Technology (NIT) Kagawa College, Japan. Currently, he is an Assistant Lecturer in NIT Kagawa College. His research interests include Kansei Engineering, electronics and machine control, image processing and sound signal processing. He is a member of IEEJ.

Dr. Tetsuo Hattori

He received the B.E. and M.E. degrees in Electrical Engineering from Doshisha University, and the Ph.D. degree in Electronics & Information Engineering from Osaka University, Japan. After he worked for Toshiba Eng. Co., Ltd., he had been with Kagawa University from 1988 to 2015. Currently, he is a Professor Emeritus at Kagawa University. His research interests include STEM education based on ETT, signal and image analysis, pattern recognition, and realization of hybrid computer. Member of IEEJ and IEEE.

Color Image Arrangement Using Histogram Matching

Yusuke Kawakami

National Institute of Technology (Kagawa College), 355 Chokushi-cho, Takamatsu, 761-8058, Japan
Email: riverjp2002@gmail.com, kawakami-y@t.kagawa-nct.ac.jp

Tesuo Hattori

Kagawa University, 1-1 Saiwai-cho, Takamatsu, 760-8521, Japan
Email: hattori@pe.kagawa-u.ac.jp

R. P. C. Janaka Rajapakse

Tainan National University of the Arts, 66 Daqi, Guantian Dist., Tainan 72045, Taiwan (R.O.C.)
Email: janakaraja@gmail.com

Abstract

For the arrangement processing of image using its histogram, we previously have presented a Histogram Matching method based on Gaussian Distribution (HMGD). However, in the case where the brightness histogram of input image has multiple peaks, the HMGD processing does not always bring good results. In this paper, we present an improved histogram matching method using the reference histogram that is made by appropriate moving average (HMMA) processing over the histogram of input image. Also in this paper, we show the experimental results.

Keywords: Image processing, Histogram matching (HM), Moving average, Kansei impression

1. Introduction

Recently, automated image arrangement processing such as enhancement of images that we have already reported is usually used in the various imaging devices, for example, Digital Camera, Smart Phone and so on [1], [2], [3].

In our previous papers, we have presented a Histogram Matching based on Gaussian Distribution (HMGD) processing for the image arrangement [4], [5], [6], [7]. And we illustrated that HMGD processing can improve feeling (or Kansei) impression of the original image [4], [5], [6], [7].

In this paper, we propose an improved Histogram Matching method using the reference histogram that is made by appropriate Moving Average (HMMA) processing over the histogram of input image [8]. Also, in this paper, we provide the experimental results for some color images and show its effectiveness of the proposed method.

2. Principle
2.1. Histogram Matching as Example of Gaussian Distribution [9]

In the section, we describe the principle of histogram matching method as an example of Histogram Matching based on Gaussian Distribution (HMGD).

Fig. 1 shows the conceptual image of HMGD. Let $f(x)$ and $h(y)$ be two probabilistic density functions (PDF) on real variables x and y, respectively. The PDF is corresponding to histogram of image brightness level which is discretely defined.

In addition, let $y=\phi(x)$ be a continuous and monotonic increase function corresponding to cumulative histogram of image brightness level between variables x and y. And let $y=\phi(x)$ be defined by Eq. (1).

$$y = \phi(x) + L\int_{0}^{x} f(x)\,dx \quad\text{................................ (1)}$$

At first, we have to expand brightness level of original image histogram and convert into uniform distribution histogram, because we aim to match Gaussian distribution. From Eq. (1) and Fig. 1, we can derive Eq. (2) and (3).

$$f(x) = h(y)\varphi'(x) = h(y)Lf(x) \quad\text{............. (2)}$$

$$h(y) = \frac{1}{L} \quad\text{... (3)}$$

Yusuke Kawakami, Tesuo Hattori, R. P. C. Janaka Rajapakse

We understand the histogram of original image $f(x)$ becomes uniform distribution $h(y)$ by Eq. (3). This means that brightness level of original image $f(x)$ is expanded to $h(y)$.

Then, let $g(z)$ and $\gamma(z)$ be the function that is defined by Eq. (4) and Eq. (5), respectively.

$$g(z) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(z-\mu)^2}{2\sigma^2}\right) \quad \text{............} \quad (4)$$

$$y = \gamma(z) = L\int_0^z g(z)\,dz$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \int_0^z \exp\left(-\frac{(z-\mu)^2}{2\sigma^2}\right) dz \quad \text{...........} \quad (5)$$

Here, Fig. 1 shows the relationship between $y=\phi(x)$ and $y=\gamma(z)$. So we can be obtained following Eq. (6) and Eq. (7).

$$L\int_0^x f(x)\,dx = \frac{L}{\sqrt{2\pi\sigma^2}} \int_0^z \exp\left(-\frac{(z-\mu)^2}{2\sigma^2}\right) dz \quad \text{....} \quad (6)$$

$$\frac{d}{dx} L\int_0^x f(x)\,dx = \frac{d}{dz}\frac{L}{\sqrt{2\pi\sigma^2}} \int_0^z \exp\left(-\frac{(z-\mu)^2}{2\sigma^2}\right) dz. \quad \text{..} \quad (7)$$

If we perform Eq. (7),

$$L\varphi'(x) = L\gamma'(z), \quad f(x) \Rightarrow g(z) \quad \text{..........} \quad (8)$$

That is, we understand that $f(x)$ becomes Gaussian distribution $g(z)$ when we take the transform function as Eq. (1) and Eq. (5).

Thus, histogram matching processing is the function which defined by cumulative histogram transformation the original histogram into arbitrary histogram.

3. Experimentation

Fig. 2 and Fig. 3 show the example of results and the corresponding histogram respectively. In this case, we understand that HMGD image is very enhancing brightness and contrast than original image.

However, the touch of brush detail transforms unnatural in this image. Also we understand that HMMA image is enhancing contrast than original image and the color tone keeps and slightly brighter than original image.

Fig. 4 and Fig. 5 show another example of results and the corresponding histogram, respectively. In this case, we understand that the simple application of HMGD processing gives poor perspective impression. And HMMA image is enhancing its perspective and detail of rocks in this image than original.

From the experimental results, we consider that HMMA processing enhances the effect of the Kansei impression effect more than HMGD processing.
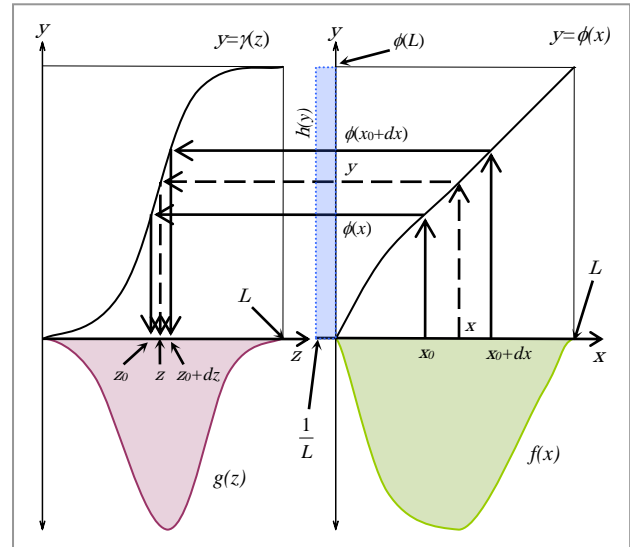
Fig.1 Conceptual image of histogram matching (example of HMGD).

4. Conclusion

In this paper, for the color image arrangement processing, we have described two histogram matching based methods: Histogram Matching based on Gaussian Distribution (HMGD) and Histogram Matching based on Moving Averaged histogram (HMMA).

And we have compared the two methods as to how each processing brings about the Kansei impression effect. Although the number of experimented images that we have shown in this paper is limited, from the results, we can find that the processing by HMMA more enhances the contrast, brightness, and detail of original image than by HMGD. That is, we consider that HMMA processing can improve the Kansei impression.

Since these evaluation of impression effects are subjective at this point, so we will objectively investigate the improvement of Kansei impression through questionnaire surveys and characteristics quantity in image, for future study.

a.   Original image

a.   Original image histogram

b.   HMGD image

b.   HMGD image histogram

c.   HMMA image
Fig. 2. Experimental results of original image, HMGD image, and HMMA image
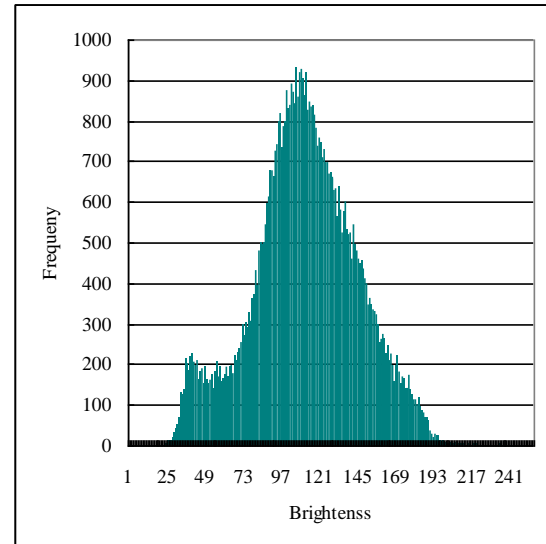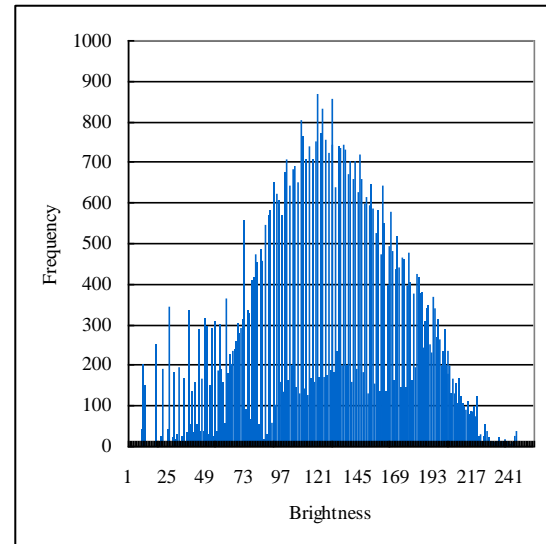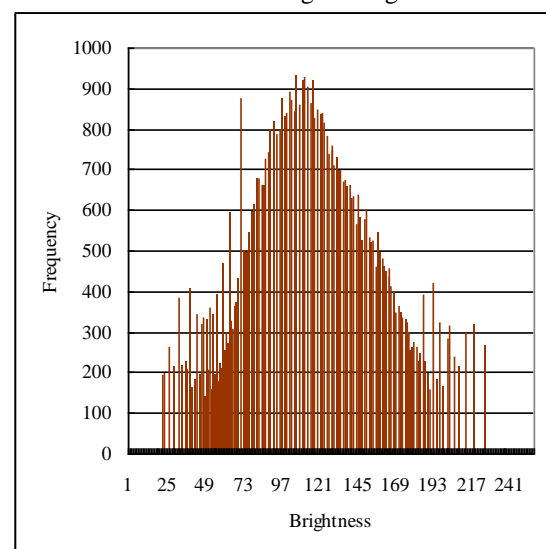
c.   HMMA image histogram
Fig. 3. Corresponding histogram of Fig. 2

a.   Original image

a.   Original image histogram

b.   HMGD image
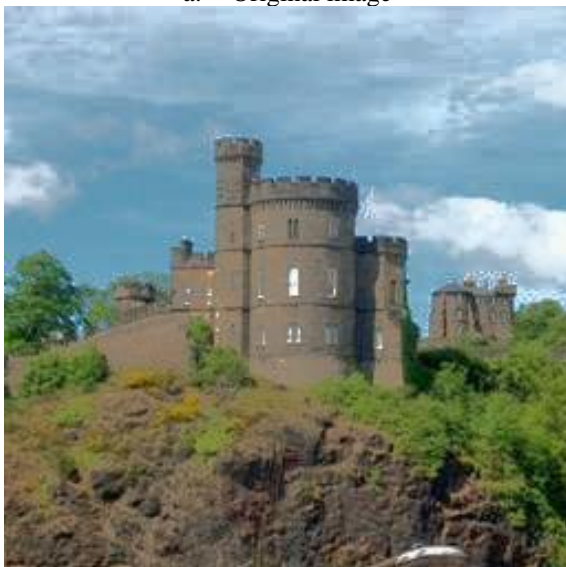
b.   HMGD image histogram
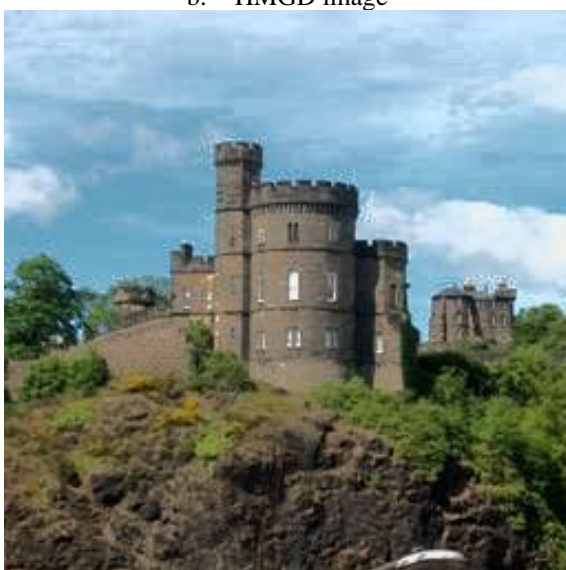
c.   HMMA image

Fig. 4. Experimental results of original image, HMGD image, and HMMA image
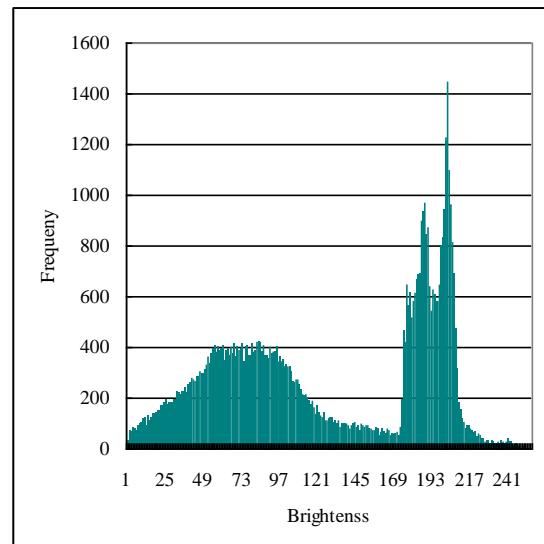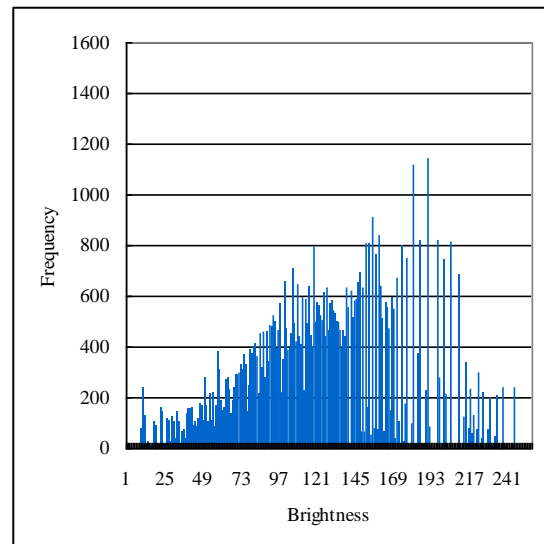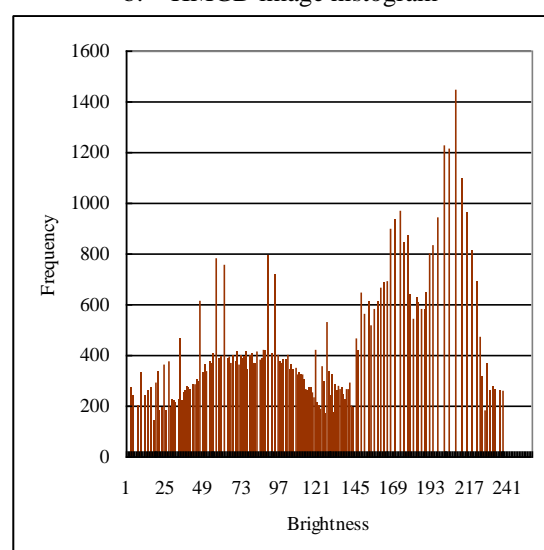
c.   HMMA image histogram

Fig. 5. Corresponding histogram of Fig. 4

References

1. R. C. Gonzalez and R. E. Woods, Digital Image Processing, Addison-Wesley Publishing Company, 1993.
2. B. Jahne, Digital Image Processing --Concepts, Algorithms, and Scientific Applications-- 4th edition, Springer, 1995.
3. E. S. Umbaugh, Computer Vision and Image Processing: A Practical Approach Using CVIP tools, Prentice Hall PTR, 1998.
4. W. Burger and J. M. Burge, Principles of Digital Image Processing: Fundamental Techniques, Springer, 2009.
5. T. Izumi, T. Hattori, S. Sugimono, and T. Takashima, Color Image Arrangement Using Elastic Transform on Principal Component Axis (in Japanese), Journal of Japan Society of Kansei Engineering Vol. 8(3) (2009), pp. 667-674.
6. Y. Kawakami, T. Hattori, D. Kutsuna, H. Matsushita, Y. Imai, H. Kawano, R.P.C. Janaka Rajapakse, Automated Color Image Arrangement Method Based on Histogram Matching - Investigation of Kansei impression between HE and HMGD -, International Journal of Affective Engineering Vol. 14(2) (2015), ISSN 2187-5413, pp. 85-93.
7. Y. Kawakami, T. Hattori, Y. Imai, H. Matsushita, H. Kawano, R. P. C. Janaka Rajapakse, Kansei Impression and Automated Color Image Arrangement Method, Journal of Robotics, Networking and Artificial Life Vol. 1(1) (2014), ISSN 2352-6386, pp. 60-67.
8. Y. Kawakami, T. Hattori, Y. Imai, K. Ando, Y. Horikawa, R. P. C. J. Rajapakse, Reference Histogram Matching Based on Moving Averaged Brightness Histogram, IEEJ Transactions on Electronics, Information and Systems Vol. 138(9) (2018), ISSN 1348-8155, pp. 1131-1132.
9. Y. Kawakami, T. Hatttori, R. P. C. J. Rajapakse, Kansei Impression of Color Image Arrangement Using Histogram Equalization, Proceedings of ICESS2023 (The 8th International Conference on Electronics and Software Science), 4pages.

Authors Introduction

Dr. Yusuke Kawakami

He received the B.E., M.E., and Ph.D. degrees in Information System Engineering from Kagawa University, Japan, in 2009, 2011, and 2014, respectively. Since 2020, he has been with National Institute of Technology (NIT) Kagawa College, Japan. Currently, he is an Assistant Lecturer in NIT Kagawa College. His research interests include Kansei Engineering, image processing and sound signal processing. He is a member of IEEJ.

Dr. Tetsuo Hattori

He received the B.E. and M.E. degrees in Electrical Engineering from Doshisha University, and the Ph.D. degree in Electronics & Information Engineering from Osaka University, Japan. After he worked for Toshiba Eng. Co., Ltd, he had been with Kagawa University from 1988 to 2015. Currently, he is a Professor Emeritus at Kagawa University. His research interests include STEM education based on ETT, signal and image analysis, pattern recognition, and realization of hybrid computer. He is a member of IEEJ and IEEE.

Dr. Janaka Rajapakse

He received the M.S. and Ph.D. degrees in Intelligence Science from Japan Advanced Institute of Science and Technology (JAIST), in 2005 and 2008, respectively. Currently, he is an Associate Professor at the Graduate Institute of Animation and Film Art, Tainan National University of the Arts, Taiwan. His research interests include Computer Graphics and Animations, Artificial Intelligence, Motion Capture Techniques, Virtual Reality, Augmented Reality, Haptic Interfaces, 3D Printing, Interactive Media and Kansei Engineering. He is also a member of IEEE.

Methodology for Creativity Oriented STEM Education Based on ETT Theory

Tetsuo Hattori, Toshihiro Hayashi, Mai Hattori, Yoshiro Imai
Kagawa University, 1-1 Saiwai-cho, Takamatsu, Kagawa 760-0016, Japan
Email: hattori@pe.kagawa-u.ac.jp

Asako Ohno
Osaka Sangyo University, 3-1-1, Nakagaito, Daito-shi, Osaka 574-8530, Japan

Takeshi Tanaka
Hiroshima Institute of Technology, Saeki-ku, Hiroshima 731-5143, Japan

Abstract

Recently, the necessity of integrated and comprehensive methodology for STEM (Science, Technology, Engineering and Mathematics) education is growing. In this paper, we propose an educational methodology utilizing the viewpoint of Equivalent Transformation Thinking (ETT) theory which has been proposed by Dr. Kikuya Ichikawa in 1955 as a principle of creativity. Especially, we show that the viewpoint is very useful not only for new technology invention but also for STEM Education in the sense that it deepens insights of the contents to be learned, motivates students to study further, and inspires their creativity.

Keywords: *STEM Education, Equivalent Transformation Thinking (ETT) Theory, Identification of Equivalence, 3-Dimensional Computer Graphics (3DCG).*

1. Introduction

Recently, the necessity of integrated and comprehensive methodology for Science, Technology, Engineering and Mathematics (STEM) education seems to be growing [1], [2], because the content of the STEM field is becoming increasingly sophisticated.

For the STEM education, we propose the methodology based on the viewpoint of Equivalent Transformation Thinking (ETT) theory which was advocated as a principle of creativity by Dr. Kikuya Ichikawa in 1955, after his investigation and analysis of the past discoveries and technological inventions.

In this paper, we briefly explain the ETT theory ([3], [4], [5], [6], [7], [8]) and describe that the viewpoint of ETT is very useful not only for new technology invention but also for STEM education. This is because the past inventions, discoveries and academic developments in STEM field can be explained according to the ETT theory.

Especially, we emphasize that the suitably combined use of dynamic visualization technology based on the viewpoint of ETT will be very effective for the students to deeply understand the learning contents, to promote their motivation, and to stimulate their creativity ([9], [10], [11]).

2. ETT Theory

As aforementioned, the ETT theory has been proposed as a principle of creativity by late Dr. Kikuya Ichikawa in 1955, after his investigation of the past discoveries and inventions to analyze how creation was done. And he revealed the investigation results that new technology developments and creations were produced along with the ETT process. Then he described the ETT process as a form of symbolic logical expression which is called "Equivalence Equation" as shown in Fig. 1.

Ao: original system, $B\tau$: transformed (or arrival) system, Σa: a set of special/peculiar properties and conditions that holds the system Ao, $c\varepsilon$: core essence or essential meaning that reasons the equivalence under some conditions, Σb: a set of necessary properties and conditions that hold the system, vi: viewpoint at a certain

Fig. 1 Equivalence Equation by Dr. Kikuya Ichikawa in his ETT Theory.

The description of "Equivalence Equation" means that, from a certain viewpoint vi, the original system Ao can be transformed to the system $B\tau$ according to the conditionalized essence $c\varepsilon$, where some properties and

conditions Σa in $\boldsymbol{Ao}$ are discarded, and new elements of property and condition Σb are added to the system $\boldsymbol{B\tau}$.

When Dr. Ichikawa presented "Equivalence Equation", he also showed the schematic flowchart of ETT. And he insisted that, if we make effort to develop something to be desired according to the ETT flowchart, creative achievement will be possible more efficiently.

As for the ETT theory, he mainly published papers at the Creativity Research Workshop that was held by Dr. Hideki Yukawa, the winner of 1949 Nobel Prize in Physics. Dr. Yukawa called this Dr. Ichikawa's theory as "Identification Theory" in his own words and highly praised it.

3. Application Examples of ETT Viewpoint
3.1. Cartesian Coordinate System

Looking back on history, it is possible to find out many facts that the act of ETT and/or Identification of Equivalence has made a progress of research in the field of STEAM (Science, Technology, Engineering, Art, and Mathematics), and that it has broadened our horizons and deepened our knowledge. Then, we consider that utilization of the viewpoint of ETT is also effective for STEM education.

For example, René Descartes provided "Cartesian coordinate system" that gives a correspondence between a point P on a plane and a pair of two numbers (x, y), where x and y are quantities based on the distances to two orthogonal axes (x-axis and y-axis), respectively, as shown in Fig. 2.
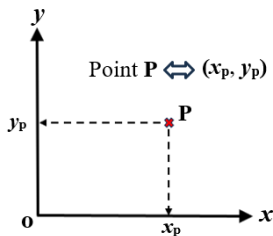
Fig. 2 Cartesian coordinate system.

As a result, taking a parabolic curve C for an example, it can be simply represented by a set of coordinates and algebra, without showing the procedure of geometric drawing, such as

C= {P(x, y) | y =ax^2, a: constant number}.

This is an example that ETT not only opens the new field of mathematics such as Analytic Geometry and Algebraic Geometry, but also brings about the new way of representation such as "Field" into physics, so we consider that it is worth to enhance the idea of ETT. Moreover, we should notice the viewpoint of ETT and appropriately point out the viewpoint of equivalence when teaching the subject of STEM.

3.2. Archimedes' Principle of Buoyancy
According to Archimedes' principle, floating power (or buoyant force) that operates on a body immersed in a fluid, is equal to the weight of the fluid in the same volume as the volume of body in the fluid (Fig. 3).

As shown in Fig. 3, if we imagine a situation that an object is quietly floating on the water, and if we think of that it is equivalent to the situation where the water in the same volume area of the object under the surface is quietly balanced (Fig. 3(c)), we understand that Archimedes' principle will be derived.
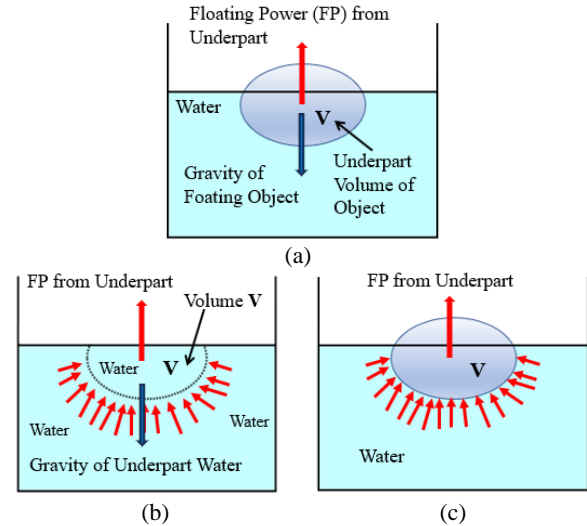
Fig. 3. Floating power (or buoyant force) that operates on a body immersed in a fluid. (a) An object is quietly floating on the water. (b) V means the same volume/area of the object under surface. (c) The water in the same volume area V is quietly balanced.

3.3. Mechanical Dynamic System (MDS) and Electric Circuit System (ECS)
For taking the relation between MDS and ECS as an example, we can deal with the equations of two systems (see Fig. 4) as equivalent equation from both meaning of mathematics and physics.

$$\begin{cases} m\dfrac{d^2x}{dt^2} + r\dfrac{dx}{dt} + \dfrac{x}{k} = \boldsymbol{F}_0 \quad \left(v = \dfrac{dx}{dt} \right) \\ L\dfrac{d^2q}{dt^2} + R\dfrac{dq}{dt} + \dfrac{q}{C} = \boldsymbol{E}_0 \quad \left(I = \dfrac{dq}{dt} \right) \end{cases}$$

Mass $\boldsymbol{m}$ $\Leftrightarrow$Inductance $\boldsymbol{L}$, Responsiveness $\boldsymbol{r}$ $\Leftrightarrow$Resistance $\boldsymbol{R}$, Compliance $\boldsymbol{k}$ (the inverse of spring constant) $\Leftrightarrow$Capacitance $\boldsymbol{C}$
(a)
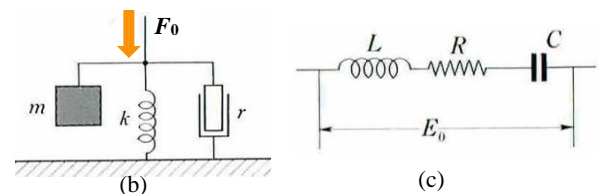
Fig. 4. (a) Differential equations and their corresponding elements in MDS and ECS. (b) Conceptual image of MDS. (c) Conceptual image of ECS.

3.4. Explanation of Faraday's World First Motor

Faraday's world first electromagnetic rotating device is shown in Fig. 5, where the current flows from top to bottom into the left and right mercury containers. Viewing from the above, when the current flows, both of bar magnet and electrode (or conductor) rotate clockwise.

(a)   (b)

(c)
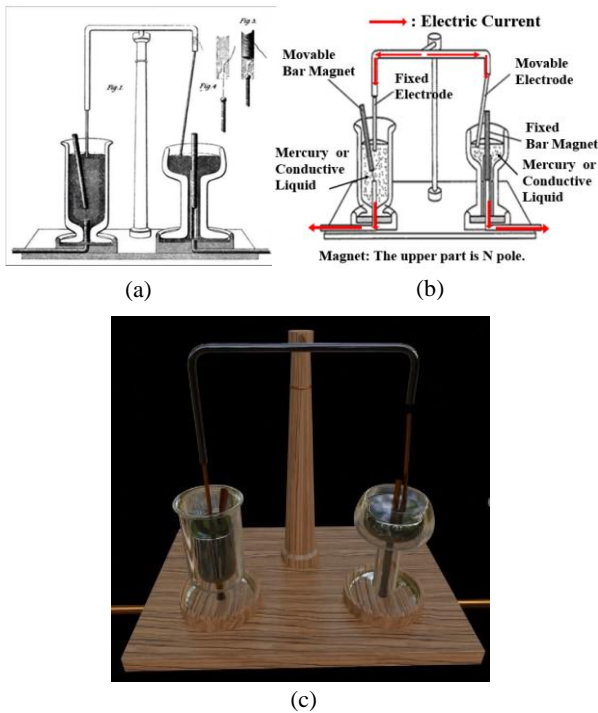
Fig. 5. Faraday's World's First Motor. (a) Documents left at the Royal Institution. (b) Explanation of each part and direction of current. (c) Motion view of the Motor made by 3-DCG "Blender".

As for the force that causes rotation, there are different explanations (or lectures) for the left and right sides. For the left side, the current flowing through conductor rotates the movable bar magnet by creating magnetic according to Ampere's law. On the contrary, for the right side, the the current in conductor rotates because it receives Lorentz force from the magnetic field made by the fixed bar magnet.

However, there is a unified explanation from law of action and reaction. In the left side, like the left side, the magnetic field caused by the current in conductor exerts a force on the fixed bar magnet, but since the bar magnet is fixed and the conductor is floating, then the the conductor rotates itself by the reaction force.

This unified explanation can be easily provided by 3-DCG video, referring that an electric current generates a magnetic field around it and receives a reaction force from the bar magnet via the magnetic field.

3.5. Recursive Programming

Recursive programming is useful, but it is not so easy to suitably write down. Taking the problem of "Tower of Hanoi" for example (Fig. 6), the problem is to move the tower consisting of n disks, from the place at pillar A (or start place) at the left end to the pillar C (or destination) at the right end, using the pillar B in the middle (as work place), under the constraints as below.
(i) Only one disk must be moved at a time, (ii) Do not go outside the three sections, (iii) Do not put a disk over a smaller disk.

(a)   (b)

Fig. 6. The "Tower of Hanoi" problem. (a) Initial state when the number of disks n=5. From the left end, we call three places as A, B, and C. (b) The smallest disk at A is moved from A to C, and the next small disk is on the way to move out.

For making the recursive program in the general case of n disks, we note a pattern of disk moving, that will appear clearly when the number of disks n=2. That is, the pattern is equivalent to what can be seen if we divide n disks into two parts: upper parts (n-1 disks) and lowest part (the bottom largest disk).

Imagine when the largest disk can be moved from A (start) to C (destination) under the constraints. Firstly, we must move the upper parts (n-1 disks) from the place A to B (workplace), after that, the largest disk (=lowest part) can be moved from A to C. As the next step, the upper parts (n-1 disks) are to be moved from B to C. This is the important turning point (Fig. 7).

(a)

(b)

(c)
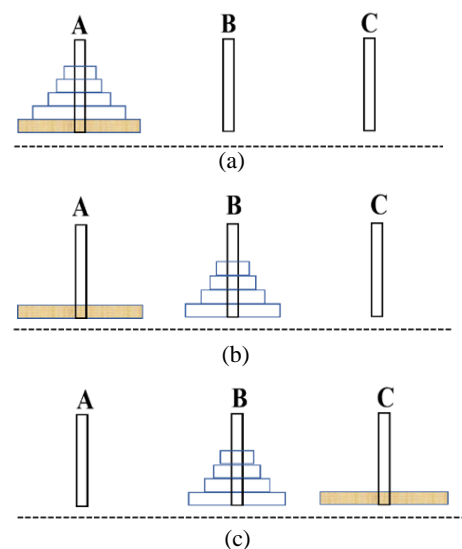
Fig.7. The turning point procedure of "Tower of Hanoi" problem in moving n disks from the place A to C using the place B. (a)Initial state. (b) and (c). The turning point when the top n-1 disks are moved from A to B, after that, largest disk left at A is moved to C.

Then, if we let TOH(n, S, W, D) represent a procedure of the "Tower of Hanoi" problem to move n disks from the start place S to destination D using workplace W, we obtain the following symbolic formulation as a recursive procedure.

TOH(n, A, B, C) =TOH(n-1, A, C, B) + Move the largest of n disks from A to C+ TOH(n-1, B, A, C).

The procedure TOH(n, S, W, D) is concretely, described by Excel VBA, as shown in the Fig. 8.

```
Sub main_Tower_of_Hanoi()
Dim n As Integer, Input_n As String
Input_n = InputBox("The number of disks ?")
n = Val(Input_n)
Call Hanoi_Tower(n, "A", "B", "C")
End Sub
'********************************
```
(a)

```
Sub Hanoi_Tower(n As Integer, S As String, W As
String, D As String)
  '/*Move n-1 disks from "S" to "W" via "D."*/
  If n >= 2 Then
       Call Hanoi_Tower(n - 1, S, D, W)
  End If
  Debug.Print "Move the top disk No. " & n & " : " & S
& "→" & D & "."
  '/*Move n-1 disks from "W" to "D" via "S."*/
  If n >= 2 Then
       Call Hanoi_Tower(n - 1, W, S, D)
  End If
End Sub
'**************************************
```
(b)

(c)
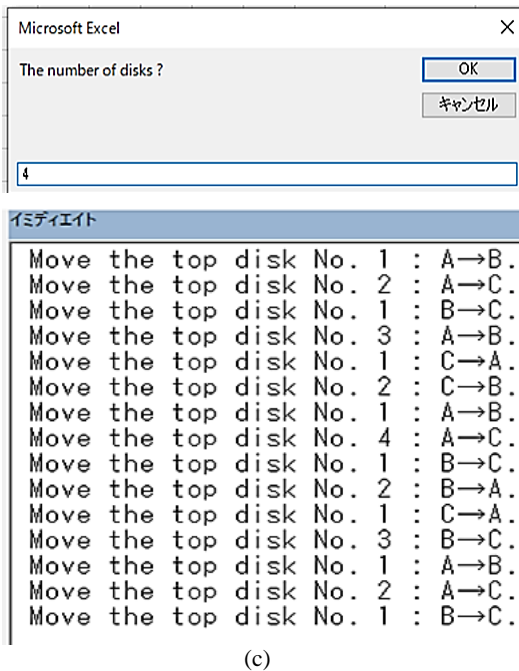
Fig. 8. Excel VBA program for the Tower of Hanoi. (a) and (b). The program list. (c) (upper) Inputbox where n=4 is input. (lower) The output at immediate window in the case n=4.

The necessary number of steps $a(n)$ to solve the problem of n disks is decided from the above algorithm.

$a(\text{n})=a(n-1) +1+a(n-1) =2a(n-1) +1$ $(n \geqq 2)$, $(a (1) =1)$.

Let $b(n)=a(n)+1$, then $b(1) =2$ and $b(n)=2b(n-1)$. Then,

$$b(n) = \frac{b(n)}{b(n-1)} \cdot \frac{b(n-1)}{b(n-2)} \cdots \frac{b(3)}{b(2)} \cdot \frac{b(2)}{b(1)} \cdot b(1) = 2^n$$

Therefore, $a(n)=b(n)-1=2^n-1$.

4. Conclusion

In this paper, we have proposed a methodology for integrated and comprehensive STEM education, based on the viewpoint of ETT. Moreover, we have presented some application examples of teaching way from the ETT viewpoint, in order to show the usefulness. We consider that the suitably combined use of visualization tool such as 3-DCG will be more effective.

As a further study, we are considering that the solution algorithm for the Tower of Hanoi problem can be represented as the tree generation in tree language notation. Then, we find out that the recursive algorithm may be described as an iterative algorithm equivalent to the recursive one, and that a fast solution algorithm can be obtained. We would like to report this on another opportunity.

References

1. Yeping Li, Ke Wang, Yu Xiao & Jeffrey E. Froyd, Research and trends in STEM education: a systematic review of journal publications, International Journal of STEM Education, Springer Open, Vol.7, No.11, https://doi.org/10.1186/s40594-020-00207-6G (2020)
2. Fitzgerald, M. S., & Palincsar, A. S., Teaching practices that support student sensemaking across grades and disciplines: A conceptual review, Review of Research in Education, Vol.43, pp.227–248(2019)
3. Kikuya Ichikawa, Science of Creativity, NHK Publishing (1970) (in Japanese)
4. Kikuya Ichikawa, Creative Engineering, Lattice Co., Ltd. (1977) (in Japanese)
5. Hideki Yukawa, Leap to Creation, Kodansha (1968) (in Japanese)
6. Equivalent Transformation Creativity Society ed., Equivalent transformation theory that can be understood graphically, NIKKAN KOGYO SHINBUNSHA (2005) (in Japanese)
7. Hikaru Matsuki, Hideaki Ogawa, Yasuaki Kajisha, A New Movement on the Theory of Equivalent Transformal Thinking, The Japanese Society for Artificial Intelligence, Vol.22, No.3, pp.392-398 (2007) (in Japanese).
8. Takeo HARUYAM, A Study on the Technology and Skill Transfer from the Creative Engineering Point of View, Journal of Japan Society for Production Management, Vol.15, No.1, pp.1-6 (2008) (in Japanese)
9. Tetsuo Hattori, Rikiya Masuda, Yoshio Moritoh, Yoshiro Imai, Yusuke Kawakami, Takeshi Tanaka, Utilization of Both Free 3DCG Software "Blender" and 3D Printing for Early STEM Education, Proc. of IEEE TALE2020 (An International Conference on Engineering, Technology and Education), ISBN 978-1-7281-6942-2, pp.971-974 (2020)
10. Tetsuo Hattori, Rikiya Masuda, Toshihiro Hayashi, Yoshiro Imai, Yoshio Moritoh, Yusuke Kawakami, Takeshi Tanaka, A STEM Education Method Utilizing 3DCG Software Blender, Proc. of the Sixth International

Conference on Electronics and Software Science (ICESS2020), ISBN978-1-941968-62-8, The Society of Digital Information and Wireless Communications (SDIWC), pp.38-43 (2020)

11. Tetsuo Hattori, Asako Ohno, Yoshiro Imai, Yusuke Kawakami, Katsunori Shimohara, Takeshi Tanaka: "Creativity Oriented STEM Education Utilizing Equivalent Transformation Thinking Theory and Dynamic Visualization Technology", Proc. on 2022 Annual Conference of Electronics, Information and Systems Society, IEE of Japan, GS3-6, pp.1058-1062 (2022) (in Japanese)

Authors Introduction

Dr. Tetsuo Hattori

He received the B.E. and M.E. degrees in Electrical Engineering from Doshisha University, Japan, and the Ph.D. degree in Electronics & Information Engineering from Osaka University, Japan. After he worked for Toshiba Eng. Co. Ltd., he had been with Kagawa University from 1988 to 2015. Currently, he is a Professor Emeritus there. Member of IEEJ and IEEE.

Dr. Toshihiro Hayashi

He received the B.E., M.E., and Ph.D. degrees in Information and System Engineering from Tokushima University, Japan, in 1989, 1991, and 1994, respectively. Since July 2004, he has been with Kagawa University, after he worked at Saga University from April 1994 to June 2004. Currently, he is a Professor at Kagawa University as the director of Integrated Center for Informatics there.

Dr. Mai Hattori

She received the Ph.D. degree in Medicine from Kagawa University, Japan, in 2023. She is currently engaged in the study on Education of Life Science and Nutrition utilizing Information Technology as a researcher of Integrated Center for Informatics at Kagawa University.

Dr. Yoshiro Imai

He received the B.E. degree in Information Engineering from Kyoto University, Japan, in 1980, and Ph.D. from Tokyo University of Agriculture and Technology, Japan, in 2008. He had been a professor at Kagawa University until the end of March, 2021. Now he is a Professor Emeritus there. His research interests include Computer Architecture and e-Learning system Design, etc.

Dr. Asako Ohno

She received the B.S. from Osaka University of Foreign Studies in 2001. She also received the M.S. and Ph.D. degrees in Human Science from Kobe University, in 2006 and 2009, respectively. Currently, she is an Associate Professor at Osaka Sangyo University. Her research interests include Intelligent Learning System, Learner Behavior Analysis, Elementary Programming Education, etc.

Dr. Takeshi Tanaka

He received the B.E., M.E., and Ph.D. degrees in Electrical and Material Engineering from Hiroshima University, Japan, in 1982, 1984, and 1990, respectively. Currently, he is a Professor at the Department of Electronics and Computer Engineering in Hiroshima Institute of Technology. His research interests include Integrated Circuit Manufacturing, Applications, and Education on Semiconductor, etc.

Efficient Campus Shuttle Tracking and Management Mobile Application for College Campus

Andrea Tantay Gonzales
INTI International College Penang, Malaysia

Kavitha Thamadharan
INTI International College Penang, Malaysia

Neesha Jothi
UCSI University, Malaysia

Abstract

The purpose of this study is to address the current paper-based shuttle booking and tracking system at INTI International College, Penang. Shuttle Stalk, a Real-Time Campus Shuttle Booking and Tracking System was a proposed solution that enables the students to register for the shuttle service through the mobile application, reducing the administrative burden and tracking its current location in real-time, allowing them to plan their journeys more effectively. The application developed serves as a comprehensive solution for both students and administrators, providing real-time shuttle tracking and streamlined management capabilities. The development of this application is informed by extensive data collected through pre- and post-acceptance surveys, to ensure its effectiveness and alignment with user needs. The mobile application not only enhances the efficiency of shuttle services but also contributes to higher user satisfaction, improving the overall campus experience.

Keywords: Campus shuttle-tracking application, real-time shuttle location, Global Positioning System (GPS), Google Maps, Firebase, Mobile Application

1. Introduction

Shuttle services are a form of public transportation that respective schools and universities often provide to enable their students to travel from the pick-up point to the institution's location. However, there is no way to track and see the shuttle's arrival and departure time. According to a study [3], the drawback of public transportation (such as buses and shuttles) is that the commuter has an unknown amount of time to wait for their transport's arrival. This results in commuters opting for private vehicles, which adds consequences in environmental, social, and economic terms [7].

A study [7] shows that (from a survey sent to 3378 correspondence), 59% of the students opt for public transportation such as buses and shuttles. Distance, time, and economic capabilities are the most important factors when choosing a mode of transportation. With that in mind, the free transportation offered by some schools and universities (e.g., INTI International College and Universities) appeals mostly to students, as it meets the criteria of choosing a transportation mode. However, the lack of information regarding the transportation's arrival time, current location, and seat availability makes commuters or students reluctant to use public transportation [8].

Therefore, a shuttle booking and tracking system was proposed as a solution for the current paper-based shuttle management at INTI International College, Penang. It enables the students to book the shuttle service through the mobile application, where they can decide whether the registration may proceed or not depending on the seating availability in the shuttle. Moreover, the students will be able to track the shuttle's real-time location as well as efficiently mark their attendance when the shuttle is nearby. Lastly, the students will be able to receive notifications from both admins and drivers if the shuttle is unavailable.

Andrea Tantay Gonzales, Kavitha Thamadharan, Neesha Jothi

2. Literature Review

The current shuttle management system in INTI International College, Penang enables the students to use the shuttle through Google Form registration every new semester. Attendance marking and tracking were paper-based; the student's name and matriculation number must be written manually by the student after riding the shuttle. There was also no administrator dashboard for the overall management of the shuttle service on campus.

Fig. 1 shows the pre-acceptance survey which was sent out to 100 students with 63 respondents. Most of the students voted on having a real-time shuttle tracking application to help plan transportation ahead of time.

Would a real-time shuttle tracking application be helpful for you?
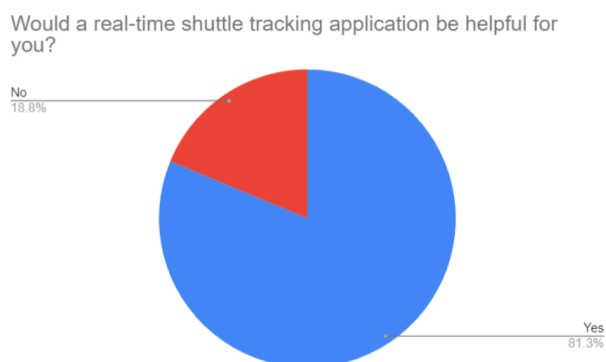
Fig. 1. Pre-acceptance survey result

A real-time location tracking application is a type of software application where the location data of any two or more devices are updated in each period to the database so that both devices' location data information synchronizes with each other. There has been much research about the implementation of real-time location tracking applications in a school and campus' public transportation system, each with different implementation methods and techniques.

Saad et al. [8] conducted research on a real-time on-campus public transportation monitoring system for Universiti Teknologi Malaysia. The researchers implemented a GPS receiver hardware and attached it to the campus bus, where it retrieves the latitude, longitude, and speed of the bus. The data is then synchronized to the student web application. This system is not as portable and accessible, however, since students mostly use mobile applications nowadays. The system must also adapt to the current design and UI trends. Lastly, each of the campus transportation vehicles must have GPS receiver hardware, which may add to the cost.

Hoque, et al. [1] developed a bus tracking system for the International Islamic University Chittagong in Bangladesh. Instead of using a physical GPS receiver, the researchers made use of an Android smartphone's embedded location-aware hardware, making the system cost and configuration time lesser. The system is mobile-based, therefore, improving student accessibility and access. There were also two added features in the system, namely the driver accident reporting and seat availability checking. These features are beneficial to students, however, the way that the seat availability checking was implemented may be hazardous (as the driver has to continuously accept/deny students' requests, even while driving). The UI for the application may also need a lot of time getting used to.

Shibghatullah, et al. [11] implemented Tracker, a real-time location tracking system that makes use of GPS-aware devices to track the current location of public transportation. It caters to a wider audience and is mobile-based. It has a much-improved UI with detailed information (e.g., real-time traffic conditions, error messages) which makes it easier for users to navigate through the application, as well as a notification feature, allowing the users to receive updates from the administrators. In terms of improvement, the application does not have a way to check seat availability, and the notification feature can be further improved by adding a push notification alert.

3. Methodology

The research methodology used for this study was the Waterfall Methodology (as shown in Fig. 2). During requirements gathering, feedback regarding the current shuttle service on the campus was compiled from administrators and students alike. The design phase implements a non-working prototype of the application with the main features. Once the idea was confirmed, the working prototype was implemented based on the pre-acceptance survey result and the initial design of the application. Testing was done afterward, where the working application was sent to the end users for the purpose of gathering UI and feature-related issues and feedback. Lastly, the application was updated based on the result of the usability testing.
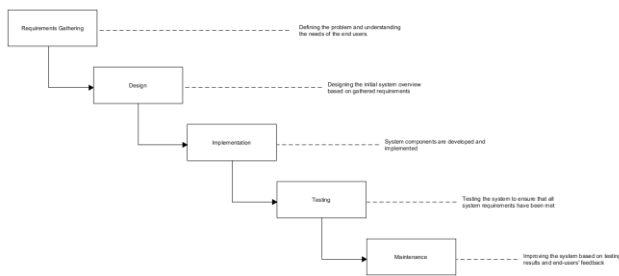
Fig. 2. Research Methodology

Fig. 3 shows the administrator website, where the admins can manage the campus' shuttle services. It contains various managerial features like shuttle and driver registration, where the admin can input the shuttle details (e.g., plate number, seating capacity) as well as create an account for the driver (to be used on the driver's side of the application). Route management is also included on the website, where the admin can assign a registered shuttle and driver to the route and enter its pickup/dropoff schedule and location. The admins can also view the number of bookings made by students, as well as the registered students.

Fig. 3. Administrator Website

A student application was also implemented as shown in Fig. 4. It allows the students to book a shuttle based on the pickup/dropoff location, as well as view the shuttle's current real-time location. The application has a built-in attendance feature where the student can mark their own attendance if the shuttle is nearby. Unmarked attendance may result in the "no show" penalty embedded in the application that denies the student from booking the shuttle. Moreover, the student may book the shuttle as long as there is available seating in the vehicle. The student can also receive push notifications from administrators and drivers alike whenever there are updates regarding the shuttle service on the campus (e.g., shuttle maintenance). Lastly, an option to cancel the shuttle booking was also in place if the journey time was not within 2 hours.
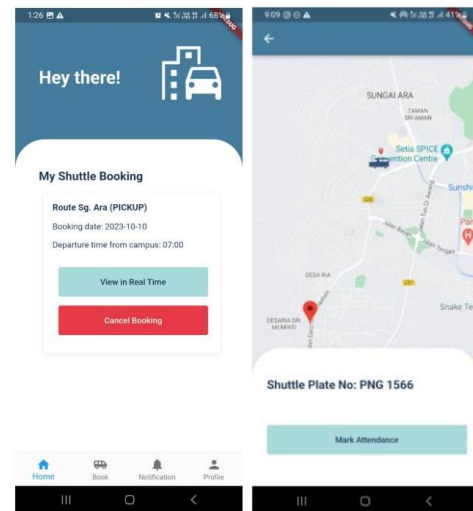
Fig. 4. Student Application

Lastly, a driver application was implemented to foster communication between the student and the driver's location (refer to Fig. 5). A listing of all available and assigned journeys to the driver can be seen on the home page. When the driver starts a journey, the application saves the device's current location and updates the database with it whenever the shuttle or driver moves 100 meters from the last location. The driver application also has an emergency reporting feature that triggers whenever the driver cancels a journey or ends a journey early. The submitted report sends a push notification to the student's application side and cancels the journey without penalizing the student.

Fig. 5. Driver Application

Andrea Tantay Gonzales, Kavitha Thamadharan, Neesha Jothi
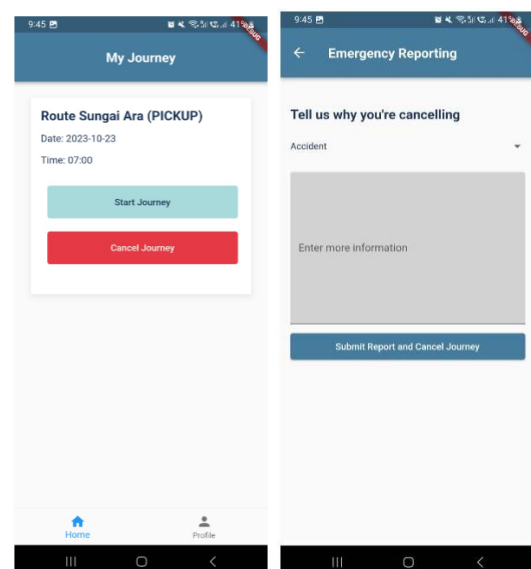
4. Results and Discussion

For this study, the application underwent usability testing. Existing research [4], [5] shows that the amount of usability testers should be 5 testers (given that the study does not focus on quantitative results, in which case, 20 testers are required). A total of 8 participants, however, were compiled for this study as it ensures that almost all usability problems were covered as much as possible.

Table 1 below shows the result and overall feedback of the participants from the usability testing.

Table 1. Usability testing results

Participant no.	Feedback
1	- Some parts of the text instructions when logging in and registering are obstructed. - Overall, the application has nice features and is smooth and fast.
2	- Application is easy to use and straightforward with minimalistic design. - There should be an option for weekly booking of the shuttle.
3	- There should be a polyline location to track the shuttle route. - Improve the arrangement of the input boxes on the booking page.
4	- The application is easy to use as the layout is pretty similar to other applications. - There was a missing toast message for cancellation of shuttle booking.
5	- The application is easy to navigate around - More optimization in the future updates
6	- Overall application experience is smooth, and the interface is nice. - There should be more customization options for personalized application experience.
7	- There should be a password confirmation during registration. - Overall, the application is straightforward and easy to use.
8	- The application is not overwhelming to use. - It takes some time to load (optimization).

It can be seen from the table that the overall impression of the system was smooth, easy to navigate, and straightforward, with most of the participants liking its

minimalistic design and similarity in layout with other applications (thus contributing to its ease of use). Further improvements of the system focused more on optimization, minor element obstruction, and other requested features by the users (e.g., personalization, polyline for Google Maps).

5. Conclusion

In this paper, an efficient shuttle tracking and management application has been developed for INTI International College, Penang campus. The proposed system proved beneficial to both the campus administrators and students, especially when it comes to managing shuttle bookings and tracking the shuttle location, compared to its initial paper-based and manual shuttle management system. Furthermore, the application was a success in terms of user acceptance as proven by the result of the usability testing.

References

1. Hoque, F., Chakma, R., Mahtab, S. S., Akter, R., & Ahmed, S. S. (2020). Design and Developing Real Time Interactive IIUC Bus Tracking System. Journal of Innovation in Computer Science and Engineering, 9(2).
2. Nama, G. F., Rasyidy, F. H., Arum, R., & Mardiana. (2018). A Real-time Schoolchild Shuttle Vehicle Tracking System Base on Android Mobile-apps. International Journal of Engineering & Technology.
3. Narkhede, P. V., Mahalle, R. V., Lokhande, P. A., Mundane, R. M., & Londe, D. M. (2018). Bus Tracking Sytem Based on Location Aware Services. International Journal of Emerging Technologies in Engineering Research (IJETER), 6(3).
4. Nielsen, J. (2012). How Many Test Users in a Usability Study? Retrieved 27 October, 2023, from Nielsen Norman Group: https://www.nngroup.com/articles/how-many-test-users/
5. Nielsen, J., & Molich, R. (15 November, 2020). 10 Usability Heuristics for User Interface Design. Retrieved 15 May, 2023, from Nielsen Norman Group: https://www.nngroup.com/articles/ten-usability-heuristics/
6. Ramya, S., Ajitha, P., Brindha, A., Gayathri, A., & Kavitha, V. (2021). QR Scan Based System for School Bus Tracking. International Journal of Scientific Research & Engineering Trends.
7. Romanowska, A., Okraszewska, R., & Jamroz, K. (2019). A Study of Transport Behaviour of Academic Communities. Sustainability. 325
8. Saad, A. S., Hisham, A. A., Ishak, M. I., Fauzi, M. M., Baharudin, M. A., & Idris, N. H. (2018). Real-time on-Campus Public Transportation Monitoring System. 2018 IEEE 14th

International Colloquium on Signal Processing & its Applications (CSPA 2018).

9. Saif, M. A., Zefreh, M. M., & Torok, A. (2019). Public Transport Accessibility: A Literature Review. Periodica Polytechnica Transportation Engineering.

10. Sham, R. B., Hussin, A., Abdamia, N., & Mohd, M. (2019). Developing a Pollution Free Environment Framework through Technology Integration (e-Hailing App). 7th AMER International Conference on Quality of Life.

11. Shibghatullah, A. S., Jalil, A., Wahab, M. H., Soon, J. N., Subaramaniam, K., & Eldabi, T. (2021). Vehicle Tracking Application Based on Real Time Traffic. International Journal of Electrical and Electronic Engineering & Telecommunications.

Authors Introduction

Andrea Tantay Gonzales

She received her Bachelor's in Computer Science (in collaboration with Coventry University) from the School of Computing, INTI International College Penang, Malaysia in 2023. She recently finished her internship in Sophic Automation SDN BHD, Malaysia.

Kavitha Thamadharan

She received her Master's from the School of Computer Sciences, Universiti Teknologi Malaysia in 2015. She is currently a Lecturer in INTI International College Penang, Malaysia. Her research interest areas are Information Security and Information Systems.

Dr. Neesha Jothi

She received her PhD from the School of Computer Sciences, Universiti Sains Malaysia in 2020. She is currently an Assistant Professor in UCSI University, Malaysia. Her research interest areas are Data Mining in Healthcare and Health Informatics.

GCN Analysis of Task-Based fMRI Data for Diagnosis of Schizophrenia

Tejaswini Thota, Reuben Stephen John
School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India

R Dhanush
School of Electronics Engineering, Vellore Institute of Technology, Chennai, India

Amutha S
School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India

Heshalini Rajagopal
Department of Electrical and Electronics Engineering, MILA University, 71800 Nilai, Negeri Sembilan, Malaysia

Abstract

This study focuses on schizophrenia, impacting 24 million people globally, characterized by distorted reality and delusions. We propose utilizing a GraphSAGE model and a Graph Convolutional Network (GCN) model on a task-based fMRI data to differentiate schizophrenia instances, with and without auditory-visual hallucinations, utilizing healthy participants for comparison. The study focuses on identifying specific auditory stimuli that can help significantly differentiate individuals experiencing hallucinations. The GraphSAGE model performed better by achieving an accuracy of 97% and 96% on the specified stimuli.

Keywords: Schizophrenia, Hallucination, Graph Convolution, GraphSAGE, fMRI

1. Introduction

Functional Magnetic Resonance Imaging(fMRI) has emerged as a powerful tool in unraveling the intricacies of mental disorders. Among the various types of fMRI, the task-based fMRI serves as an ideal method for understanding the functioning of neurons as it captures real-time brain activity as individuals perform specific cognitive tasks. In the context of schizophrenia, the analysis of task-based fMRI data has proven to provide valuable insights that in turn aid the diagnosis and treatment of this disorder.

Graph Convoluted Network (GCN) is a type of neural network architecture designed to process data structured as graphs. In these graphs, the nodes represent entities, and the edges represent the relationship between the entities. **GCN′s** leverage information from the neighboring nodes to perform traditional convolution operations. This methodology is suitable for analyzing fMRI data as they model dynamic changes in brain connectivity. In our investigation, we designate regions of interest (ROIs) as nodes, and the edges represent the functional connectivity between these brain regions. We leverage pertinent node attributes to encapsulate the variability in brain activity corresponding to specific stimuli, serving as distinctive features for our classifier. This methodology enables us to unveil patterns and relationships in brain connectivity, offering enhanced insights into the complex dynamics of schizophrenia and the responses of individuals affected by it to auditory stimuli.

2. Literature Review

This section provides concise information about prior literature which has been consulted as reference during the model preparation. The application of graph based neural networks in the analysis of neural activity is an emerging area of study. Saeidi *et al* [1] have proposed an end-end GCN network on the benchmark HCP dataset, which makes use of statistical features extracted from the connectivity matrices as a node feature. Additionally, standard node embedding algorithms such as NetMF, RandNE, Node2Vec and Walklets were used to extract node features and their performance was analyzed. The paper acknowledges the importance of individual differences in brain function by testing the model on a gender-based and gF score-based sub-dataset. The effect of batch size was also accounted for. The proposed framework achieved an accuracy of 97.7% and concluded that NetMF and RandNE algorithms outperformed the rest.

Li, X. *et al* [2] have employed a custom framework BrainGNN, composed of Ra-GNN and R-pool blocks which are responsible for generating the node embeddings. The node embeddings are aggregated with its neighbors and the updated representations are projected onto a learnable vector which is fed to a classifier. The model was utilized on two distinct datasets: the Biopoint Autism Study Dataset (Biopoint) and the Human Connectome Project (HCP) 900 Subject Release [3]. The primary objective for the Biopoint dataset [4] was the classification of autism affected individuals and healthy controls. Conversely, for the HCP dataset, the focus was on classifying seven

Tejaswini Thota, Reuben Stephen John, R Dhanush, Amutha S, Heshalini Rajagopal

cognitive and behavioral task states. BrainGNN demonstrates an enhancement in average accuracy ranging from 3% to 20% for autism classification. Additionally, it achieves an average accuracy of 93.4% for the classification on Human Connectome Project dataset.

3. Methodology

3.1. *fMRI dataset and preprocessing*

We obtained task fMRI data for 71 participants, 46 schizophrenic patients and 25 healthy controls subjected to three experimental conditions, word lists, sentence lists and reversed speech. White noise was the low-level control condition to which the task results were compared to. The schizophrenic patients are further classified into two classes, 23 subjects who experience auditory visual hallucinations (AVH+) and 23 subjects who do not (AVH-) (Table 1). The data was obtained using a 3T Philips Ingenia scanner employing a T2*-weighted echo-planar imaging (EPI) sequence. The acquisition consisted of 341 volumes, and the following parameters were applied: a repetition time (TR) of 2000ms, echo time (TE) of 30ms and a flip angle of 70°. The in-plane resolution was set to 3.5×3.5 mm, with a field of view (FOV) measuring 238×245 mm. The slice thickness was 3.5 mm, and an inter-slice gap of 0.75 mm was maintained. Slices, totaling 32 per volume, were acquired in an interleaved order parallel to the AC-PC plane [5]. Refer (Fig. 1.a) for raw data.

Table 1: Participants description

	HC N=25	SCHZ N=46	AVH+ N=23	AVH- N=23	Difference
Age	39.8 (14.08)	42.52 (10.72)	40.09 (12.97)	44.96 (7.38)	HC vs SCZ: t = -0.84, p = 0.404 AVH+ vs AVH-: t = -1.57, p = 0.127
Sex (M: F)	18:7	36:10	20:3	16:7	HC vs SCZ: χ2 = 0.35, p = 0.555 AVH+ vs AVH-: χ2 = 2.04, p = 0.284
Current IQ (WAIS III)	108.00 (18.71)	95.44 (12.96)	92.09 (13.76)	99.32 (11.08)	HC vs SCZ: t = 2.76, p = 0.010 AVH+ vs AVH-: t = -1.83, p = 0.072
Duration of illness (years)	-	17.9 (10.40)	14.76 (10.57)	21.37 (9.26)	t = -2.11, p = 0.042
PSYRATS – Hallucination subscale	-	-	24.91 (7.32)	-	-

The preprocessing of the data was carried out using the FEAT module of FSL(FMRIB Software Library) [6]. The

first 5 volumes were discarded to account for signal stabilization. Brain extraction was done using the BET tool. The functional data was then subjected to motion correction using the MCFLIRT algorithm and spatial smoothening (FWHM=5mm) was also applied (Fig 1.b) . FLIRT was used to register the low-resolution functional images to a high-resolution scan, and registration of the high-resolution scan to a standard MNI152) image. Using the transformation matrix obtained from the previous steps, we registered the functional sequence to the standard MNI space (Fig 1.c).
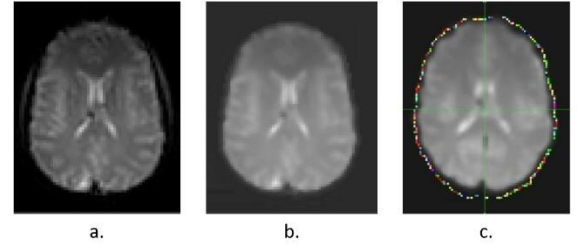
Fig 1 (a) depicts raw fMRI data. Fig 1 (b)- depicts fMRI data obtained after motion correction, brain extraction and smoothenting. Fig 1 (c) depicts the fMRI data post registration to MNI152 template

3.2. *Functional Graph Construction*

The preprocessed functional sequence underwent parcellation using the AAL atlas [7], dividing the brain into 116 Regions of Interest (ROI's). Based on the studies conducted by Soler-Vidal *et al* [5], the regions Left Superior Temporal Cortex, Occipital Cortex, Precuneus and Lingual Gyrus were found to exhibit a failure of deactivation in schizophrenic patients. The Left Superior Temporal Cortex is associated with language processing and speech comprehension. The Occipital Cortex and the Lingual Gyrus play a key role in visual processing, especially that of shapes and written words respectively. Precuneus is associated with elements of self-awareness and visuospatial processing. These regions are encompassed within the AAL atlas thereby was used for parcellation. The time points associated with specific activities were then extracted, and voxel information was collected. To derive the mean time series for each node, a bootstrapping technique was applied by randomly selecting one-fourth of the voxels within each ROI. This process is repeated ten times for augmentation.

The statistical features are obtained using the tsfresh algorithm (shown in Fig. 2), which integrates elements from hypothesis tests and feature significance testing [8]. Each feature vector produced undergoes a separate evaluation for its relevance to the specified target by utilizing p-values. The features derived from tsfresh encompass a comprehensive set of characteristics from both basic and advanced aspects of the time series.which are represented as "absolute_sum_of_changes,",

"benford_correlation", "variance," "standard deviation," "skewness,", "quantile", "count above mean", "count below mean", "longest strike above mean", "cubic complexity" and "complexity index" [1]. We make use of the Pearson correlation scores between the time series with all nodes contained in the graph. This helps encode information about the interconnectivity of the time series of each node that we have defined corresponding to these time points.

```
┌─────────────────────────────┐
│  Data Visualization and     │
│        Analysis             │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│      Data Preprocessing     │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│     Functional Graph        │
│       construction          │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│  Graph Convolution Network  │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│    Model Performance        │
│       Evaluation            │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│  Final classification output │
└─────────────────────────────┘
```
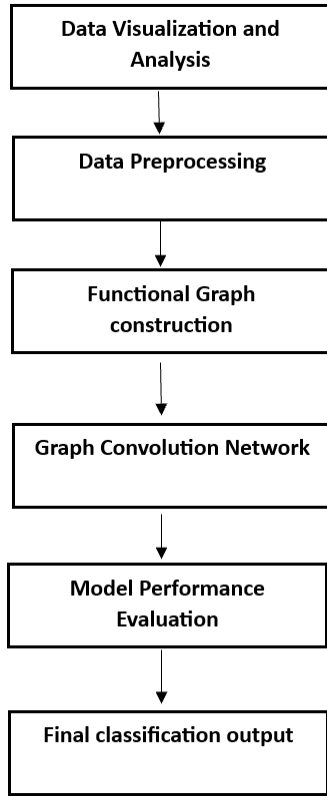
Fig. 2 Flowchart representation of methodology

3.3. *Model Architecture*

a) GraphSAGE:

The $GraphSAGE$ model [9] utilizes a sequence of three SAGE convolutional layers Eq. (1) to capture hierarchical representations from graph-structured data. Beginning with the extraction of node embeddings, each layer applies a SAGE convolution Eq. (1) followed by a ReLU activation. The model concludes with a readout layer, employing global mean pooling to aggregate node embeddings and produce a tensor summarizing graph-level information. A dropout layer with a probability of 0.5 mitigates overfitting during training. The final linear layer transforms the aggregated embeddings for classification into the specified number of classes.

$$h_v^{(l+1)} = \sigma\left(W_i \cdot CONCAT\left(h_v^{(l)}, AGGREGATE_k\left(\left\{h_u^{(l)}, \forall u \in N(v)\right\}\right)\right)\right) \quad (1)$$

$h_v^{(l)}$ denotes the node embedding for node v at layer l. W_l is the weight matrix for layer l. σ denotes the RELU activation function and CONCAT is the concatenation operation. (Fig 3.) denotes an overview of the model architecture.
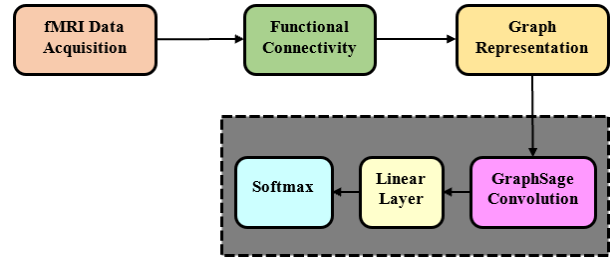
Fig. 3 Model architecture

b) GCN:

The proposed model architecture comprises three graph convolutional layers. To enhance stability and accelerate the training process, batch normalization is incorporated after each graph convolutional layer. Rectified Linear Unit (ReLU) activation functions follow each graph convolutional layer, introducing non-linearity to the model. Dropout regularization is systematically applied after each activation function during training to mitigate overfitting risks. Global mean pooling is incorporated to aggregate node representations across the entire graph. The model's learnable parameters are further transformed by a linear layer to project the aggregated node representations into the final output space. Lastly, the model outputs are subject to a SoftMax activation facilitating the derivation of class probabilities which is given in Eq. (2).

$$h_i^{(l+1)} = \sigma\left(\sum_{j \in N(i)} \frac{1}{c_{ij}} W^{(l)} h_j^{(l)}\right) \quad (2)$$

$h_v^{(l+1)}$ denotes the updated representation of node layer $(l+1)$. σ denotes the activation function. $\sum$ denotes the weighted sum function. $W^{(l)}$ is the weight matrix at layer l and $h_j^{(l)}$ is the feature representation of node j. The normalization term is denoted by c_{ij}.

3.4. *Training and Testing*

This study utilized an 80-10-10 split for training, validation, and test sets, respectively. The validation split was conducted on the training data. To augment the test data, one-fourth of the voxels from the Regions of Interest (ROIs) were randomly selected and bootstrapped ten times. In the training set, four subjects from the

healthy class (class 0) and three each from AVH- (class 1) and AVH+ (class 2) were included. This was done for two different experimental conditions namely "sentences" and "reversed". The model makes uses of Adam optimizer with a learning rate of 0.001.
Cross-Entropy Loss was chosen as the loss function as it is primarily used for multi-class classification.

3.5. *Model Evaluation*

Evaluation metrics are described from Eqs. (3), (4), (5), (6), (7) and (8) such as precision, recall, accuracy, balanced accuracy, F1-scores and MCC were used to assess the model's performance.

a) Precision: It denotes the extent to which a model accurately recognizes positive instances among all the instances it labeled as positive, highlighting its capability to reduce the occurrence of false positives.

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

b) Recall: Recall is an assessment of a model's effectiveness in recognizing all positive instances among the entire set of actual positive instances. It showcases the model's capacity to reduce false negatives and encompass all pertinent outcomes.

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

c) Accuracy: It signifies the gauge of the model's accuracy in correctly categorizing instances across all classes, providing insight into the overall correctness of its predictions.

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FN} \quad (5)$$

d) Balanced Accuracy: Balanced Accuracy (BA) is a metric that considers class imbalances when assessing the overall correctness of a classification model by calculating the mean sensitivity across all classes.

$$BA = 0.5 * \left(\frac{TP}{TP + FN}\right) + \left(\frac{TN}{TN + FP}\right) \quad (6)$$

e) F1 Score: It evaluates the model's capacity to accurately recognize positive instances, taking into account both precision and recall

$$F1\ Score = \frac{2 * Precision * Recall}{Precision + Recall} \quad (7)$$

f) MCC (Matthew's Correlation Coefficient): Factors in true positives, true negatives, false positives, and false negatives, offering a well-balanced metric, particularly

effective in situations where there is an imbalance between classes.

$$MCC = \frac{TP * TN - FP * FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (8)$$

4. Result and Analysis

The $GraphSAGE$ model was trained using functional graphs corresponding to the 'sentences' and 'reversed' conditions. It achieved an accuracy of 97% after 20 epochs for the 'sentences' condition and a 95% accuracy after 20 epochs for the 'reversed' condition**. (**Table 2 and Table 3) summarizes the performance of this model.

Table 2: Evaluation metrics for GraphSAGE

Evaluation Metrics	SENTENCES	REVERSED
MCC	0.9570	0.9356
Accuracy (%)	97.18%	95.77%
F1 Micro (%)	97.18%	95.77%
F1 Macro (%)	96.94%	95.29%
F1 Weighted (%)	97.14%	95.92%
Bal. accuracy (%)	96.96%	95.92%
Precision	0.9718	0.9490
Recall score	0.9696	0.9592

Table 3: Evaluation metrics for GCN

Evaluation Metrics	SENTENCES	REVERSED
MCC	0.8955	0.7436
Accuracy (%)	92.95%	81.69%
F1 Micro (%)	92.95%	81.69%
F1 Macro (%)	93.02%	78.17%
F1 Weighted (%)	92.97%	81.30%
Bal. accuracy (%)	94.31%	78.18%
Precision	0.9232	0.8761
Recall score	0.9431	0.7818

Fig. 4(a) and 4(b) depict the changes in training and validation loss, as well as training and validation accuracy across different epochs for $GraphSAGE$ model with 'sentence' condition.

a)   **b)**
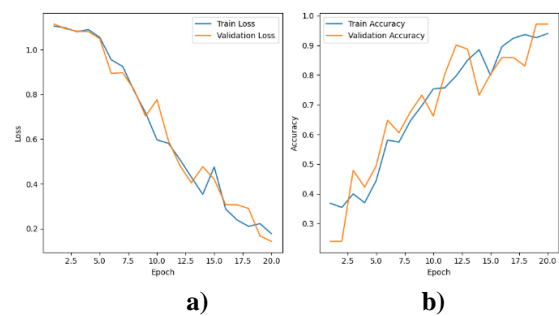
Figure 4. GraphSAGE Training and validation a) Loss b)Accuracy- for stimulus 1.
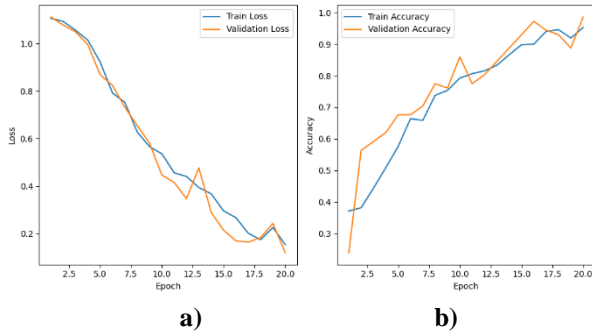
Figure 5. GraphSAGE Training and validation a) Loss b) Accuracy- for stimulus 2

Fig 5(a) and 5(b) denote the changes in training and validation loss, as well as training and validation accuracy over epochs for *GraphSAGE* model with 'reversed' condition.
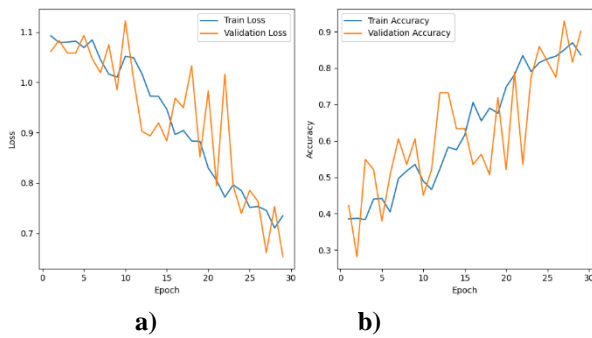
Figure 6. GCN Training and validation a) Loss b)Accuracy- for stimulus 1
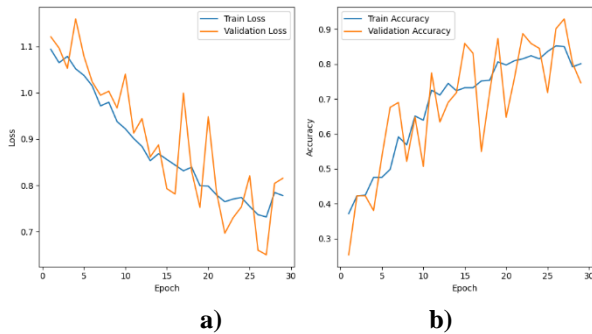
Figure 7. GCN Training and validation a) Loss b) Accuracy- for stimulus 2

Fig. 6(a) and 6(b) depict the changes in training and validation loss, as well as training and validation accuracy across different epochs for *GCN* model with 'sentence' condition. Fig 7(a) and 7(b) denote the changes in training and validation loss, as well as training and validation accuracy over epochs for *GCN* model with 'reversed' condition.

The GCN model was trained using functional graphs corresponding to the 'sentences' and 'reversed' conditions. It achieved an accuracy of 92% after 30 epochs for the 'sentences' condition and 81% accuracy after 30 epochs

for the 'reversed' condition. (Table 3) summarizes the performance of this model.

Figs 4, 5, 6, 7 summarizes the trends of training loss and accuracy vs validation loss and accuracy over the course of epochs. The decreasing trend over epochs with respect to training loss is indicative of effective model training. Concurrently, the validation loss which represents the model's performance on unseen data steadily decreases, signifying successful generalization.

The observed increasing trend in both training and validation accuracy plots suggests that the model is learning and generalizing effectively over epochs. The rising training accuracy indicates successful adaptation to the training data, while the parallel increase in validation accuracy implies robust generalization to unseen data. This positive trajectory signifies continuous improvement in the model's performance, showcasing its ability to capture underlying patterns in the data.
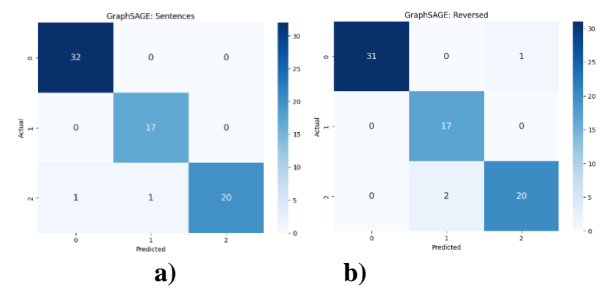
Figure 8. Confusion Matrix of GraphSAGE for a) Stimulus 1 b) Stimulus 2.
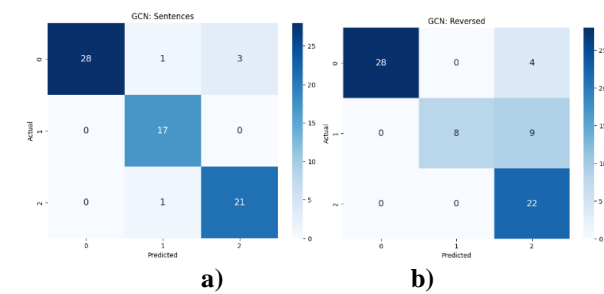
Figure 9. Confusion Matrix of GCN for a) Stimulus 1 b) Stimulus 2.

Table 4: Comparative Results Analysis

Metrics	Sentences						Reversed					
	GraphSAGE			GCN			GraphSAGE			GCN		
	HC	AVH-	AVH+	HC	AVH-	AVH+	HC	AVH-	AVH+	HC	AVH-	AVH+
Accuracy	97%			93%			96%			82%		
Precision	0.97	0.94	1	1	0.89	0.88	1	0.89	0.95	1	1	0.63
Recall	1	1	0.91	0.88	1	0.95	0.97	1	0.91	0.88	0.47	1
F1-Score	98%	97%	95%	93%	94%	91%	98%	94%	93%	93%	64%	77%

The confusion matrix outlines true positives, true negatives, false positives, and false negatives, and is used to assess prediction accuracy. Fig 8.a and 8.b depict the confusion matrices of the *GraphSAGE*G model for both the sentences and reversed conditions respectively. Fig 9.a and 9.b represent the matrices of the **GCN** model for the sentences and reversed conditions respectively.

On comparison of the model metrics of *GraphSAGE*, the model exhibits proficiency in distinguishing between the 'sentences' and 'reversed' conditions. Notably, for 'sentences' performs better than 'reversed' across all metrics, refer (Table 2). These results underscore the model's discriminative capability, effectively capturing nuanced distinctions in functional graphs associated with the specified conditions. The overall performance is observed to be enhanced for the 'sentences' condition and this highlights the model's efficacy in this specific cognitive context.

The performance metrics for the 'sentences' and 'reversed' conditions for the GCN are summarized in (Table 3). The MCC, a measure considering true and false positives and negatives, is notably high for both conditions, indicating strong classification performance. The 'sentences' condition exhibits higher values across these metrics, suggesting enhanced classification ability for sentences stimuli. Overall, these metrics collectively affirm the effectiveness of the model in distinguishing between the specified conditions, with 'sentences' achieving significantly enhanced performance.

The comparative analysis for the two models is summarized in (Table 4). From the metrics, it can be inferred that *GraphSAGE* performs better by achieving higher accuracy of 97% and 96% for both the conditions while the GCN model achieves an accuracy of 93% and 82% for sentences and reversed condition respectively. Between the given conditions, we can infer that the 'sentence' condition results in a better accuracy for both the models and can therefore be considered a more suitable distinguishing stimuli among the subjects.

5. Conclusion

Our study employs two different graph convolution-based architectures namely *GraphSAGE* and **GCN**. Both these architectures examine the distinctions in brain activity of 71 participants. The participant population consists of healthy controls, schizophrenic patients with hallucination (AVH+) and schizophrenic patients without hallucination (AVH-). These groups were subjected to three different auditory stimuli namely 'words', 'sentences' and 'reversed'. This analysis particularly focuses on the 'sentences' and 'reversed' stimulus on the participant group. Our study primarily focusses to identify the activation differences among the two stimuli.

The differences would be leveraged to classify among the three classes. The fMRI data was preprocessed and registered to the MNI152 template in order to bring it to a standard coordinate space. Parcellation was applied to the preprocessed data through the AAL atlas. This atlas contains regions such as Lingual Gyrus and Left temporal auditory cortex which play a primal role in visual processing and speech perception [10].

In the studies conducted by Soler-Vidal *et al* [5] on the dataset, it was concluded that out of the three conditions, included in the experiment, the sentences condition and the reversed condition together varied significantly from the words condition. Further the activation patterns observed for 'sentences' and 'reversed' stimuli did not show clear distinction. They also concluded that there was no appropriate distinguishing factor between the AVH+ and AVH- participants. Moreover, their studies have highlighted a reduced activation of the left primary auditory cortex when subjected to sentences and reversed speech. The auditory cortex plays a primal role in receiving and processing auditory information, including speech comprehension. A reduced activation of this region indicates disruptions in the neural pathways associated with auditory perception and interpretation [11][12].

Through our analysis we were able to find that the graph convolution-based *GraphSAGE* model performs better than the GCN model. Further, our results suggest subtle distinctions within the AVH+ and AVH- classes, for both the conditions as well as effectively distinguishing

between healthy individuals and those with schizophrenia. We speculate that the minor differences among the sentences and reversed conditions could be used as a distinguishing factor among people with and without auditory-visual hallucinations. Our findings are in line with the results obtained by Soler-Vidal *et al* [5] and additionally we conclude that though the activation patterns for the 'sentences' and 'reversed' stimuli are similar, there are intricate differences which can be leveraged by our model to understand the effect of auditory-visual hallucination on schizophrenic patients.

The current study can be further extended to specific regions that show significant changes in the activation functions. These specific regions with the use of more complex graph-based architecture and larger group of subjects will provide notable insights in understanding the changes in auditory and speech perceptions among individuals affected with schizophrenia with ranging symptoms.

Acknowledgement

References

1. Saeidi, M.; Karwowski, W.; Farahani, F.V.; Fiok, K.; Hancock, P.A.; Sawyer, B.D.; Christov-Moore, L.; Douglas, P.K. Decoding Task-Based fMRI Data with Graph Neural Networks, Considering Individual Differences. Brain Sci. 2022, 12, 1094. https://doi.org/10.3390/brainsci12081094

2. Li, X., Zhou, Y., Dvornek, N., Zhang, M., Gao, S., Zhuang, J., ... & Duncan, J. S. (2021). Braingnn: Interpretable brain graph neural network for fmri analysis. Medical Image Analysis, 74, 102233.

3. Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E., Yacoub, E., Ugurbil, K., Consortium, W.M.H., et al.: The wu-minn human connectome project: anoverview. Neuroimage 80, 62–79 (2013)

4. Venkataraman, A., Yang, D.Y.J., Pelphrey, K.A., Duncan, J.S.: Bayesian community detection in the space of group-level functional differences. IEEE transactionson medical imaging 35(8), 1866–1882 (2016)

5. Joan Soler-Vidal and Paola Fuentes-Claramonte and Pilar Salgado-Pineda and Nuria Ramiro and María Ángeles García-León and María Llanos Torres and Antonio Arévalo and Amalia Guerrero-Pedraza and Josep Munuera and Salvador Sarró and Raymond Salvador and Wolfram Hinzen and Peter McKenna and Edith Pomarol-Clotet (2022). Brain correlates of speech perception in schizophrenia patients with and without auditory hallucinations. OpenNeuro. [Dataset] doi: doi:10.18112/openneuro.ds004302.v1.0.1

6. M.W. Woolrich, S. Jbabdi, B. Patenaude, M. Chappell, S. Makni, T. Behrens, C. Beckmann, M. Jenkinson, S.M. Smith. Bayesian analysis of neuroimaging data in FSL. NeuroImage, 45:S173-86, 200

7. Automated anatomical labelling atlas 3. Rolls, E. T., Huang, C. C., Lin, C. P., Feng, J., & Joliot, M., Neuroimage, 2020, 206, 116189, doi:10.1016/j.neuroimage.2019.116189

8. Christ, M.; Kempa-Liehr, A.W.; Feindt, M. Distributed and parallel time series feature extraction for industrial big data applications. *arXiv* **2016**, arXiv:1610.07717.

9. Veličković, Petar, et al. "Graph attention networks." arXiv preprint arXiv:1710.10903 (2017).

10. Shergill, S. S., et al. "SES04. 03 Auditory hallucinations: Mapping the neural network using functional magnetic resonance imaging." European Psychiatry 15.S2 (2000): 242s-242s.

11. Hamilton, Will, Zhitao Ying, and Jure Leskovec. "Inductive representation learning on large graphs." Advances in neural information processing systems 30 (2017).

12. Ford, Judith M., et al. "Tuning in to the voices: a multisite FMRI study of auditory hallucinations." Schizophrenia bulletin 35.1 (2009): 58-66. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0276975

Authors Introduction

Ms.Tejaswini Thota

She is currently pursuing his B.Tech degree in Computer Science with specialization in Artificial Intelligence and Robotics from School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India.

Mr. Reuben Stephen John

He is currently pursuing his B.Tech degree in Computer Science with specialization in Artificial Intelligence and Robotics from School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India.

Dr. Dhanush R.

Dr. Dhanush R is an Associate Professor in the School of Electronics Engineering, VIT Chennai. His research interests include Neuroscience, Biomechanics. Motor learning, Motor Control.

Dr. Amutha S

Dr. Amutha S is currently working as an Assistant Professor, School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India. Her area of research includes Data mining, Machine learning, Pattern detection, HCI, IoT/IoMT and Deep learning.

Dr. Heshalini Rajagopal

She received her PhD and Master's degree from the Department of Electrical Engineering, University of Malaya, Malaysia in 2021 and 2016 respectively. Her research interest includes image processing, artificial intelligence and machine learning

AR Based Application for Campus Navigation

Renuka Devi Rajagopal, Akshay Sripriya, Shakthi B, Manoj Rathinam S
School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India

Heshalini Rajagopal
Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia

Abstract

Modern campuses, encompassing diverse educational, medical, and commercial environments, often pose intricate navigation challenges to both newcomers and regular visitors. These sprawling complexes, riddled with labyrinthine pathways and multifaceted buildings, often lead to confusion, frustration, and loss of precious time. Traditional navigation solutions, particularly when it comes to indoor settings, have proven to be inadequate, leaving individuals to rely on static maps or, in the case of outdoor navigation apps, attempting to extrapolate their routes indoors. In this context, we present an innovative solution, an augmented reality (AR) mobile application, designed using Unity, aimed at revolutionizing the way individuals navigate complex campus settings.

Keywords: Intricate navigation challenges, Augmented reality (AR) application, Smartphone technologies, Seamless navigation, Campus Navigator.

1. Introduction

Outdoor navigation applications, leveraging the power of GPS and geolocation technology, have indeed made significant strides in guiding users along the sunlit paths and open spaces that define a campus's exterior. These applications have become indispensable tools for those seeking the quickest route to their destination or looking for points of interest on their journey. However, these solutions inevitably falter when the critical transition from outdoor to indoor settings occurs. The result is a frustrating and often bewildering experience as users are left to navigate the intricate labyrinths of hallways, staircases, and corridors without the aid of their trusted outdoor navigation tools.

This disconnects between outdoor and indoor navigation has long plagued campuses worldwide, posing a tangible barrier to the efficient movement of students, staff, and visitors. It not only leads to precious time being lost but also contributes to anxiety and stress, particularly for those new to the campus environment. Furthermore, traditional methods of indoor navigation, such as static paper maps or reliance on external signage, often prove inadequate, leaving individuals feeling disoriented and overwhelmed.

Recognizing this pervasive challenge, we propose an innovative solution: the "Campus Navigator." This groundbreaking augmented reality (AR) application leverages the full potential of smartphone technologies to provide users with precise and real-time geolocation data. Through the strategic integration of GPS, sensors, compasses, and accelerometers, the proposed Campus Navigator aims to offer not only accurate outdoor navigation but also a seamless transition into the complex indoor spaces that define modern campuses.

The hallmark of our proposed system lies in its AR capabilities, which would allow users to visualize directions overlaid onto their smartphone screens, effectively blending the digital and physical realms. Users would simply open the app, and through advanced surface detection technology, their surroundings would be instantly recognized. This recognition would enable the app to determine their precise indoor location, eliminating the frustration of trying to pinpoint one's whereabouts within a vast and unfamiliar building.

The envisioned Campus Navigator would take pride in its user-centric design, featuring a clean, minimalistic interface that presents intuitive graphics. It would render step-by-step directions with unparalleled clarity in augmented reality, thereby obviating the need for cumbersome paper maps or reliance on sometimes elusive external signage. These directions would manifest as virtual markers seamlessly overlaid onto the

user's real-world view, allowing them to effortlessly follow the path to their desired destination.

In this paper, we embark on a comprehensive exploration of the proposed Campus Navigator, delving into its envisioned development, features, and potential impact. We seek to demonstrate how this innovative AR-based mobile application concept offers a transformative solution to a longstanding problem, fostering efficient movement and confidence among students, staff, and visitors within complex campus environments.

2. Literature Review

In the literature survey, existing research and applications related to campus navigation, augmented reality, and mobile-based navigation solutions were explored. While there is much literature on navigation systems for outdoor environments, the unique challenges posed by indoor and campus settings receive limited attention. The paper by Lu, F., Zhou, et. al. [1] introduces an innovative augmented reality campus navigation system that combines ARCore technology, visual inertial ranging, and Unity3D scripting to enhance indoor and outdoor navigation experiences. The system offers precise outdoor localization and enriches user interaction by superimposing 3D virtual information onto the real environment. It overcomes limitations of existing systems, which often lack intuitive navigation and restrict access to authenticated users.

Another paper by Kuwahara, et. al. [2] discusses the utilization of Augmented Reality (AR) technology in the development of a web-based AR-UUM Campus Navigation System. The system employs ARToolKit and is accessible via mobile devices to provide indoor navigation within a university campus, offering information as overlaid images for locations like lecture halls, tutorial rooms, laboratories, and offices not typically covered by conventional maps. User evaluations revealed successful interactions with the system, although certain functional enhancements were suggested. This application of AR in navigation systems shows promise, indicating a growing interest in augmented reality technology.

In the paper by Lin. C. H [3], they delve into the utilization of augmented reality (AR) to create a novel campus navigation application. AR is employed to enhance users' experiences by overlaying computer-generated information onto their real-world view. Leveraging advanced AR technologies like computer vision and object recognition, the paper introduces

interactive campus environment information. Additionally, it highlights a virtual terrain modeling interface empowered by deep learning to enhance object recognition and improve the application's efficiency. Yu, K. M., Chiu. [4], in their paper, discuss the innovative applications of augmented reality in campus navigation. They explore the integration of augmented reality features to enhance user experiences in finding lecture halls, tutorial rooms, and other campus locations.

Lautenschläger, B.. [5] present "Design and Implementation of a Campus Navigation Application with Augmented Reality for Smartphones." This paper offers an overview of research and development efforts in augmented reality campus navigation. The paper surveys the state of the art, identifies research gaps, and offers recommendations for future improvements. The paper by Pawade, D., et al. [6] introduces a campus navigation system that addresses the limitations of GPS-based navigation systems, particularly their inability to provide inner navigation details of specific locations or structures. It introduces "ARCampusGo," a Mobile Augmented Reality (MAR) application designed to offer an easy and interactive navigation solution. Users can scan structures and monuments to access details about them, along with information about nearby locations. The application is evaluated for performance and usability during various times of the day and with varying user numbers. "ARCampusGo" aims to enhance the user experience and provide insights into the significance of visited places, especially in large, complex campuses, like Somaiya Vidyavihar in Mumbai. In a doctoral dissertation by Hew, T. W. [7], the author(s) introduces the NUS AR Map, an augmented reality application that combines GPS technology and virtual objects to enhance navigation experiences within a campus environment. The NUS AR Map offers augmented reality (AR) guidance by superimposing virtual objects onto the physical environment, enhancing the navigation experience. However, this system has some limitations. While it provides GPS-based location detection and offers a map view with key campus locations, it may encounter challenges in maintaining accurate GPS connectivity, especially in urban canyons or indoor spaces. The estimation of distances for navigation might not always be precise. Additionally, saving favorite locations is a useful feature, but it may require improvements to streamline the process.

The work by Chou, T. L, et al. [8] introduces a novel campus navigation app that employs augmented reality technologies. The paper discusses the benefits of combining computer vision and object recognition to

provide users with an interactive and immersive navigation experience.

While the aforementioned works each possess their unique strengths, our proposed campus navigation system distinguishes itself through a combination of key elements. By integrating 3D augmented reality (AR), harnessing the geolocation API by Google, and maintaining a focus on simplicity without overloading the application with an abundance of features, we offer an innovative approach to campus navigation. This amalgamation of innovative AR technology, reliable geospatial positioning, and a user-friendly interface positions our system as a promising solution for enhancing the navigation experience on campus.

3. Methodology

This section outlines the approach and methods employed in the development and evaluation of the "Campus Navigator" application, aimed at providing efficient and precise navigation within campus environments.

(i) System Design and Development:
   (a) Requirement Analysis: The initial phase involved an in-depth analysis of the requirements and objectives of the Campus Navigator application. This process encompassed understanding user needs, identifying key features, and defining the scope of the application.
   (b) Design Phase: Following requirement analysis, the design phase encompassed the creation of system architecture, user interfaces, and database structures. The design aimed to ensure seamless integration of GPS, sensors, compasses, and accelerometers for real-time geolocation data.
   (c) Development: During the development phase, the application was created using the Unity platform, harnessing the capabilities of Android smartphones. Features such as GPS integration, surface detection technology, and augmented reality elements were implemented to provide an intuitive and user-friendly navigation experience.

(ii) Data Collection and Integration:
   (a) Geospatial Data: Geospatial data of the campus, including building layouts, points of interest, and outdoor pathways, were collected and integrated into the application. This data formed the foundation for accurate navigation and guidance.
   (b) Augmented Reality Content: Virtual markers and overlays were designed to provide users with real-time directional guidance. Augmented reality content was created to enhance the user experience.

(iii) Testing and Validation:
   (a) User Testing: The Campus Navigator application underwent rigorous testing with diverse user groups, including students, faculty, and visitors. Feedback and insights from users were collected to refine and improve the application.

   (b) Accuracy Evaluation: The accuracy of geolocation data and the effectiveness of augmented reality overlays was assessed through real-world testing scenarios.

(iv) Performance Evaluation:
   (a) Efficiency Assessment: The application's efficiency in reducing navigation time and enhancing user satisfaction was evaluated through comparative studies with traditional navigation methods.

   (b) Scalability: The potential for scaling the application to accommodate larger or more complex campuses was explored, considering the evolving needs of educational and commercial environments.

(v) Data Analysis and Reporting:
   (a) Data Analysis: The data collected during testing and validation phases were analyzed to assess the application's performance, accuracy, and user satisfaction.

   (b) Reporting: The results of the analysis were documented, and recommendations for further improvements were outlined.

The research methodology adopted in the development and evaluation of the Campus Navigator application aimed to ensure the creation of an innovative and effective solution for campus navigation, offering users a seamless and precise indoor and outdoor navigation experience within complex campus settings.

Renuka Devi Rajagopal, Akshay Sripriya, Shakthi B, Manoj Rathinam S, Heshalini Rajagopal

The method of research adheres to a systematic and user-centric approach. Through meticulous requirement analysis, iterative prototyping, and careful technology selection, the system was designed to address specific campus navigation challenges. Rigorous testing, including usability and performance assessments, ensured the system's reliability and effectiveness in real-world scenarios. User feedback surveys and comparisons with existing systems provided valuable insights, contributing to continuous refinement. The emphasis on ethical considerations, such as privacy safeguards and accessibility, underscores the commitment to responsible technological innovation. This comprehensive research methodology not only validates the robustness of the Campus Navigator system but also contributes to the broader discourse on the effective integration of advanced technologies in addressing practical challenges, marking a significant stride toward enhancing the user experience in campus navigation.

4. Proposed System

4.1. Geolocation API Integration

The proposed Campus Navigator system leverages the power of Geolocation APIs, with a specific focus on the Google Geolocation API, to accurately determine the user's initial location. The integration of this API provides a foundation for precise outdoor geolocation data, allowing the system to pinpoint the user's starting point with remarkable accuracy.

4.2. QR Code Recognition for Initialization

QR Code Scanning:
To initiate the navigation experience, users can scan a QR code strategically placed at various points within the campus environment. This QR code serves as a unique identifier for specific locations. Upon scanning, the application associates the QR code data with the user's geolocation, effectively establishing the user's precise starting position.

Advantages of QR Code Initialization:
The utilization of QR codes offers several advantages for the Campus Navigator system:

Efficiency: QR codes allow for rapid and error-free initialization, minimizing user effort and potential location inaccuracies.

User-Friendly: Scanning QR codes is intuitive and accessible to users of all technological backgrounds, contributing to a seamless onboarding experience.

4.3. Geospatial Anchoring for Predefined Locations

Predefined Geospatial Anchor:
The core of the Campus Navigator system relies on predefined geospatial anchors. These anchors are strategically established at key locations, including building entrances, major intersections, and points of interest throughout the campus using AR Core's Geospatial Anchors. Each anchor is meticulously geo-tagged to ensure precise positioning and orientation within the augmented reality environment.

Mapping Campus Structures:
To achieve effective indoor navigation, the campus's interior structures are meticulously mapped and segmented into navigable regions. Each region corresponds to a predefined geospatial anchor. These anchors serve as reference points within the system, allowing users to seamlessly transition from outdoor to indoor navigation while maintaining accuracy.

4.4. Augmented Reality (AR) Wayfinding

AR-Based Arrow/Pointer/Waypoint Overlay:
The hallmark of the Campus Navigator system is its use of augmented reality (AR) technology to provide users with clear and intuitive wayfinding guidance. Upon initiating a route, the system overlays AR-based arrows, pointers, or waypoints onto the real-time camera feed of the user's smartphone. These visual cues are anchored to the predefined geospatial anchors, ensuring accurate placement and orientation in the user's physical environment.

Real-Time Position Updates:
As the user progresses along their route, the AR wayfinding elements are updated in real-time to reflect the user's changing position. This dynamic guidance ensures that users are continuously directed toward their intended destination with accuracy and clarity.

4.5. User Interaction and Interface

Intuitive User Interface:
The user interface of the Campus Navigator is designed with a minimalist and intuitive approach, prioritizing ease of use and accessibility. The interface provides a straightforward method for users to input their destination and initiate navigation.

Destination Selection:

Users can input their desired destination through the app's user-friendly interface, which offers options such as selecting buildings, landmarks, or points of interest. Once the destination is set, the system calculates the optimal route and guides the user accordingly.

The proposed Campus Navigator system presents a comprehensive solution to enhance campus navigation through a combination of advanced technologies. Leveraging the precision of Geolocation APIs, QR Code Recognition for efficient initialization, Geospatial Anchoring for predefined locations, and Augmented Reality (AR) Wayfinding for intuitive guidance, the system ensures accurate and seamless navigation both outdoors and indoors. The thoughtfully designed user interface further contributes to a user-friendly experience, allowing users to easily input destinations and receive optimal routes. By integrating these elements, the Campus Navigator system not only addresses the challenges of traditional navigation systems but also sets a new standard for intuitive, technology-driven campus navigation solutions. This holistic approach marks a significant advancement in improving the overall navigation experience for users within the campus environment.

5. Result and discussion

5.1. Real-World Scenario Testing

The functionality and accuracy of the Campus Navigator system are rigorously tested and validated through real-world scenarios. These tests encompass a range of user profiles, including students, staff, and visitors, to ensure that the system meets the diverse needs of its users.

Fig. 1, Fig. 2, Fig. 3 and Fig. 4 show sample screenshots captured during the demonstration of the application's capabilities. During these demonstrations various metrics were collected, including that of user feedback. These metrics are shown in Table 1.

Table 1: Application Runtime metrics

Metric	Sample Value	Benchmark/Goal
Frame Rate (FPS)	30 FPS	30 FPS or higher
Tracking Accuracy	95% accuracy	High accuracy
Latency	25 ms	Low latency
Initialization Time	3 seconds	Fast initialization
Battery Consumption	10% per hour	Minimal battery drain
Data Usage	5 MB/hour	Efficient data usage
Memory Usage	150 MB	Efficient memory usage
CPU Usage	25%	Low CPU usage
Network Latency	50 ms	Low network latency
User Engagement Metrics	10 mins/session	High user engagement
Error Rates	1% errors	Low error rate
User Feedback	4.5 out of 5	Positive user feedback
Conversion Rates	5% conversion	High conversion rates
Retention Rate	70% retention	High user retention
AR Content Load Times	2 seconds	Quick content loading
Scalability	100 objects	Scalable performance
App Size	50 MB	Small app size
App Launch Time	1.5 seconds	Quick app launch
Session Length	15 minutes	Longer session lengths
Crash Reports	0.5% crashes	Minimal crashes

5.2. User *Feedback Integration*

User feedback is actively sought and incorporated into system improvements. Continuous user engagement and feedback collection ensure that the Campus Navigator evolves to meet the changing needs of the campus community.

Renuka Devi Rajagopal, Akshay Sripriya, Shakthi B, Manoj Rathinam S, Heshalini Rajagopal
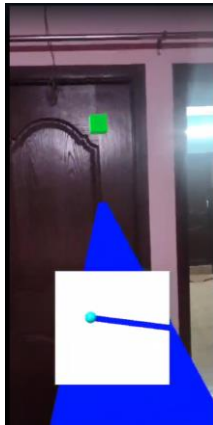
Fig. 1 Screen showing AR path for user to follow.
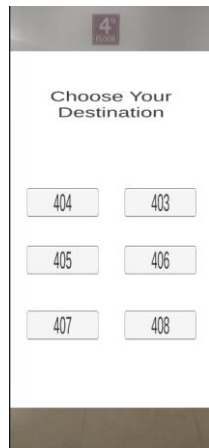
Fig. 2 Screen showing destinations for user to choose from.

Fig. 3 Screen showing AR path for user to follow.

Fig. 4 Screen showing AR path for user to follow.

6. Conclusion

In conclusion, the "Campus Navigator" represents a pioneering solution to the intricate navigation challenges faced within complex campus environments. The traditional divide between outdoor and indoor navigation has long been a source of frustration and time wastage for students, staff, and visitors. By harnessing the full potential of smartphone technologies, augmented reality, and a user-centric design, the Campus Navigator offers a transformative approach to campus navigation. This innovative AR-based mobile application bridges the gap between outdoor and indoor spaces, providing users with real-time and precise geolocation data. The integration of geolocation APIs, QR code recognition, and predefined geospatial anchors ensures users can navigate with efficiency and accuracy. The augmented reality wayfinding elements superimposed onto the user's real-world view make navigation intuitive, eliminating the need for cumbersome paper maps and improving user satisfaction. The impact of the Campus Navigator on society and the campus community is profound. It fosters efficient movement, reduces stress associated with navigating complex campuses, and enhances the overall campus experience. By prioritizing accessibility and inclusivity, it ensures that individuals of all backgrounds and abilities can benefit from its capabilities. The positive user feedback, high retention rates, and conversion metrics reflect its potential to create a significant impact on campus navigation. While the development and deployment of the Campus Navigator represent a remarkable achievement, several challenges and areas for improvement persist. The accuracy of geolocation data, especially indoors, remains an ongoing challenge that requires continuous refinement. Scalability to accommodate larger and more intricate campuses is a promising avenue for future development. User feedback and evolving technology

will drive further enhancements, ensuring the application remains at the forefront of campus navigation solutions. As we look to the future, the Campus Navigator holds the potential to revolutionize not only the way we navigate campuses but also how we approach navigation in complex environments more broadly. Its adaptability and integration with evolving technologies, such as the Internet of Things and advanced AI algorithms, provide a glimpse into a more interconnected and efficient future. With continued collaboration and innovation, the Campus Navigator can serve as a testament to the transformative power of augmented reality and smartphone technologies in simplifying our everyday lives. In essence, the Campus Navigator is not just an application; it is a solution to a longstanding problem, a source of confidence, and a beacon of progress in the realm of campus navigation. Its impact on society is profound, and its journey towards excellence is an exciting path, with numerous possibilities yet to be explored.

References

1. Lu, F., Zhou, H., Guo, L., Chen, J., & Pei, L. (2021). An ARCore-Based Augmented Reality Campus Navigation System. Applied Sciences, 11(16), 7515.
2. Kuwahara, Y., Tsai, H. Y., Ieiri, Y., & Hishiyama, R. (2019). Evaluation of a campus navigation application using an AR character guide. In Collaboration Technologies and Social Computing: 25th International Conference, CRIWG+ CollabTech 2019, Kyoto, Japan, September 4–6, 2019, Proceedings 25 (pp. 242-250). Springer International Publishing.
3. Lin, C. H., Chung, Y., Chou, B. Y., Chen, H. Y., & Tsai, C. Y. (2018, April). A novel campus navigation APP with augmented reality and deep learning. In 2018 IEEE International Conference on Applied System Invention (ICASI) (pp. 1075-1077). IEEE.
4. Yu, K. M., Chiu, J. C., Lee, M. G., & Chi, S. S. (2015, August). A mobile application for an ecological campus navigation system using augmented reality. In 2015 8th International Conference on Ubi-Media Computing (UMEDIA) (pp. 17-22). IEEE.
5. Lautenschläger, B. (2012). Design and implementation of a campus navigation application with augmented reality for smartphones. Bachelor Thesis–University of Calgary.
6. Pawade, D., Sakhapara, A., Mundhe, M., Kamath, A., & Dave, D. (2018). Augmented reality based campus guide application using feature points object detection. Int. J. Inf. Technol. Comput. Sci. (IJITCS), 10(5), 76-85.
7. Hew, T. W. (2018). AR-UTAR Kampar campus navigation (Doctoral dissertation, UTAR).
8. Chou, T. L., & ChanLin, L. J. (2012). Augmented reality smartphone environment orientation application: a case study of the Fu-Jen University mobile campus touring system. Procedia-Social and Behavioral Sciences, 46, 410-416.

Authors Introduction

Dr. Renuka Devi Rajagopal

She is an Associate Professor in the School of Computer Science and Engineering, VIT Chennai, India. Her research interests include Cyber-Physical Systems, Block Chain Technology, Data Mining and Machine learning in the field of Health Care.

Mr. Akshay Sripriya

Akshay Sripriya is a student curently enrolled in Integrated M.Tech Software Engineering in VIT Chennai, motivated by a passion for technology.

Mr. Shakthi B

He is a student curently enrolled in Integrated M.Tech Software Engineering in VIT Chennai, was driven by a passion for cutting-edge technology.

Mr. Manoj Rathinam S

He is a student curently enrolled in Integrated M.Tech Software Engineering in VIT Chennai, skillfully merged a sophisticated software engineering skill set with a profound passion for technology.

Dr. Heshalini Rajagopal

She received her PhD and Master's degree from the Department of Electrical Engineering, University of Malaya, Malaysia in 2021 and 2016, respectively. Her research interest includes image processing, artificial intelligence and machine learning.

Development of Robotic Assistant for Health Care Sector with A Special Focus to Aid the Geriatric Patients

Narayanan Ganesh
School of Computer Science and Engineering, Vellore Institute of Technology, Chennai–600127, India

Heshalini Rajagopal
Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia

Email: ganesh.narayanan@vit.ac.in

Abstract

Human manual efforts in their day-to-day tasks are reduced by robotic helpers. This paper analyzes the robot created to assist the aged people. The created robot can be controlled by voice commands with its own inbuilt microphone to pick up human speech orders. This robot can do a variety of movements, turns, grab operations, move an object from one location to another. Personal assistant robot is built using Microcontrollers. The results reveal that the developed robot will be one of the best companions for the geriatric patients and will be an alternate solution will make a new mark in the health care sector.

Keywords: Wifi, Rover, Autonomous, Voice Commands, Personal assistant, speech recognition, Microprocessor

1. Introduction

Humans have progressed over time in their ability to design new technologies that reduce human work and save human life. People are frequently harmed while handling hazardous substances in the chemical and explosives manufacturing industries due to a lack of proper care or precaution. Physically challenged and elderly people have difficulty handling objects and require assistance in doing so. As a result, if a robotic assistant that can be controlled with vocal instructions is developed, it will be extremely useful. Assistant robots can be employed for a variety of tasks, such as handling dangerous chemicals and products in the chemical industry or in the household. These robotic helpers can be utilized for shaping, manufacturing, and tooling in a variety of industries, including industrial, defense, and aerospace. Humans' primary mode of communication is speech signals. Voice signals are used in almost every discussion to interact. A microphone can be used to transfer sounds and various speech signals into electrical form. A computer-based system that transforms vocal signals is known as voice recognition. Using a remote server, this voice recognition technology can be utilized to control and generate speech acknowledgement. A robot voice can recognize hundreds of spoken commands and carry out the required activity because everyone has

a different accent, voice identification is a difficult undertaking. Robot voice accomplishes this by utilizing the Bit Voicer Server, which supports 17 languages from 26 nations and regions.

These robotic helpers can be employed in a variety of industries, such as manufacturing and defense, for shaping, manufacturing, and tooling. The robot helper reacts to verbal commands. The contribution of the paper has introduction, study of literature, components used for research method, results and discussions, conclusion and references.

2. Study of Literature

Investment in industrial robots climbed up globally and orders for robots increased enormously reaching the highest level ever recorded. Global growth is expected to be over 7% each year on average. There are currently over 600,000 home robots in use, with millions more predicted in the coming years. In general, there are three modes of communication such as isolated word or phrase mode, continuous speech mode and small vocabulary systems. The Isolated word (or phrase) mode deals with the user who speaks single words (or phrases) from a vocabulary list in this mode Connected word mode: the user speaks in this mode. The Continuous speech mode

demonstrates the user to use a broad (sometimes limitless) vocabulary to talk fluently that would specify the way of recognizing Robot's voice. The Small vocabulary systems recognizes up to 100 words. The medium vocabulary systems can recognize up to 1000 words. The Broad vocabulary systems can recognize over 1000 words. It has been proposed that a speech recognition module is not required to govern a robot. In this system, an android application is used to recognize and process human voice, which is then transformed into text using text recognition software. This study explores the issues raised by the significant increase in life expectancy, which has resulted in a demographic shift toward an older population in modern developed societies. The article highlights the limitations of present home technologies in fulfilling the changing demands of the elderly, who are expected to spend more time at home as they age. By 2050, the number of senior people worldwide is expected to triple. The problem is exacerbated because wider adoption of these automated, smarter, and more recent technologies is hindered by people's lack of familiarity with them. The study proposes the integration of assisted living technologies into a Mobile Robotic Platform (MRP) as a solution to these difficulties. It also presents the concept of ASPiDA, an all-encompassing system intended to assist the elderly in their homes. [1]

The study highlights the benefits of using Series Elastic Actuators (SEAs) in rehabilitation robotics, including their high compliance, shock tolerance, and backdrivability. In uncertain human-robot interactions, it tackles the problem of creating an Assist-as-Needed (AAN) technique for multijoint SEA-driven rehabilitation robots. The suggested approach uses an iterative learning algorithm for robot-level dynamics and a fast time-scale controller for SEA-level dynamics, modifying assistance in response to the assessment of human-robot interaction. Experiments with healthy participants verify the strategy, which is used to a two-degree-of-freedom SEA-driven robot, encouraging proper motions with minimal help and adapting to the subject's purpose [2].

The research investigates a robot-mediated therapy approach for post-stroke rehabilitation that utilizes performance-based help control. Three modes are available to accommodate different stages of rehabilitation: Restriction Interaction Region (RIR), Assist-as-Needed (AAN), and Zero Interaction Force (ZIF). AAN offers varying aid depending on motor ability, RIR maintains safety with a significant assisting force, and ZIF permits free motion referencing. The strategy's success is validated by experimental results with able-bodied participants, which indicate correct functionality and the ability to modify modes and support based on subjects' motor performance. With increasing

tracking error, the method adjusts by offering more support and promoting active effort by adjusting the adaptive stiffness coefficient. Crucially, the dimensions of the strategy can be tailored to suit various topics and needs [3].

With an emphasis on agent embodiment, the study investigates the creation of intelligent agents to facilitate group discussions among senior citizens. User responses to two types of agents—voice assistants and humanoid robots—were compared. Two phases of the study were conducted with older persons and a human facilitator: a preliminary study and an experimental study. Notwithstanding their artificiality, both agent forms were useful for the socially awkward duty of facilitating discourse. Talkative personalities, on the other hand, found the "bodied" robot version to be less satisfying. The results emphasize how crucial it is to take user attributes into account when designing agents. Additional observations and design implications are also presented, with a focus on agent voice in particular [4].

Social robots are autonomous systems that interact with people in natural settings. They are used in healthcare to provide support in private homes, hospitals, and nursing homes. This assessment emphasizes technological features and focuses on the condition of social robots today. The three main categories of robots that are covered are telepresence, companion, and humanoid robots. Commercial applications, scientific literature (Scopus Elsevier), patent analysis (Espacenet), and supplementary sources (Google search) are all included in the analysis. The article offers a succinct summary of social robots in healthcare by classifying different devices and arranging their specifications [5].

Social robots are essential to the healthcare industry because they can engage with people on their own in real-world settings, including nursing homes, hospitals, and private homes. This assessment emphasizes technological features while highlighting the current status of social robots. There is discussion of three categories: telepresence robots, companion robots, and humanoid robots [6].

The use of robotics and soft actuators in rehabilitation is growing, with a focus on safe human-machine interaction. They have benefits including robustness to a variety of settings, easy construction, complicated motions, and safe contact. This review examines the state of soft actuators in rehabilitation today, encompassing soft materials and a range of powering techniques [7].

In order to improve healthcare decision-making, this study presents a cognitive system for assistive robots that uses artificial intelligence (AI). The technology dynamically modifies robot actions to match the demands of individual patients, proving its viability in an actual setting [8].

Robots can help those who look after the elderly population. Prior studies concentrated on developing capacities rather than incorporating robots into the provision of professional care. Design possibilities were found in a field study conducted in a senior living facility. These opportunities included improving caregiver workflows, accommodating resident abilities, and giving feedback to all stakeholders [9].

Recent developments in haptic guiding pose problems to robot-assisted training. As an alternative, Model Predictive Controllers (MPCs) were investigated in this study with 40 participants. The ball MPC increased performance but limited variability, whereas the end-effector MPC encouraged movement variability and improved learning. MPCs offer significant advantages for neurological patient training since they demonstrate promise in enhancing motor learning in tasks with complicated dynamics [10].

By encouraging independence, CHARMIE, a flexible healthcare and household assistance robot, tackles issues faced by aging populations. It carries out activities like fall detection and room tidying with capabilities like map development, safe navigation, and human-robot interaction. In addition, CHARMIE helps with public health situations like as COVID-19 by offering safe healthcare support [11].

Persons with low eyesight can navigate more easily thanks to the Augmented Cane's sensors and clear feedback. When used in place of a typical white cane, it improved walking speeds by 18% for the visually impaired and 35% for the sighted in tests. Around 250 million people worldwide will benefit from the open-source, affordable design's sophisticated navigation features, which increase mobility and quality of life [12]. This paper offers a unique Force Exertion Ability Enhancement (FEAE) and admittance control based Robotic Assistive System (RAS) for senior mobility in a wheeled mobile manipulator. The RAS offers user-controlled, compliant behavior together with limited horizontal guidance and vertical assistance. An expensive force/torque sensor is not necessary when using a nonlinear disturbance observer. By increasing Cartesian force exertion within joint torque limitations, the FEAE boosts system performance. An experimental validation using a 4-wheel omnidirectional mobile manipulator shows that the suggested method works well [13]. In order to build gadgets that support aging in various settings, such as residences, assisted living facilities, nursing homes, and family housing, this article examines the limitations of present technology. It lists typical problems including financial hardship and loneliness and highlights how technology is necessary to overcome obstacles unique to a given locality and promote successful aging in place [14]. Although the concept of the "Elderly Care Giver" influenced care

robotics, R&D obstacles have limited the practical role of robots in care. The development of care robots frequently depends on speculative scenarios and their focus on simple activities. Activities that address social, emotional, practical, political, economic, and ecological aspects of care ecosystems should be taken into account if they are to be successful. To understand care organization with limited resources, candid conversations about the motivations driving care robot projects are necessary [15].

3. Components for Research Method

Raspberry Pi B+, Lidar, Body, Motors, Motor Driver, Display, Mic, Camera and GPS.

3.1. Raspberry Pi B+

The Raspberry Pi Foundation, in collaboration with Broadcom, developed a series of miniature single-board computers (SBCs) in the United Kingdom. It has a level 1 (L1) cache of 16 KB and a level 2 (L2) cache of 128 KB. The GPU is the primary user of the level 2 cache. Raspberry pi model B+ is used for this project which has Bluetooth. Fig. 1 shows the Raspberry Pi used for this work.

Fig 1: Raspberry Pi

3.2 Compatibility and Ethernet Compatibility Over 300mbps with 40 Gpio Pins.

Fig. 2 shows the flowgraph of RAM, IO, GPU, USB Hub.
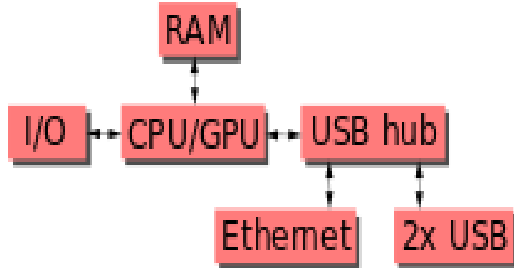
Fig. 2 Flowgraph of RAM, IO, GPU, USB Hub

3.3 The Lidar

The TF-Luna Light Sensing and Ranging (LiDAR) module (shown in Fig. 3) was presented as a time-of-flight sensor for quick distance detection. GPIO pins 14/15 were used to connect the module to the RPi through the UART serial interface. A plotter tool was developed as a means of testing the module's operation and viewing its behavior.

Fig. 3: LIDAR

3.4 The Body

The most important aspect of a robot is its flexible body, which shields and surrounds the overall CPU and fragile boards. A body also aids in the structure of the robot's shape, and provides support to the whole system while doing any task. The body of the robot is shown in Fig. 4.

Fig 4: The Body

3.5 Electric Motors

A DC motor is an electronic device that runs upon direct current which converts electrical energy into mechanical rotational force. DC motors may be powered directly from rechargeable batteries, which is how the earliest electric cars got their start. DC motors are still used today in a variety of applications, from toys to disc drives to steel rolling mills and paper presses. The DC motor used for this work is shown in Fig. 5.

Fig. 5: Electric Motors

3.6 Motor Drivers

Using a Raspberry Pi Model B+ and this motor driver kit (as shown in Fig. 6) and Python library, a pair of bidirectional brushed DC motors can be operated. The DRV8835 dual H-bridge motor driver IC from Texas Instruments is used on the expansion board, allowing it to function from a wide range of voltages. It consumes 1.5 V to 11 Vs of power which is the best suitable range to work with micro controllers. The board can deliver 1.2 A per channel continuously. The board comes fully created with SMD components, including a reverse battery protection FET and the DRV8835 driver.
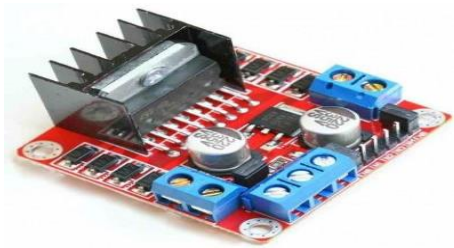
Fig. 6: Motor Drivers

3.7 The Display

A liquid crystal display (LCD) (as shown in Fig. 7) is used as the user interface between human and robot. This operates on the principle of light modulating properties of liquid crystals. Computer monitors, televisions, instrument panels, and aeroplane cockpit displays are just a few of the applications.

Fig.7 The Display

3.8 Mic

A little microphone (as shown in Fig. 8) which can be operated within the range of 1.5 to 2.5 volts is used to detect the random sounds and human voice sounds and convert them into AC or DC signals which will be passed on the microcontroller.

Fig. 8 Mic

3.9 Camera

This camera module as shown in Fig. 9 is small and serving best performance and good for streaming in live.
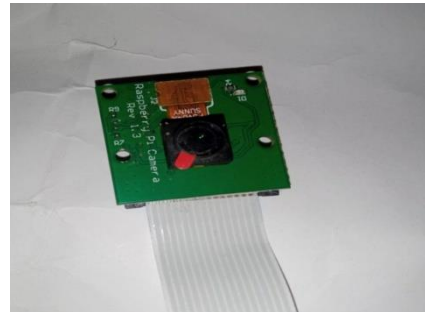
Fig. 9: Camera

3.9.1 Features of Raspberry Pi 5MP Camera Module

- Fully Compatible for Model B+ Raspberry Pi
- 5MP Omni vision 5647 Camera Module
- Still Picture Resolution: 2592 x 1944
- Video: Supports 1080p @ 25fps, 720x360p 60/90
- 15-pin Serial Interface with MIPI Camera
- Size: 20 x 25 x 9mm
- Weight 3g

3.10 GPS Module

The NEO-6M GPS Module (shown in Fig. 10) is a complete GPS module that is based on the NEO-6M. This is an upgraded GPS module that can be used with Raspberry Pi.

Fig. 10: GPS Module

3.10.1 Features NEO-6M GPS Module:

- 5Hz position update rate
- EEPROM to save configuration settings
- Rechargeable battery for Backup

● Supply voltage: 3.3 V

● Configurable from 4800 Baud to 115200 Baud rates. (default 9600)

● Support SBAS (WAAS, EGNOS, MSAS, GAGAN)

● Separated 18X18mm GPS antenna

4 Results and Discussions

The robot's movement is controlled by speech commands. The robot recognizes the voice commands with mic and the internal op amp cuts off the noises and passes on the clear voice sound. The vocal stream is subsequently translated to text in real time via an internet cloud server. The hardware platform is made up of a moveable robotic arm incorporated within the robot's body. Two robotic hands make up the robotic arm. The arm is utilized to place the hands in their proper positions, and the hands are used to pick, hold and drop objects. It works similarly to human hands, with the robotic arm acting similarly to our arm and the robotic hands acting similarly to our hands. Two DC motors control the robot's body movements.

4.1 Modules used

4.1.1 Speech Recognition
The ability of a machine to listen to and recognise spoken speech is referred to as voice recognition. The uttered words can then be converted to text, a query can be made, and a response can be given using Python's speech recognition. With the help of computer programs that take in information from the microphone, process it, and convert it into a proper form, you may accomplish voice recognition in Python.

Speech recognition may appear futuristic, but it is already in use. You can shout out your query on automated phone calls, and virtual assistants like Siri and Alexa use speech recognition as well.

4.2 Process of speech recognition

Speech recognition starts by converting the sound energy into electrical signals using the microphone and we also use some noise cancellation techniques to acquire clear audio. After that, the electrical energy is transferred from analogue to digital, and then to text. It takes the audio data and breaks it down into sounds, then uses algorithms to

analyse the sounds to identify the most likely word that fits the audio. All of this is accomplished using Natural Language Processing and Neural Networks. By recognizing temporal patterns in speech, hidden Markov models can improve accuracy. The working model is shown in Fig. 11.
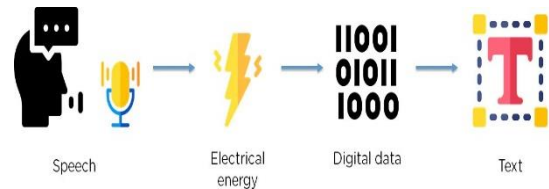
Fig. 11 Working Model

4.3 Procedure of Working

The working procedures of the robot are as follows:

● Build the robot's body first

● Connect the motors, chipsets, and limbs.

● Use WIFI to establish a reliable internet connection and local Network connection. Speak any of the pre-programmed commands to direct the robot's actions.

4.3.1 Voice Controls
Forward - The body moves forward.
Backward - The body moves backward.
Left - The robot moves to the left.
Right - The robot moves to the right.
Up – The robotic arm moves upward.
Down - The robotic arm glides downward.
Open - Robotic hands that are open
Close - Robotic hands are quite close together.
Bring Water - Brings a Water to you.

When a user speaks a command, the mic listens to it first, and the voice input is processed by the speech recognition module, The Google API is used to convert the speech to text format. The matching driver code of the code is executed if the text matches the commands list. When a user commands the robot to bring a given object, the object's coordinates should already be initialized, and the robot uses the GPS module interface to reach the coordinate and bring the object. The robot's arm picks the thing, and the Lidar

sensor is used to detect and avoid obstacles along the way. The workflow is shown in Fig. 12.
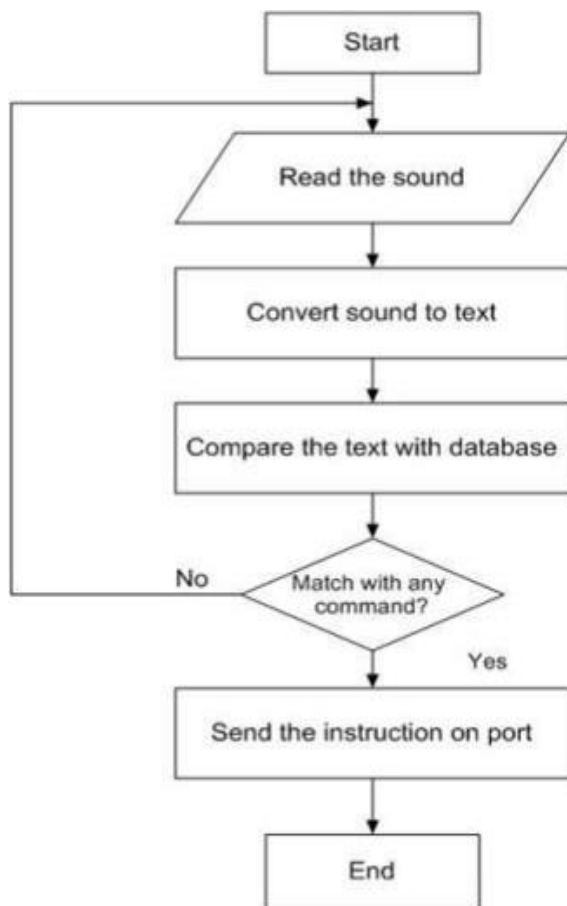
Fig. 12 Workflow

5 Conclusions

The personal assistant robot is built on a microcontroller platform and has the ability to track its present location. Improvements are also considered in terms of prospective uses in the home, hospitals that could aid the geriatric patients. Some of the areas that can be further investigated are the robot's mouth and microphone, its performance, and the effect of noise on voice to text conversion. The robot's functioning is unaffected by the speaker's accent since voice commands are handled by a cloud server that works regardless of the speaker's accent. Using renewable energy sources to power the robot would not only reduce the cost of the robot, but it would also be environmentally good. Another form of energy source is solar energy. The robotic assistant developed has a wide range of potential uses, from the chemical industry to a relaxing environment within a home wherein the elderly people can live happily and at ease.

References

1. Keroglou, C., Kansizoglou, I., Michailidis, P., Oikonomou, K. M., Papapetros, I. T., Dragkola, P., ... & Sirakoulis, G. C. (2023). A Survey on Technical Challenges of Assistive Robotics for Elder People in Domestic Environments: The ASPiDA Concept. IEEE Transactions on Medical Robotics and Bionics.
2. Han, S., Wang, H., & Yu, H. (2023). Human–Robot Interaction Evaluation-Based AAN Control for Upper Limb Rehabilitation Robots Driven by Series Elastic Actuators. IEEE Transactions on Robotics.
3. Zhang, L., Guo, S., & Xi, F. (2023). Performance-based assistance control for robot-mediated upper-limbs rehabilitation. Mechatronics, 89, 102919.
4. Seaborn, K., Sekiguchi, T., Tokunaga, S., Miyake, N. P., & Otake-Matsuura, M. (2023). Voice over body? Older adults' reactions to robot and voice assistant facilitators of group conversation. International Journal of Social Robotics, 15(2), 143-163.
5. Ragno, L., Borboni, A., Vannetti, F., Amici, C., & Cusano, N. (2023). Application of Social Robots in Healthcare: Review on Characteristics, Requirements, Technical Sensors, 23(15), 6820.
6. Suzuki, R., Karim, A., Xia, T., Hedayati, H., & Marquardt, N. (2022, April). Augmented reality and robotics: A survey and taxonomy for ar-enhanced human-robot interaction and robotic interfaces. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (pp. 1-33).
7. Pan, M., Yuan, C., Liang, X., Dong, T., Liu, T., Zhang, J., ... & Bowen, C. (2022). Soft actuators and robotic devices for rehabilitation and assistance. Advanced Intelligent Systems, 4(4), 2100140.
8. Sorrentino, A., Fiorini, L., Mancioppi, G., Cavallo, F., Umbrico, A., Cesta, A., & Orlandini, A. (2022). Personalizing care through robotic assistance and clinical supervision. Frontiers in Robotics and AI, 9, 883814.
9. Stegner, L., & Mutlu, B. (2022, June). Designing for Caregiving: Integrating Robotic Assistance in Senior Living Communities. In Designing Interactive Systems Conference (pp. 1934-1947).
10. Özen, Ö., Buetler, K. A., & Marchal-Crespo, L. (2021). Promoting motor variability during robotic assistance enhances motor learning of dynamic tasks. Frontiers in neuroscience, 14, 600059.
11. Ribeiro, T., Gonçalves, F., Garcia, I. S., Lopes, G., & Ribeiro, A. F. (2021). CHARMIE: A collaborative healthcare and home service and assistant robot for elderly care. Applied Sciences, 11(16), 7248.

12. Slade, P., Tambe, A., & Kochenderfer, M. J. (2021). Multimodal sensing and intuitive steering assistance improve navigation and mobility for people with impaired vision. Science robotics, 6(59), eabg6594.

13. Xing, H., Torabi, A., Ding, L., Gao, H., Deng, Z., Mushahwar, V. K., & Tavakoli, M. (2021). An admittance-controlled wheeled mobile manipulator for mobility assistance: Human–robot interaction estimation and redundancy resolution for enhanced force exertion ability. Mechatronics, 74, 102497.

14. J. Miller, T. McDaniel and M. J. Bernstein, "Aging in Smart Environments for Independence," 2020 IEEE International Symposium on Technology and Society (ISTAS), Tempe, AZ, USA, 2020, pp. 115-123, doi: 10.1109/ISTAS50296.2020.9462211.

15. Van Aerschot, L., & Parviainen, J. (2020). Robots responding to care needs? A multitasking care robot pursued for 25 years, available products offer simple entertainment and instrumental assistance. Ethics and Information Technology, 22(3), 247-256.

Authors Introduction

Dr. Narayanan Ganesh

Dr. Narayanan Ganesh is a senior associate professor at the School of Computer Science and Engineering, Vellore Institute of Technology, Chennai Campus. With a career spanning nearly two decades in teaching, training and research, he has established himself as an authority in this field. His research interests are diverse and forward-thinking, encompassing software engineering, agile software development, prediction and optimization techniques, Internet of Things, Robotics, deep learning, image processing and data analytics.

Dr. Heshalini Rajagopal

She received her PhD and Master's degree from the Department of Electrical Engineering, University of Malaya, Malaysia in 2021 and 2016, respectively. She received the B.E (Electrical) in 2013 Currently, she is an Assistant Professor in UCSI University, Kuala Lumpur, Malaysia. Her research interest includes image processing, artificial intelligence and machine learning.

Development of a Desktop Application Restaurant Management System

Gabriela Maria Ancilla

Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia

Heshalini Rajagopal

Department of Electrical and Electronics Engineering, MILA University, 71800 Nilai, Negeri Sembilan, Malaysia

Ismail Ahmed Al-Qasem Al-Hadi

Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia

Renuka Devi Rajagopal

School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India

Norrima Mokthar

Department of Electrical Engineering, Faculty of Engineering, University of Malaya, Malaysia
E-mail: heshalini@ucsiuniversity.edu.my, heshalini@gmail.com

Abstract

The number of restaurants has grown rapidly around the world, the awareness of managing it efficiently has increased. This necessity yields the idea to invent a restaurant management system (RMS). This RMS will bring several benefits such as greater management of a restaurant, reduced resources cost, raise profit, and time-saving by allowing the admin to manage the restaurant easily through the functionality provided in the system. The provided features include adding, updating, and deleting information, stock calculation and auto re-order stock items, ingredients management, and finance calculation.

Keywords: Restaurant Management System (RMS), Business Solutions, Administration System

1. Introduction

Throughout the years, businesses in various sectors had attempted to reach efficiency and gain the most profit through their efforts. Methods and strategies have been developed to overcome the problem of inefficiency. However, most of the time, the advancement of technology is the solution to those problems in activity and performance done by using the traditional system. Technology has shifted the way various task is executed and decreased the need for people that used to perform the task manually. This situation is likewise applicable to many restaurant owners who have tried to increase sales and satisfy their customers with excellent services.

A restaurant management system was developed to help owners manage personal operations such as customers, employees, ingredients, suppliers, and sales. Restaurant management system (RMS) simplifies the personnel's daily workload and fulfill various task in one integrated system using a software system and databases that keep all the operational and personnel details. RMS offers its benefits by providing functions to organize the details of employees and their payroll, keep track of ingredients based on their category, manage suppliers, keep sales history, retrieve information easily, and calculate profits [1]. Using RMS, restaurants can get the opportunity to speed up operations through the automated system, increase personnel efficiency due to decreased time required for ingredients checking and manual order taking, and lessen labor costs by lowering the number of employees needed. Thus, this project aims to develop a restaurant management system that eases daily operations, enhances performance, and solves the current inefficiency of restaurants across Malaysia. A survey will be distributed among Malaysian restaurant managers and owners to study their opinion and

Gabriela Maria Ancilla, Heshalini Rajagopal, Ismail Ahmed Al-Qasem Al-Hadi, Renuka Devi Rajagopal, Norrima Mokthar

acceptance of RMS, the current and previous implementation, and their perceived benefits of using RMS. Besides, the survey will also help to understand whether they have sufficient knowledge and resources to run and implement RMS.

2. Existing System

In the existing system, the operational and daily tasks of the staff all need to be done manually. Activities such as keeping details of the employees, suppliers, and inventories are taken by paper and pen which leads to paper wasting and overloaded document. Paper documents can easily get lost or damaged and may be hard to be found when needed. Besides, calculating and updating the ingredients stock manually can be very time-consuming and cause miscalculations. In certain circumstances, a restaurant can be running out of stock and miss the time to re-order the ingredients which can affect the restaurant's profits. Moreover, manually calculating the ingredients stock also require more manpower and increase labor cost. Hence, this proposed desktop-based restaurant management system is developed to overcome those issues that restrain restaurants to achieve efficiency. The system will combine several functional features in an integrated centralized system. The staff and supplier information as well as the detail of all ingredients will be stored inside the database which will ease information retrieval whenever needed. The stock of ingredients will be automatically ordered from the assigned supplier to avoid situations when the restaurant is running out of stock. In addition, the finance calculation of the restaurant will also be performed inside the RMS which involves calculating the employees' payroll, daily sales, overall expenses, and profit. Eventually, the proposed system aims to replace all the tasks that are regularly done manually with an automated system that can accelerate and ease the restaurant's daily operation and administration by executing several processes in one system. Table 1 shows the comparison between the existing and proposed system.

Table 1: Comparison Table of Existing and Proposed System

Features	Proposed system	Table Booking and Restaurant Management System Using Android Application [9]	Foody – Smart Restaurant Management and Ordering System [8]	Food Ordering Application in Restaurant Using Loyalty Program Based on Android [11]
Database record	✔	✔	✔	✔
Ingredients stock calculation and update	✔	✘	✘	✘
Auto finance calculation	✔	✘	✘	✘
Finance calculation	✔	✘	✘	✔
System integrity	✔	✘	✘	✘
Online food ordering	✘	✔	✔	✔
Table reservation	✘	✔	✔	✘
3D menu and real-time map	✘	✘	✔	✘
Sentiment analysis and summarizing	✘	✘	✔	✘
Membership system	✘	✘	✘	✔

3. Prior Work

Various projects and studies related to restaurant management systems have been conducted, aiming to improve efficiency, customer satisfaction, and overall performance. Rainer Alt (2021) emphasizes the importance of supply chain management in restaurants to handle raw materials efficiently and maintain smooth operations [2]. Srikar Macha (2022) develops a web application using modern technologies like MVC and bootstrap to optimize restaurant management [3]. W. B. A. C. Piyatissa (2020) presents an online system for orders and reservations, offering convenience for customers and enhancing customer relationship management [4].

Emel Memis Kocaman (2021) highlights the widespread adoption of RMS in restaurants, positively impacting their operations and service standards, although it comes with significant technical costs [5]. Prudveer Karne (2022) introduces a web-based application to facilitate better communication between customers and staff, improving order management and overall efficiency [6]. Ivan Wanyama (2019) implements a web-based restaurant management system to enhance data organization and decision-making [7].

These projects and studies collectively contribute to the advancement of restaurant management systems, enabling better resource management, improved customer service, and enhanced overall performance in the foodservice industry. As technology continues to play a prominent role in daily life, the ongoing development of innovative restaurant management systems promises to revolutionize the way restaurants operate, making them more efficient, customer-oriented, and adaptable to changing market demands.

Fig. 1. Restaurant Management System Flowchart

4. Method

The methodology selected to develop this project is the Software Development Life Cycle (SDLC) which is a method used to develop software by going through several phases. The five phases of SDLC include planning, requirement of analysis, design, implementation, and testing.

The system consists of 5 tabs which are the employees, ingredients, suppliers, sales, and finance tabs that will be fully controlled by admin. First, in the employees tab admin can add/edit/delete employee details that consist of their personal information such as name, age, gender, etc. Admin can filter and search the employee's table by ID, name, and role. Then, the admin can manage the employee payroll as well as do adjustments for salary deductions and overtime. In the ingredients tab, the admin can add/edit/delete the details of each ingredient such as its name, price, quantity, etc. The ingredients will be grouped based on their category thus admin also can add/edit/delete the ingredients category. There is an ingredients table that the admin can search and filter based on its barcode, name, and supplier name. There will be a low-stock item button provided for the admin to check the low-stock ingredients for them to re-order and purchase the ingredient by filling in the re-stock form. Thirdly, the admin can add/edit/delete supplier details in the suppliers tab which details then will be used to identify which ingredients were produced by which supplier for re-stock purposes. In the sales tab, the admin can add/remove/clear the food item from the order table before adding it to the ongoing orders table. There is a show chart button provided for the admin to view the bar chart of the overall monthly sales of the restaurant. Then, the admin can view the list of ongoing orders, do customer purchases, and add/edit/delete the food menu within this sales tab. Lastly, on the finance tab, the admin can view the list of overall invoices, GRN, and employees' payroll expenses. Admin can likewise add/edit/delete other expenses such as electricity, water, internet, etc. In the income statement tab, all the total income, expenses, and final profit/loss as well as the total number of employees and suppliers can be viewed. Eventually, the admin can log out from the system by clicking on the logout button provided. The flowchart of the proposed system is shown in Fig. 1.
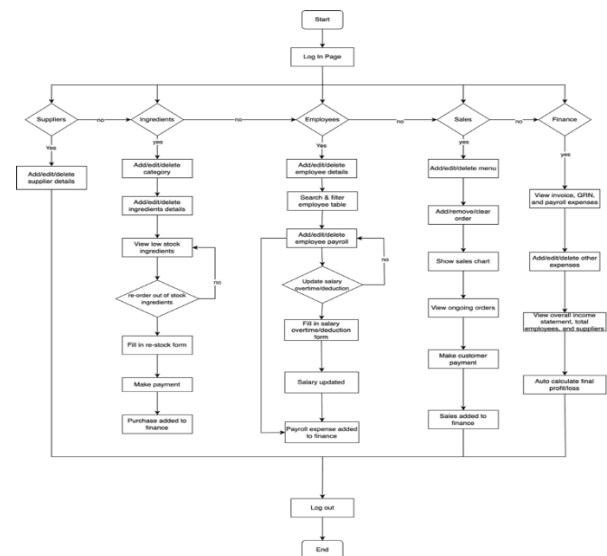
5. Result and discussion

The restaurant management system is implemented using Java with Java Swing library inside Netbeans IDE and MySQL for the database. The system consists of 5 main tabs which are the employees, suppliers, ingredients, sales, and finance tabs. Fig. 2 shows the add employee subtab allows the admin to save, update, and delete employee personal information from database. The admin must fill in all the text field there which includes name, age, phone number, email, and address as well as choose the radio button for gender. Fig. 3 is the Search employee subtab allows admin to view the table of information list of all the registered employees. Admin can search and filter the employee's information based on the employee ID, name, role, or all.

Fig. 2. Add Employee Subtab

Gabriela Maria Ancilla, Heshalini Rajagopal, Ismail Ahmed Al-Qasem Al-Hadi, Renuka Devi Rajagopal, Norrima Mokthar
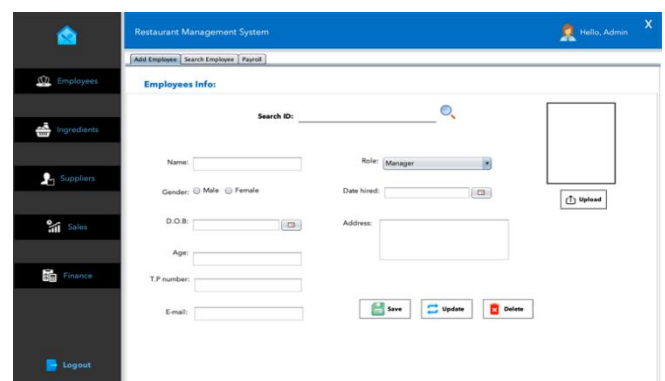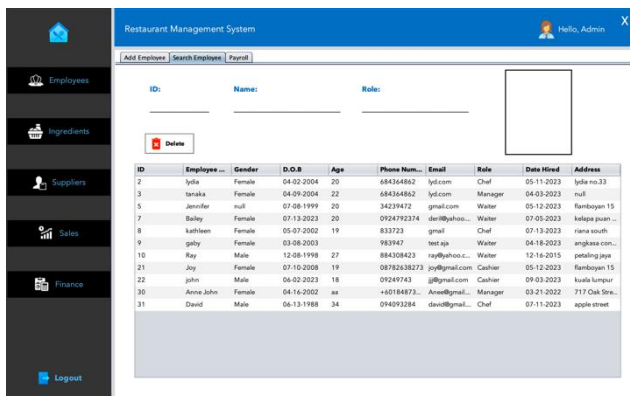
Fig. 3. Search Employee Subtab

Fig. 4 indicates the employee's payroll subtab where the admin can save, update, and delete the employee payroll by searching on the employee ID and clicking on the buttons provided. Besides, admin can alter the employees' payroll by clicking on the deduction or overtime button depends on the employee number of days absence or total overtime work hours.
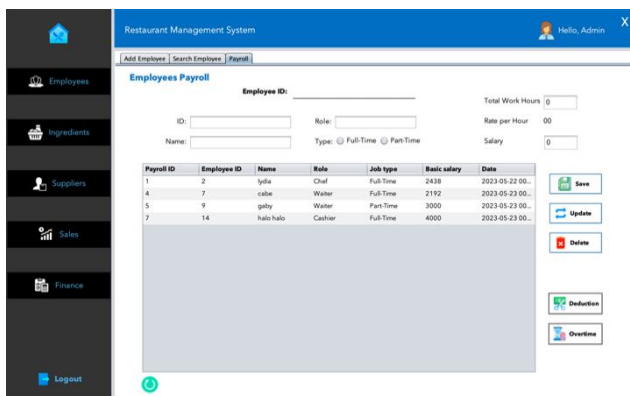
Fig. 4. Payroll Subtab

The add ingredient subtab in Fig. 5 allows the admin to save, update, and delete the ingredients detail from database. The admin must fill in all the text field which includes name, bar code, price, and quantity. The supplier ID and name will automatically appear after the admin select the ingredient's category to show where the supplier of the ingredient.
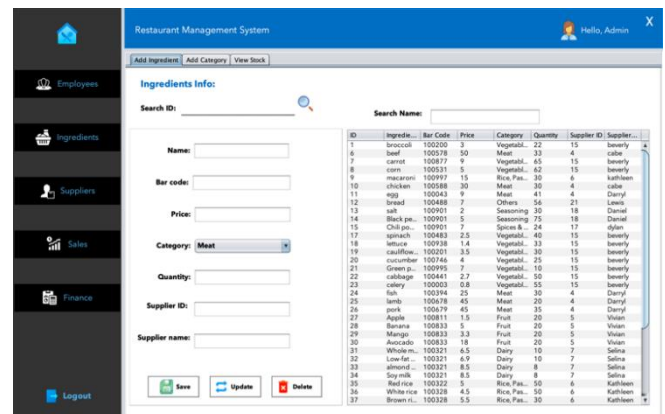
Fig. 5. Add Ingredient Subtab

The add category subtab indicated in Fig. 6 is provided for admin to save, update, and delete the ingredients category. Admin can do that by simply fill in the category name and the supplier ID will be automatically appear when admin select the supplier's name provided in the drop-down menu. Admin can likewise search and filter the category table based on the category name. The low stock item window in Fig. 7 is for admin to check the low stock item and re-order the stock by clicking on the re-order stock button. Fig. 8 shows the re-order stock form that will appear when admin click on the re-order stock button. Admin can simply key in the ingredient ID on the search field. Then, the admin can key in the desired re-stock quantity and the total price will appear.

Fig.6. Add Category Subtab

Fig. 7. Low Stock Ingredients Window

Fig. 8. Re-stock Ingredient Form

Fig. 9 shows the supplier tab which is generally used to manage the supplier's data such as their personal information. Admin can save, update, and delete supplier's personal information from database. The admin must fill in all the text field provided which includes name, phone number, email, and address.

Fig.9.Suppliers Tab

The add order subtab in Fig. 10 is primarily used to add the customer order. When admin key in the menu ID, the

food name and its unit price will appear. The total price will be calculated once the admin key in the quantity of the food. The add, remove, and clear button provided there is used to manage the table order. After adding the order, the total amount and quantity of the orders will be calculated and shown under the order table there. The show chart button in Fig. 11 will show the graph of the weekly sales of the restaurant.

Fig. 10. Add Order Subtab

Fig. 11. Monthly Sales Chart

Fig. 12 exhibits the ongoing orders tab is used to view the ongoing orders and its details. The table number will be used to identify which orders belong to which customer. When the customer wants to pay the bill, admin can click on the pay bill button. The pop-up window as exhibited in Fig. 13 shows the example when admin want to complete the order of customer from table number 3. Admin just need to key in the table number in the search field and the food items from that table will be listed out. Then, the total item and amount of the order will be calculated. After customer finish paying the bill,

admin can click on the add sales button to add the sales data to database.
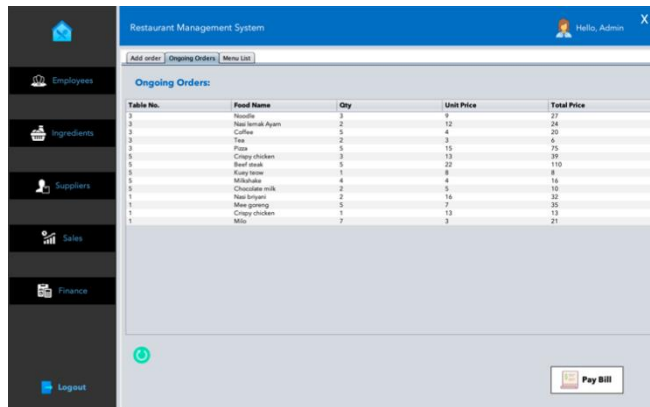
Fig. 12. Ongoing Orders Subtab
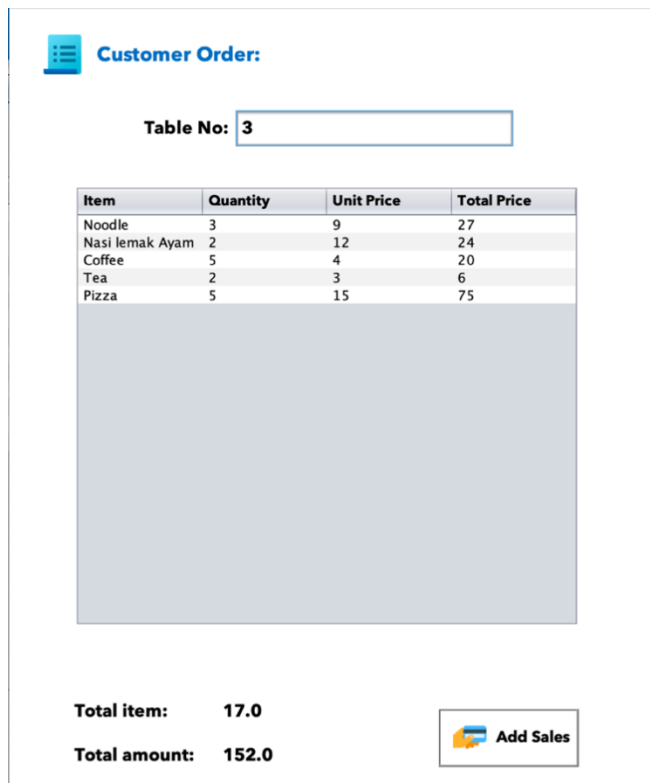
Fig.13.Re-stock Ingredient Form

Fig. 14 shows the menu list sub tab where admin can save, update, and delete the restaurant menu item. Admin can just simply type the food name and its price in the text field provided.
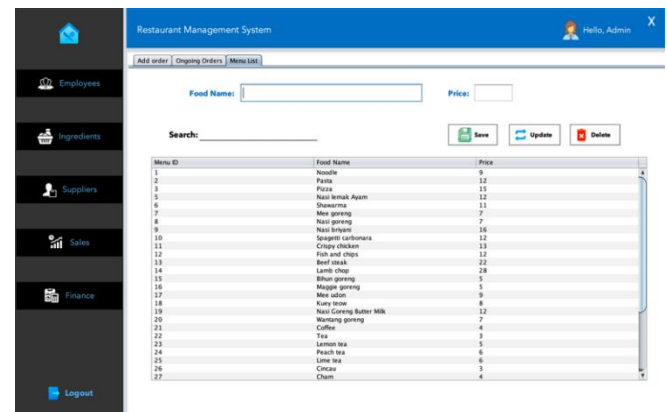
Fig.14. Menu List Subtab

The invoice subtab in Fig. 15 is used to list out all the sales of the restaurant. It includes the sale ID, total item, total amount, and date. Admin can delete the sales when there is a mistake by key in the sale ID or clicking on the list and press on the delete button. The total of sales item and amount will be automatically calculated. The income statement subtab in Fig. 16 is where admin can see the total income, expenses, employees, and suppliers of the restaurant. The profit or loss will be automatically calculated based on the sales and purchase of the restaurant. Lastly, admin can logout from the system by clicking on the logout button provided at the bottom left of the system.
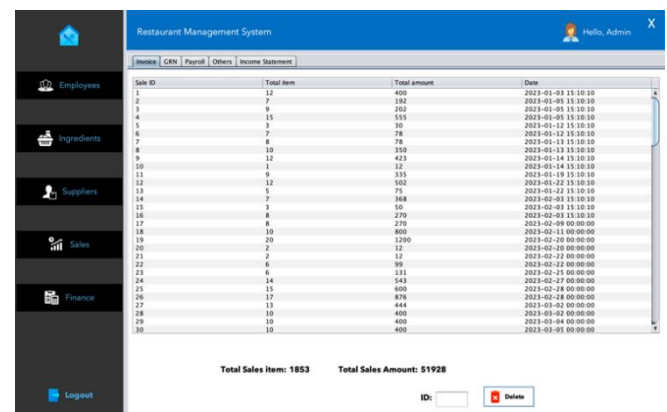
Fig. 15. Invoice Subtab

Fig.16.Income Statement Subtab

Based on the survey taken from 30 respondents of restaurant managers or owners, RMS has received positive feedback where all of them agree with the efficiency of using RMS and would like to use the system for their restaurant. Thus, this project can bring the solution to restaurant owners who wants to take the benefits of using RMS. Fig. 17 shows the positive feedback from users.
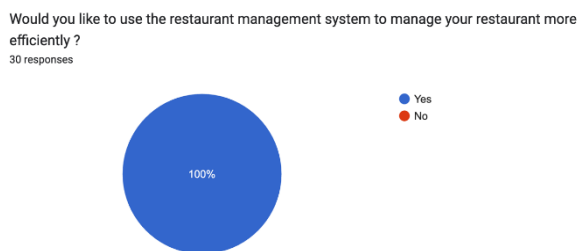
Fig. 17. User Acceptance Test Question on Willingness to Use the RMS

6. Conclusion

The proposed restaurant management system is developed to address the current issues in the existing system and reduce the manual process for restaurant daily operations. The traditional system that nowadays is still implemented in many restaurants is time-wasting, demands a high cost, and often cause human error which as a result cannot promote restaurant to perform in the best efficient way. Thus, the proposed restaurant management system is time-saving and error-free compared to the manual system. It aims to fulfill all the needs of the restaurant owners by developing a desktop application restaurant management system which able to ease the operational management of a restaurant,

increase management efficiency, minimize labor and resource costs, and enhance the performance of restaurant management. The system development is based on the SDLC methodology in a determined time and scope. User acceptance test is performed and has received positive feedback from users. Overall, the system is targeted to bring benefits to most restaurants and create efficiency by developing an integrated system that includes several functional features. This system can be enhanced and improved for future development by adding more functional features.

References

1. Alt, R. (2021). Digital transformation in the restaurant industry: Current developments and implications. Journal of smart tourism, 1(1), 69-74..
2. K. Kuligowski, "Small Business Guide to a Restaurant Management System," 1 July 2022.
3. Macha, S. (2022). Management System for a Restaurant..
4. Piyatissa, W. B. A. C. (2021). Web Based Restaurant Management System (Doctoral dissertation).
5. Kocaman, E. M. (2021). Operational effects of using restaurant management system: An assessment according to business features. International Journal of Gastronomy and Food Science, 25, 100408.
6. Karne, P. (2022). Management System for a Restaurant.
7. Wanyama, I., Arinaitwe, A., & Adraako, F. (2019). Web-based restaurant management system (Doctoral dissertation, Makerere University).

Authors Introduction

Gabriela Maria Ancilla

She has completed her B. Sc. (Hons) Business Information Systems, UCSI University, Malaysia.

Dr. Heshalini Rajagopal

She received her PhD and Master's degree from the Department of Electrical Engineering, University of Malaya, Malaysia in 2021 and 2016, respectively. Her research interest includes image processing, artificial intelligence and machine learning.

Dr. Ismail Ahmed Al-Qasem Al-Hadi

He received the Ph.D. degree in intelligent computing from Universiti Putra Malaysia (UPM), in 2017. His research interests include recommendation systems, machine learning, and deep learning approaches.

Dr. Renuka Devi Rajagopal

Dr. R. Renuka Devi is an Associate Professor in the School of Computer Science and Engineering, VIT Chennai, India. Her research interests include Cyber-Physical Systems, Block Chain Technology, Data Mining and Machine learning in the field of Health Care.

Dr. Norrima Mokhtar

Norrima Mokhtar was appointed as a lecturer to serve the Department of Electrical Engineering, University of Malaya immediately after graduating with her Master of Engineering. As part of her career development, she received SLAB/SLAI scholarship to attain her Ph.D. in 2012. She is now serving Department of Electrical Engineering, University of Malaya.

Face Recognition based on Attendance System

Koh Pei Cong

Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia

Heshalini Rajagopal

Department of Electrical and Electronics Engineering, MILA University, 71800 Nilai, Negeri Sembilan, Malaysia

Ghassan Saleh Hussein Al-Dharhani

Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia

Norrima Mokthar

Department of Electrical Engineering, Faculty of Engineering, University of Malaya, Malaysia

Dhanush Rachaveti

School of Electronics Engineering, Vellore Institute of Technology, Chennai, India
E-mail: heshalini@ucsiuniversity.edu.my, heshalini@gmail.com
www.ucsiuniversity.edu.my

Abstract

The use of face recognition technology for attendance tracking has grown popular in recent years. Hence, the main goal of this project is to produce an accurate, fast, and robust face recognition based on an attendance system. The system detects the user's unique features, understand the identity of the user through face recognition technology, and thus records the attendance from the face recognition dataset to the user that matches the user, in an attempt to help the user automatically check in with real-time attendance date. The system includes features such as face detection, face recognition, distance estimation, and attendance recording.

Keywords: Face Recognition, Attendance, Convolutional Neural Network

1. Introduction

A face recognition based on attendance system uses machine learning techniques to identify individuals and track their attendance automatically [1]. It captures an image of the person's face, processes it to extract unique features, and compares it with the face recognition datasets to determine the known faces [2]. If the known face is detected, the system updates his or her attendance record. This technology is commonly used in workplace environments, educational institutions, and other organizations to streamline the attendance-taking process and reduce manual errors [3].

In the 1980s and 1990s, researchers started to use linear algebra for face recognition. This approach was based on the idea that faces could be represented as mathematical objects and compared using linear algebraic operations. The first linear algebra-based face recognition algorithm was introduced by Turk and Pentland in 1991 [4]. They proposed a method called "Eigenface" which used eigenvectors and eigenvalues to represent the face space. This method involved transforming the images into a lower-dimensional space where faces could be represented as points in space. This approach was later improved upon with the introduction of "Fisherface" in 1997 by Belhumeur, Hespanha, and Kriegman [5]. This method used linear discriminant analysis to transform the face space into a space where faces could be separated more easily. This made it possible to use face recognition in a variety of applications, including security and surveillance, border control, and airport security [5].

In the 2000s and 2010s, the development of mobile devices and the widespread availability of the internet paved the way for cloud-based face recognition systems, which allowed for easy access and use from anywhere in the world. This also paved the way for the development of mobile face recognition applications, which could be used on smartphones and other mobile devices. In education, face recognition systems were used for attendance tracking and to enhance student safety [4], [6].

However, many conditions can affect the face recognition attendance system, such as the camera angles, the lighting conditions of the environment, the quality of the image, and various other factors. In addition, the images will get low contrast and low brightness if they are captured under low light conditions [1]. There are several algorithms used by experts to improve face recognition attendance systems, such as the Eigenface algorithm, Principle Component Analysis (PCA), Fisherface algorithm, and Convolutional Neural Networks (CNNs) [5], [7], [8]. The choice of algorithm will depend on the specific requirements of the application. However, CNN is widely used in face recognition attendance systems due to their high accuracy and ability to learn complex features from large amounts of data.

Therefore, this project uses a built-in camera to capture video from the camera, presents automatic attendance tracking, develops a Graphical User Interface(GUI), prevents system warnings and wrong execution from unknown faces, and recognizes faces with high accuracy and fast speed. The expected outcomes aim to ensure that the system can recognize faces under the camera angles, the lighting conditions of the environment, and the quality of the image.

2. Prior Work

This section reviews the methods which were used in those studies to implement face recognition technology and the corresponding to their results.

Shuhui and Xiaochen (2022) compared different SOTA deep face recognition models to extract features from a face image even in the presence of masks. By using data augmentation, the masked face recognition system became more robust to variations in the face images. The result proved that an appropriate margin aided the model's establishment of flexible embedding [9].

Winarno et. al (2019) proposed a combination of CNNs and PCA to identify and verify the attendance of individuals. CNN was used to extract features from a face image while PCA was used to reduce the dimensionality of the data and reduce computational costs. The result showed that the combination of CNNs and PCA can lead to improved performance in face recognition compared to using PCA alone [10].

Cahyono et. al (2020) designed a face recognition system for automating the process of tracking employee attendance. The linear SVM was used to classify the faces of employees in real-time as they entered or left the workplace. K-fold cross-validation was used to evaluate the performance and accuracy of a model. As a result, the combination of FaceNet and SVM achieved 100% accuracy [2].

Qu et. al (2018) compared algorithms, such as PCA, LBP, and CNN algorithms, for handling recognition speed and accuracy. PCA was suitable for fast recognition but with reduced accuracy, while CNN was suitable for high-accuracy recognition but at the cost of computational resources. LBP was a good compromise between the two, offering a balance between recognition speed and accuracy. It was possible to implement a CNN on FPGA due to its high performance and low power consumption [7].

Early work by Ilyas et. al (2019) was concerned with the use of the Viola-Jones algorithm to detect faces in images. Histogram Equalization was used in image processing to enhance the contrast of an image. The result showed that the recognition method using the CMU PIE face database has a higher recognition rate than the Extended Yale B face database [11].

KB and J (2020) designed a real-time face recognition system using CNNs. Viola-Jones algorithm was used to capture real-time input images from a camera. Then, the AT&T database was used to detect faces in each image and crop the image to include only the face. It was possible to simplify the feature extraction process by having all images of equal size [12].

Rahouma and Mahfouz (2021) implemented a face recognition system that leveraged the capabilities of the API mobile vision for face detection. API Google's mobile vision was used to detect faces in the pre-processed images while Pearson correlation was used to compare two face representations and determine their similarity. As a result, the Pearson correlation spent less time but also produced a higher recognition rate than other recognition methods [13].

Sarwar et. al (2021) developed an LBPH-based face recognition system for visually impaired people. Haar Cascade Classifier worked as a pre-processing step to detect faces in an image or video frame. The use of the Euclidean distance helped to determine the similarity between two faces and decide whether they belong to the same person or not. The result showed that LBPH gained higher average accuracy than the Fisherface algorithm [14].

Wang et. al (2020) compared algorithms, such as SGD, RMSprop, and Adam for evaluating the accuracy of the face recognition system. ReLU allowed the network to

model complex features in the input images, such as the shapes of eyes, noses, and mouths. The result proved that CNNs have significant practical application potential when used for face recognition [15].

Early work by Wandzik et. al (2018) was concerned with the use of an SVM classifier to distinguish between real faces and morphed faces. SVMs were able to handle high-dimensional data and can work well with noisy data. The Euclidean distance was used to compare the features of two faces and determined if they were significantly different. The result showed that the performance of CNN-based approaches dramatically increased [16].

3. Method

The face recognition based on attendance system goes through five stages:

(1) Set up the hardware: This involves installing the cameras and other necessary hardware components to capture the images and process them in real-time.

(2) Collect and prepare the training data: A large dataset of face images is collected and processed to extract the features needed for training the face recognition model.

(3) Train the face recognition model: The extracted features are used to train a deep Convolutional Neural Network (CNN) or other machine learning model to recognize faces.

(4) Integrate face recognition system with attendance system: The face recognition model is integrated with the attendance system so that the attendance of individuals can be recorded and tracked in real-time.

(5) Test and evaluate the system: The system is tested on a variety of real-world data to evaluate its accuracy and performance. Any essential modifications are performed to raise the system's efficiency.

Fig. 1 represents two actors: the user and the admin. A user can enter his or her personal information, such as name, id, and course code. This information will be stored in the system. Besides that, users can take face images, train a model, view the result, or play an audio file. If the username matches the result, the system marks user attendance automatically. Users can exit from the program after using the system. On the other hand, an admin can check user information, view the attendance records, and download an attendance list.
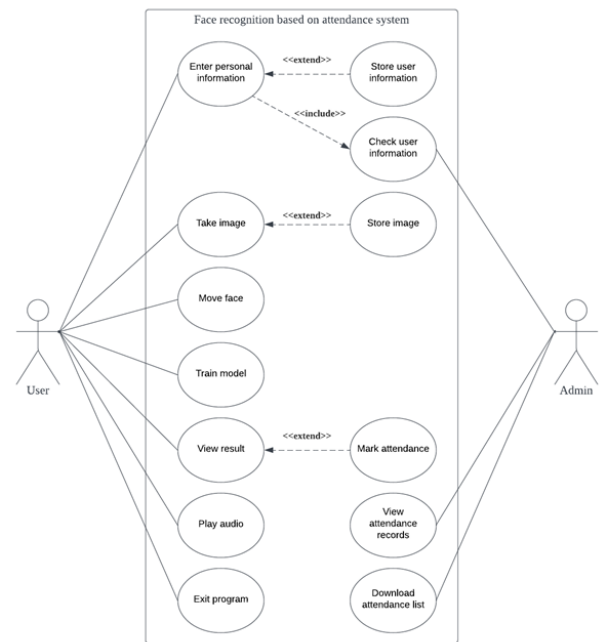
Fig. 1. Use Case Diagram

4. Result and discussion

There is an interface design of face recognition based on attendance system, as shown in Fig. 2. It is designed to be simple to use and user-friendly for end users of all ranges.
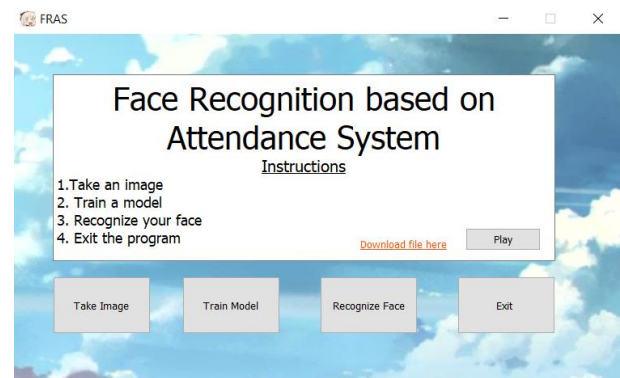
Fig. 2. Graphical User Interface (GUI) using QT Designer

The system includes six main functions, such as:

1. Play button: The user can click the play button to play an audio file. This audio file introduces the instructions should follow to complete all the processes.

2. Take image button: The user can click the take image button to register their faces. This button will bring the user to another window, as shown in Fig.3.
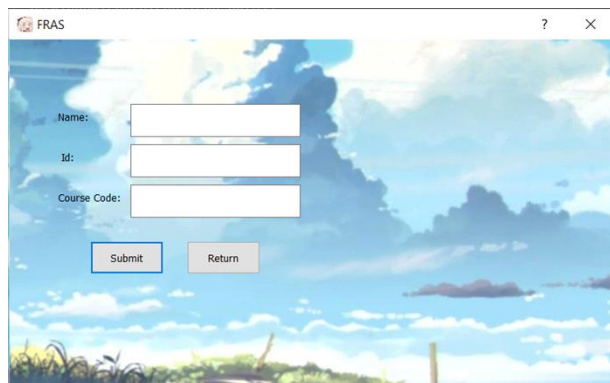
Fig. 3: Take image window

3. Train model button: The user can click the train model button to train a CNN model. The training process runs on the back-end interpreter. Please give a few moments of patience until the window pops up a success message.

4. Recognize face button: The user can click the recognize face button to recognize his or her face. The result shows the name, id, and the distance between an individual and the built-in camera. The system will automatically record the name, attendance date, and time.

5. Download link: The user can click the download file here to select the file to download and select the destination folder to store the file. The system will automatically download files in the destination folder.

6. The user can click the exit button to exit from the program immediately.

The system converts the images to grayscale, equalizes the histogram, and normalizes pixel values between 0 and 1., as shown in Fig. 4.

Fig. 4. Preprocessing images

Next, the system preprocesses the data, creates the model, applies image augmentation, and starts the training loop, as shown in Fig. 5.

Fig. 5. Train CNN model

After the CNN model is trained, the model is evaluated on the test set to find the test score and test accuracy, as shown in Fig. 6.
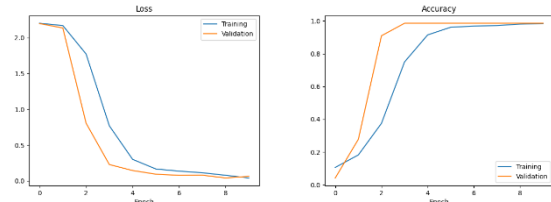
Fig. 6. Test Score(Loss) and Test Accuracy

Then, the system uses focal length calculation to estimate the distance of the recognized face from the camera, as shown in Fig.7.

Fig. 7. Focal Length Calculation

After that, the system predicts images from each detected face in the frame, as shown in Fig. 8.

Fig. 8. Predict Image

Lastly, the system updates the CSV file every time a face is recognized, as shown in Fig. 9.

Fig. 9. Attendance records

The testing and evaluation phases are crucial for the developed system. The system is tested on a variety of real-world data to evaluate its accuracy and performance. Any essential modifications are performed to raise the system's efficiency. This project covers unit testing, functional testing, and user acceptance testing.

Unit testing involves testing individual components of the system to ensure they function correctly. In this project, the unit testing helps to check the images are captured successfully, as shown in Fig. 10. Besides that, this testing helps to check the model has been trained successfully, as shown in Fig. 11. Moreover, it helps to check the face recognition done successfully, as shown in Fig. 12.
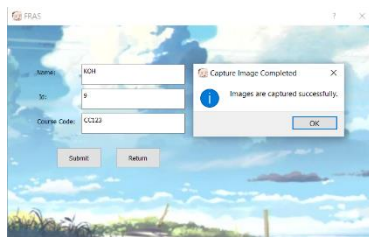
Fig. 10. Capture Image Completed
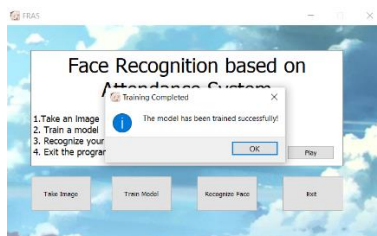
Fig. 11. Training Completed
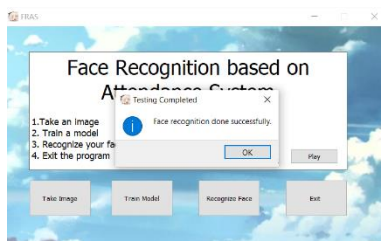
Fig.12. Testing Completed

Functional testing involves testing the system as a whole to ensure that it meets the specified functional requirements and performs correctly. In this project, the system recognizes known faces and unknown faces, as shown in Fig. 13. Moreover, the system downloads an attendance file in the destination folder.
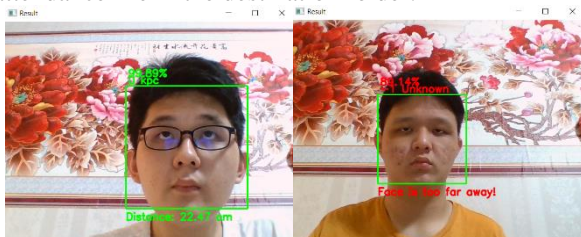
Fig. 13: Known face and Unknown face

5. Conclusion

In conclusion, face recognition based on attendance system provides a functional solution for tracking attendance automatically. The system includes features such as face detection, face recognition, distance estimation, and attendance recording. It exploits a pre-trained Convolutional Neural Network (CNN) model for face recognition. Moreover, the system can recognize faces with high accuracy and fast speed. Furthermore, the system can accurately recognize known faces and unknown faces. In addition, the system stores attendance records in a CSV file, making it easy to manage and track their attendance. Besides that, the system works on a graphical user interface to interact with end users easily.

References

1. K. Alhanaee, M. Alhammadi, N. Almenhali, and M. Shatnawi, "Face recognition smart attendance system using deep transfer learning," in *Procedia Computer Science*, 2021, vol. 192, pp. 4093–4102. doi: 10.1016/j.procs.2021.09.184.

2. F. Cahyono, W. Wirawan, and R. Fuad Rachmadi, "Face recognition system using facenet algorithm for employee presence," in *4th International Conference on Vocational Education and Training, ICOVET 2020*, Sep. 2020, pp. 57–62. doi: 10.1109/ICOVET50258.2020.9229888.

3. H. Yang and X. Han, "Face recognition attendance system based on real-time video processing," IEEE Access, vol. 8, pp. 159143–159150, 2020, doi: 10.1109/ACCESS.2020.3007205.

4. I. Adjabi, A. Ouahabi, A. Benzaoui, and A. Taleb-Ahmed, "Past, present, and future of face recognition: A review," Electronics (Switzerland), vol. 9, no. 8. MDPI AG, pp. 1–53, Aug. 01, 2020. doi: 10.3390/electronics9081188.

5. T. Schenkel, O. Ringhage, A. Márki, and J. Bae, "A Comparative Study Of Facial Recognition Techniques With focus on low computational power Bachelor Degree Project in Information Technology Basic level 30 ECTS Spring term 2019."

6. V. Kumar, G. Jacek, M. Zurada, B. Raman, and G. R. Gangadharan Editors, "Studies in Computational Intelligence 885 Modern Approaches in Machine Learning and Cognitive Science: A Walkthrough Latest Trends in AI."

7. X. Qu, T. Wei, C. Peng, and P. Du, "A Fast Face Recognition System Based on Deep Learning," in *Proceedings - 2018 11th International Symposium on Computational Intelligence and Design, ISCID 2018*, Jul. 2018, vol. 1, pp. 289–292. doi: 10.1109/ISCID.2018.00072.

8. H. Zhi and S. Liu, "Face recognition based on genetic algorithm," *J Vis Commun Image Represent*, vol. 58, pp. 495–502, Jan. 2019, doi: 10.1016/j.jvcir.2018.12.012.

9. W. Shuhui and M. Xiaochen, "Dual-Proxy Modeling for Masked Face Recognition," in *Procedia Computer Science*, 2022, vol. 208, pp. 145–151. doi: 10.1016/j.procs.2022.10.022.

10. Winarno, E., Al Amin, I. H., Februariyanti, H., Adi, P. W., Hadikurniawati, W., & Anwar, M. T. (2019, December). Attendance system based on face recognition

system using cnn-pca method and real-time camera. In 2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI) (pp. 301-304). IEEE.

11. Ilyas, B. R., Mohammed, B., Khaled, M., & Miloud, K. (2019, November). Enhanced face recognition system based on deep CNN. In 2019 6th International Conference on Image and Signal Processing and their Applications (ISPA) (pp. 1-6). IEEE.

12. K. B. Pranav and J. Manikandan, "Design and Evaluation of a Real-Time Face Recognition System using Convolutional Neural Networks," in *Procedia Computer Science*, 2020, vol. 171, pp. 1651–1659. doi: 10.1016/j.procs.2020.04.177.

13. K. H. Rahouma and A. Z. Mahfouz, "Design and Implementation of a Face Recognition System Based on API mobile vision and Normalized Features of Still Images," in *Procedia Computer Science*, 2021, vol. 194, pp. 32–44. doi: 10.1016/j.procs.2021.10.057.

14. M. G. Sarwar, A. Dey, and A. Das, "Developing a LBPH-based Face Recognition System for Visually Impaired People," in *2021 1st International Conference on Artificial Intelligence and Data Analytics, CAIDA 2021*, Apr. 2021, pp. 286–289. doi: 10.1109/CAIDA51941.2021.9425275.

15. D. Wang, H. Yu, D. Wang, and G. Li, "Face recognition system based on CNN," in *Proceedings - 2020 International Conference on Computer Information and Big Data Applications, CIBDA 2020*, Apr. 2020, pp. 470–473. doi: 10.1109/CIBDA50819.2020.00111.

16. Wandzik, L., Kaeding, G., & Garcia, R. V. (2018, September). Morphing detection using a general-purpose face recognition system. In 2018 26th European Signal Processing Conference (EUSIPCO) (pp. 1012-1016). IEEE.

Authors Introduction

Koh Pei Cong

He has completed his Bachelor of Science (Hons) in Computing, UCSI University, Malaysia.

Dr. Heshalini Rajagopal

She received her PhD and Master's degree from the Department of Electrical Engineering, University of Malaya, Malaysia in 2021 and 2016, respectively. Her research interest includes image processing, artificial intelligence and machine learning.

Dr. Ghassan Saleh Hussein Al-Dharhani

He is currently an Assistant Professor with the Institute of Computer Science and Digital Innovation, UCSI University, Malaysia. His research interests include Artificial Intelligence, Data Mining and Knowledge Discovery, Machine Learning, Data Science & Analysis

Dr. Norrima Mokhtar

Norrima Mokhtar was appointed as a lecturer to serve the Department of Electrical Engineering, University of Malaya immediately after graduating with her Master of Engineering. She is now serving Department of Electrical Engineering, University of Malaya.

Dr. Dhanush R.

Dr. Dhanush R. is an Associate Professor in the School of Electronics Engineering, VIT Chennai, India. His research interests include Neuroscience, Bio mechanics, Motor learning, Motor Control.

U-Reserve: Development of a Facility Reservation System for UCSI University

Esther Chong Jun Lynn, Neesha Jothi, Ismail Ahmed Al-Qasem Al-Hadi
Institute of Computer Science and Digital Innovation, UCSI University, Kuala Lumpur, Malaysia

Yeo Sook Fern
Faculty of Business, Multimedia University, Jalan Ayer Keroh Lama, Melaka 75450,
Malaysia Department of Business Administration, Daffodil International University, Dhaka 1207, Bangladesh
Email: neesha@ucsiuniversity.edu.my

Abstract

In today's technology-driven era, there is an increasing emphasis on efficient and user-friendly methods to elevate user satisfaction. UCSI University, Kuala Lumpur, currently relies on a manual facility reservation process, necessitating either in-person or email submissions. This outdated approach frequently leads to double bookings and underutilization of resources. Thus, this project endeavours to create a web-based facility reservation system called U-Reserve, employing the prototype methodology. U-Reserve was crafted using HTML, CSS, JavaScript, and core Java (JSP and Servlet), while its database is administered through phpMyAdmin. Additionally, U-Reserve places emphasis on incorporating preferred colour tones and font types of the target users, aiming to enhance user experience and satisfaction. Through U-Reserve, the reservation request and cancellation processes are streamlined, and a data-driven dashboard facilitates informed decision-making regarding facility management.

Keywords: facility reservation system, web-based system, colour tone, design principle, data-driven decision-making, user permission management

1. Introduction

UCSI University's ongoing expansion has posed challenges in keeping facilities' information up-to-date, leading to students and staff being unaware of available resources. The Group Logistics Management Office (GLMO) depends on a manual reservation process, requiring students to email GLMO admin for facility availability confirmation, leading to a three-day response time. Subsequent booking procedures exacerbate inefficiencies, risking overlapping reservations and increasing the administrative workload.

Additionally, the manual confirmation process for Block B's discussion rooms at the library counter exacerbates these challenges, leading to time wastage and increased workload for library admins. This inefficiency significantly impacts the overall experience and satisfaction of students, staff, and lecturers at UCSI University, potentially impeding academic performance and participation in extracurricular activities [1], [2].

Therefore, this project introduces U-Reserve, a web-based facility reservation system aiming to address the issues associated with the current reservation process at UCSI University. The system has been developed using HTML, CSS, JavaScript, and core Java (JSP and Servlet) to establish its functionality. For database management, phpMyAdmin is used alongside the MySQL database management system. The primary objectives include streamlining the reservation process, enhancing security measures, and enabling data-driven resource optimization, which will benefit students, staff, lecturers, and administrators at UCSI University.

The remainder of this paper includes sections on system comparison, literature review, development methodology, conclusion, and future work, offering a comprehensive overview of the development process, system features, and potential areas for further improvement.

2. System Comparison

Comparison among four (4) existing facility reservation systems—Bookeo [3], MRBS [4], Reservator [5], and RoomBooker [6] highlights commonalities and distinctive features, detailed in Table 1. These systems share fundamental functionalities, enabling users to make reservations, check reservation history, and granting administrators slot management capabilities. These essential features eliminate manual processes, prevent double bookings, and ensure efficient slot control, forming the foundation for U-Reserve. Table 1 illustrates that Bookeo, MRBS, and Reservator offer reporting tools for analyzing reservation trends. U-Reserve utilizes graphs for visualizing data, aiding in decision-making. Bookeo, MRBS, and RoomBooker send email notifications to requesters—an absent feature in Reservator. U-Reserve ensures prompt updates for users upon request approval or rejection. Furthermore, MRBS and Reservator display booked slots, whereas Bookeo and RoomBooker exhibit available slots only. U-Reserve

Esther Chong Jun Lynn, Neesha Jothi, Ismail Ahmed Al-Qasem Al-Hadi

combines both, providing a comprehensive Home Page table for clearer planning. Additionally, RoomBooker features admin reservation management and user-friendly slot searches—attributes adopted by U-Reserve for enhanced control and convenience.

Table 1. Features Comparison Between Different Facility Reservation Systems and The Proposed System.

Features	Bookeo	MRBS	Reservator	RoomBooker	U-Reserve
Make reservation.	√	√	√	√	√
Reservation history checking.	√	√	√	√	√
Slot management.	√	√	√	√	√
Report and analytics.	√	√	√	×	√
Email notifications.	√	√	×	√	√
Table showing booked slot.	×	√	√	×	√
Reservation management.	×	×	×	√	√
Search for available facilities or slots.	×	×	×	√	√
Batch reservation.	×	√	×	×	√
Track reservation details.	×	×	×	×	√
View announcements and contact information.	×	×	×	×	√

Moreover, MRBS introduces batch reservation options for weekly, monthly, and yearly bookings. U-Reserve customizes this feature for weekly reservations, offering users flexible date selection to minimize unnecessary bookings. Furthermore, U-Reserve empowers administrators to track reservation details, ensuring seamless user contact. Users can also access department announcements and contact information, streamlining communication.

In summary, U-Reserve integrates functionalities from Bookeo, MRBS, Reservator, and RoomBooker, while incorporating unique elements to provide a comprehensive solution.

3. Literature Review

3.1. *Security*

User Permissions: Managing user permissions is a critical and multifaceted aspect within web-based systems, serving as the cornerstone for security, privacy protection, web personalization, and access control. As emphasized in the referenced scholarly works, adopting Role-Based Access Control (RBAC) models and permission-based mechanisms equips web-based systems with essential tools for handling user permissions. The benefits are extensive, encompassing the mitigation of security risks, safeguarding user privacy, tailoring user experiences, and facilitating secure, controlled resource access. In the evolving landscape of web-based systems' influence on our digital lives, effective user permission management remains a pivotal component. It ensures system integrity, operational efficacy, and fosters trust among both users and administrators [7], [8], [9], [10].

3.1.1 Data-Driven Resource Optimization

Resource Management: Web-based systems have revolutionized resource management, particularly in domains like facility management. These systems provide a comprehensive platform for data analysis, decision support, accessibility, and collaboration. In the context of facility management, they enable quantitative and conceptual analysis, data sharing, remote accessibility, and secure data management. Ultimately, adopting web-based systems for resource management empowers businesses to optimize their operations, reduce costs, enhance safety, and improve overall efficiency. By leveraging the capabilities of these systems, organizations can ensure that their resource management processes are both effective and efficient, ultimately contributing to their success and competitiveness [11], [12].

Decision Making: Web-based systems have revolutionized the decision-making landscape for businesses by offering tools and frameworks that empower organizations to make informed, data-driven decisions. Decision support systems, business intelligence, analytics, interactive decision aids, and provenance-based strategies all contribute to more effective and efficient decision-making processes. In the era of big data and data science, web-based systems are essential for organizations aiming to stay competitive and improve business performance. The effective harnessing of data is a key driver of success, with web-based systems playing a pivotal role in achieving this goal [13], [14], [15].

3.1.2 Design Principle

System Design: It's the core foundation for web-based systems, addressing resource allocation, reliability, security, usability, and learning system design. Its critical role ensures reliable performance, secures sensitive data, enhances usability, and promotes effective web-based learning. Good system design is essential for seamless operation and user satisfaction in digital environments, be it large-scale enterprises or local host websites [16], [17].

Colour Theory: Color tone is a potent tool in web design, influencing users' perceptions, emotions, and behaviours when used intentionally. With a well-matched colour scheme that complements content, designers craft visually appealing designs and guide users effectively. This approach fosters trust, improves user experience,

and drives engagement and satisfaction in web-based systems [18], [19], [20].

Visual Appeal: Crucial in web system design, the layout establishes flow and hierarchy, guiding users effectively. Applying visual principles and aesthetics, designers craft engaging, user-friendly systems, contributing to satisfaction and success [21], [22].

3.2 Applications of Security, Data-Driven Resource Optimization, and Design Principle on U-Reserve

Table 2. Applications of Security, Data-Driven Resource Optimization, and Design Principle on U-Reserve.

Area	Applications
Security.	● A role-based security system prevents errors and safeguards system integrity.
Data-Driven Resource Optimization.	● Decision support design for administrators: comparing requests and enhancing efficiency. ● Dashboard presenting visualised data for data-driven facility management. ● Downloadable Excel report providing valuable insights into facility utilisation.
Design Principle.	● User preferences on colour tones and font types are gathered for UI design. ● Logical page structure, consistent navigation bar, and organised content ensure coherence. ● Font size variations create a clear visual hierarchy on pages.

4. Development Methodology

The project was completed using the prototype methodology, as illustrated in Fig. 1. This approach was selected due to its capacity to involve users in identifying missing functionalities at an early stage [23].
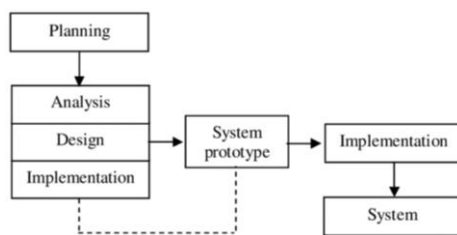
Fig. 1 Prototype Methodology [24]

4.3 Planning

Location of Study: All research activities were conducted at UCSI University, Kuala Lumpur, Malaysia, and remotely via Microsoft Teams using an Internet connection.

Target Population: Students, lecturers, staff, and admins of UCSI University, Kuala Lumpur.

Sampling Methods: Purposeful sampling entails preselecting participants based on specific criteria to ensure diverse insights into various system aspects. Snowball sampling recruits additional participants meeting specific criteria, allowing respondents to refer others.

Data Collection: Primary data collection involves identifying functional and non-functional requirements, as well as preferences for colour tone and font type for U-Reserve. Secondary data from sources such as journals guides the acquisition of user requirements, supporting project insights and results.

4.4 Analysis

Functional and Non-functional Requirements Analysis: A total of seventy-five (75) users actively participated in this survey via Google Forms, providing valuable insights and feedback.

Functional requirements include checking availability, making reservation requests, checking request status, request history, and tracking reservation details. All highlighted functional requirements identified by respondents are integrated to enhance user satisfaction. Priority will be given to enabling users to check availability, make reservation requests, and check request status, as these received the highest percentages: 93.33%, 86.67%, and 74.67%, respectively.

Moving to non-functional requirements, feedback on the ease of use of a facility reservation system shows that 73.33% prefer high user-friendliness. U-Reserve prioritizes easy navigation and efficient user interactions in line with this preference. The next query assesses reliability and security, with 53.33% rating them as very important, guiding U-Reserve's robust security design. The importance of 24/7 availability is evident, with 65.33% considering it crucial for user convenience and efficiency. Lastly, 70.67% emphasize scalability's importance for a reservation system's effectiveness. Despite its local host development, U-Reserve needs to meet availability and scalability as vital non-functional requirements when adopted by UCSI University.

User Interface (UI) Preferences Analysis: The feedback received from these one hundred (100) users via Google Forms provided insights on their preferred colour tones and font types for U-Reserve.

Four (4) commonly used colour tones - blue, red-violet, green, and red - were selected for the user interface (UI) of U-Reserve. Respondents could choose their preferred colour tone for each questionnaire question. Blue, linked to familiarity and professionalism, emerges as the top choice across multiple criteria: visual appeal (32%), suitability (39%), calming effect (42%), professionalism (34%), modernity (30%), and accessibility (47%). This preference for blue aligns with research highlighting its universal appeal, trustworthiness, and readability in digital environments. Notably, cool colours like blue and green consistently dominate preferences (60.83%), reflecting users' preference for calming and trustworthy associations. Emphasizing the importance of contrast in U-Reserve's UI, achieved through thoughtful design considerations, is crucial for readability and enhancing the overall user experience [25].

Additionally, respondents were asked to select a font type—Josefin Sans + Dosis, Sans-Serif, Verdana, Arial, or Georgia—based on suitability, readability, and aesthetic appeal for the UI design. Verdana emerges as the top choice, favoured by 38% for suitability, 48% for readability, and 49% for aesthetic appeal. Sans-Serif and the combination of Josefin Sans + Dosis also receive varied support. The consistent preference for Verdana aligns with research emphasizing its screen-centric design, readability advantages, and aesthetic appeal, making it the optimal font type for a user-friendly experience in a facility reservation system [26].

4.5 Design

System Design:

The following figures show the main pages of U-Reserve.

Fig. 2 displays a login page with ID number and password fields. Regardless of their role, users must provide the correct ID number and password for authentication.

Fig. 2 Login Page of U-Reserve

After a successful login, the home page, as depicted in Fig. 3, is displayed, showcasing a navigation bar, announcement slides, contact information, and an availability table for facility reservations.

Fig. 3 Home Page of U-Reserve

Fig. 4 Block B Facility Checking and Reservation Page of U-Reserve

Users can check availability by viewing the cell's colour and make reservations by clicking on the cells. Additionally, users can navigate to specific facility reservation pages for Block B, as illustrated in Fig. 4, or Block C, as illustrated in Fig. 5.
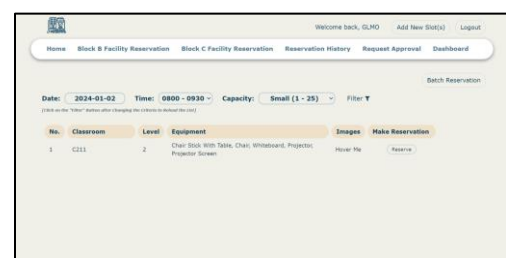
Fig. 5 Block C Facility Checking and Reservation Page of U-Reserve

After selecting a facility, users are redirected to the reservation page, as depicted in Fig. 6 which is used to fill in the details. Additionally, users can review their reservation request list and check the status from the reservation history page, as shown in Fig. 7. Finally,

admins can approve or reject reservation requests on the reservation approval and checking page, as illustrated in Fig. 8.

Fig. 6 Reservation Page of U-Reserve

Fig. 7 Reservation History Page of U-Reserve

Fig. 8 Reservation Approval and Checking Page of U-Reserve

Use Case Diagram:

The U-Reserve use case diagram delineates three user roles: students and staff, lecturers, and administrators (Group Logistics Management Office (GLMO) and library admin), illustrated in Fig. 9.
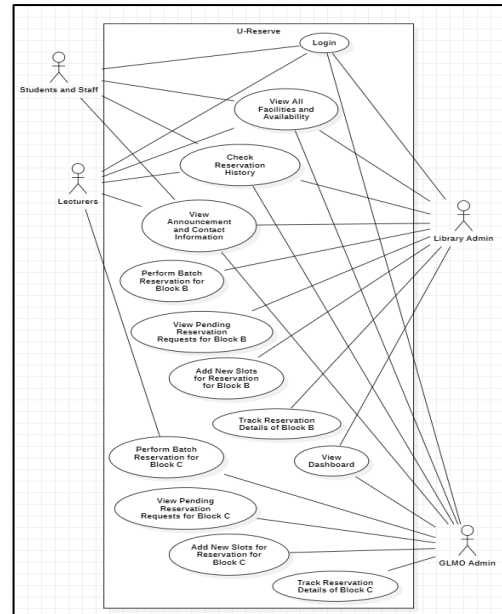
Fig. 9 Use Case Diagram for U-Reserve

Students and staff possess various capabilities, including logging in, password management, accessing announcements and contact details, checking facility availability, making reservations, reviewing reservation history, and canceling reservations. Additionally, lecturers enjoy extended privileges, allowing batch reservations specifically for Block C classrooms. GLMO and library admins have the highest permissions, sharing fundamental functionalities with students and staff while also conducting batch reservations, adding new slots, and overseeing reservation requests. GLMO admins can access a dashboard and download usage reports for Block C classrooms, whereas library admins focus on managing discussion rooms in Block B, significantly contributing to efficient oversight and management.

Entity Relationship (ER) Diagram:
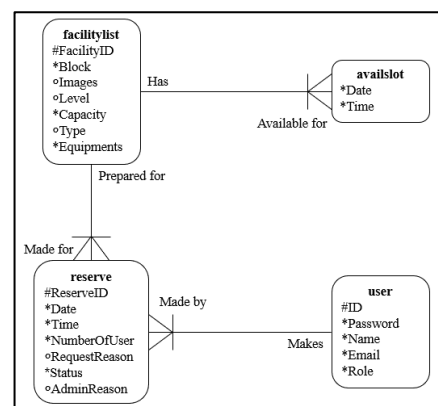
Fig. 10 ER Diagram for U-Reserve

In the ER Diagram displayed in Fig. 10, four entities are represented: "facilitylist", "availslot", "reserve", and "user". The "facilitylist" entity stores data about all facilities, utilizing "FacilityID" as its unique identifier, and includes mandatory attributes like "Block", "Capacity", "Equipments," as well as optional attributes such as "Images", "Level", and "Type". The "availslot" entity comprises mandatory attributes "Date" and "Time." Identified by "ReserveID," the "reserve" entity contains mandatory attributes like "Date," "Time," "NumberOfUser," and "status," along with optional attributes like "RequestReason" and "AdminReason". The "user" entity, marked by "ID," encompasses mandatory attributes like "Password", "Name", "Email", and "Role". Relationships are established using UID bars: a facility can have multiple available slots, each uniquely linked to a facility through the "FacilityID" attribute. Similarly, a facility can be prepared for multiple reservations, each associated with a facility through the same "FacilityID". Finally, a user can make multiple reservations, with each reservation linked to a user through the "ReserveID" attribute.

Activity Diagram:
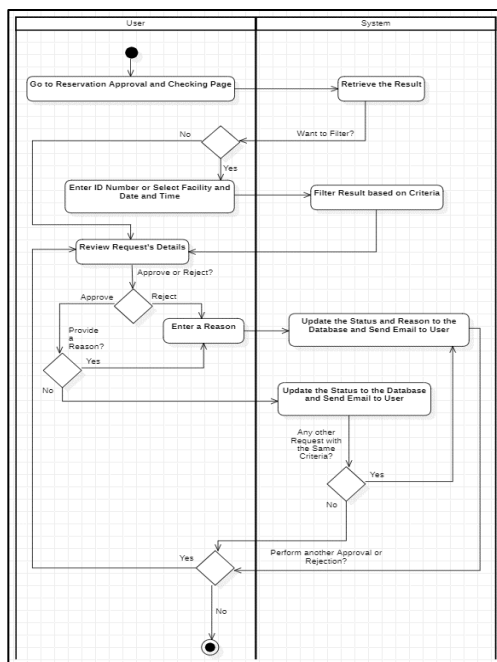
Fig. 11 Activity Diagram for U-Reserve – Approve or Reject Reservation Requests

In Fig. 11, users handle reservation requests by accessing the reservation approval and checking page through the "Request Approval" link. Here, they can browse and filter pending reservation requests by ID number, facility, date, and time. Users review request details and decide to approve or reject. Giving a reason for rejection is necessary. The system updates the status and notifies requesters via email. Approved requests prompt the system to check for similar ones,

automatically rejecting them, updating the status, and sending notifications.
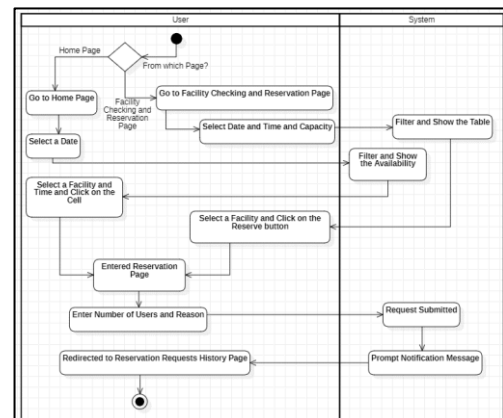
Fig. 12 Activity Diagram for U-Reserve – Make Reservation

As per Fig. 12, users reserve through two methods: via the home page's availability table or Block B/Block C facility checking and reservation page. To reserve from the home page, users scroll, pick a date from the date picker, check facility availability, select time and facility, and click a cell. This redirects to the reservation page where they input user count and reason, get confirmation, and are sent to the reservation history. Alternatively, from the facility checking and reservation page, users choose Block B/C, set date, time, capacity, select a facility, and click "Reserve". This proceeds to the reservation page following the same steps.
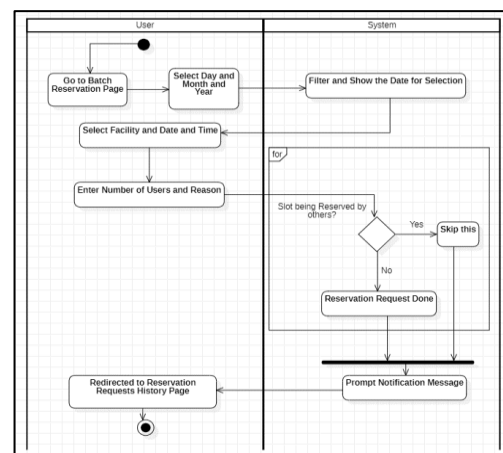
Fig. 13 Activity Diagram for U-Reserve – Batch Reservation

In Fig. 13, users perform batch reservations on the batch reservation page. Here, users select the day of the week, month, and year. The system retrieves corresponding dates from the database and presents them. Users then pick the facility, date, and time, enter user count and reasons, and submit the request. The system checks for existing reservations for the selected slots. If available,

the system adds the new reservation request to the database; otherwise, it skips it. This cycle repeats for each selected date until completion. Users are notified of the submitted requests and redirected to the reservation requests history page.
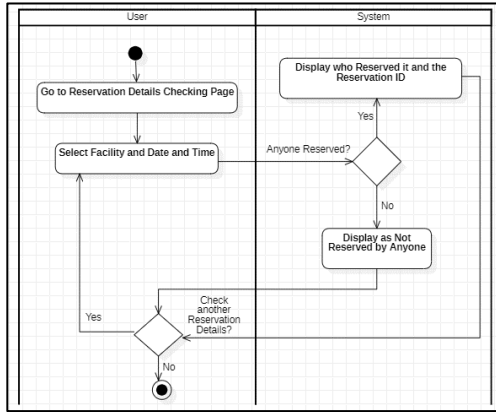
Fig. 14   Activity Diagram for U-Reserve – Track Reservation Details

In Fig. 14, users can monitor reservations for a particular facility, date, and time via the reservation details checking page. They choose the desired parameters, and the system scans the database for reservations that match the criteria. If found, the system displays the Reservation ID and requester's ID number; otherwise, it indicates no reservations.
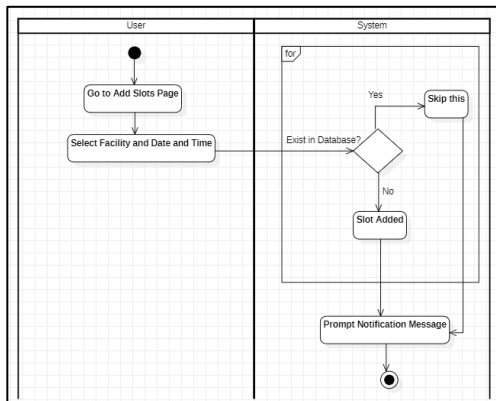
Fig. 15   Acytivit Diagram for U-Reserve – Add New Slots

In Fig. 15, admins can include reservation slots by clicking the "Add New Slot(s)" button to access the Add Slots page. On this page, they select the facility, date, and time, and submit the details to the system. For each slot, the system verifies its existence in the database; if already present, the system skips that slot. Upon completion of the loop, the system informs admins about the number of submitted slots.
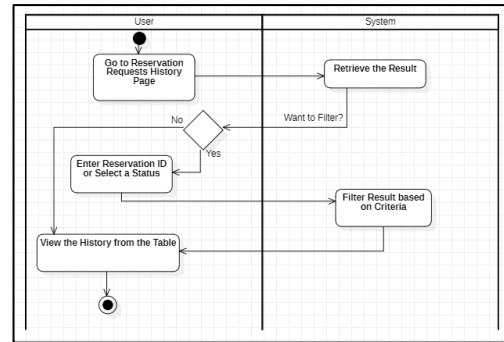
Fig. 16   Activity Diagram for U-Reserve – Check Reservation History

Users can access their reservation request history by clicking the "Reservation History" link on the navigation bar, as depicted in Fig. 16. Upon reaching the reservation requests history page, users can observe all their reservation requests. They have the option to input the Reservation ID, select a status, or both to refine the history. Subsequently, the system connects to the database, filters the results based on the provided criteria, and displays them to the users.
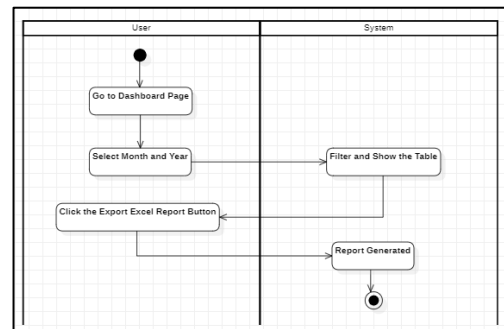
Fig. 17 Activity Diagram for U-Reserve – Export Excel Report

In Fig. 17, users can generate an Excel format utilization report from the dashboard page by clicking the "Dashboard" link on the navigation bar. They choose the preferred month and year, and the system filters and displays the corresponding data. Clicking the "Export Excel Report" button prompts the system to generate and download the report.

4.6 Development and Implementation

Table 3 displays the tabulated data from the user acceptance test (UAT) forms, utilising a Likert scale ranging from 1 for strongly disagree to 6 for strongly agree, as suggested by research [27], as shown in Fig. 18. The analysis used the arithmetic mean technique, following the methodology of [28]. A comprehensive evaluation of different sections can be accomplished by calculating the mean rating for all items within each section. A mean rating exceeding five (5) indicates user agreement. In this three-scale representation of

Esther Chong Jun Lynn, Neesha Jothi, Ismail Ahmed Al-Qasem Al-Hadi

agreement, encompassing 4 - Agree, 5 - Somewhat Agree, and 6 - Strongly Agree, five (5) serves as the midpoint.

Fig. 18 Likert Scale [27]

Table 3 User Acceptance Test (UAT) Result.

Item	1	2	3	4	5	6	Mean Rating
User Interface Design [27]							
Appropriate colour tone.	0	1	1	5	10	13	5.1000
Appropriate font type.	0	0	0	3	9	18	5.5000
Appropriate font size.	0	0	0	3	7	20	5.5667
Navigation is easy.	0	0	0	3	7	20	5.5667
Navigation is clear and concise.	0	0	0	1	8	21	5.6667
Perceived Usefulness [29]							
Using the system in my job/study increases my productivity.	0	0	0	4	9	17	5.4333
Using the system enhances my effectiveness in my job/study.	0	0	0	1	10	19	5.6000
I find the system to be useful in my job/study.	0	0	0	2	11	17	5.5000
Perceived Ease of Use [29]							
My interaction with the system is clear and understandable.	0	0	0	1	13	16	5.5000
Interacting with the system does not require a lot of my mental effort.	0	0	0	1	10	19	5.6000
I find the system to be easy to use.	0	0	0	1	8	21	5.6667
I find it easy to get the system to do what I want to do.	0	0	0	1	10	19	5.6000
Output Quality [29]							
I have no problem with the quality of the system's output.	0	0	0	3	11	16	5.4333
I rate the results from the system to be excellent.	0	0	0	2	9	19	5.5667
Result Demonstrability [29]							
The results of using the system are apparent to me.	0	0	0	1	11	18	5.5667
I would have no difficulty explaining why using the system may be beneficial.	0	0	0	4	11	15	5.3667
Behaviour Intention [29]							
Assuming I had access to the system, I intend to use it.	0	0	0	0	10	20	5.6667

In addition, Table 3 presents survey results indicating a favourable user perception of U-Reserve across various aspects, such as UI design, effectiveness, ease of use, output quality, result demonstrability, and behavioural intention. The system meets user expectations and preferences satisfactorily. However, potential areas for improvement have been identified, and several suggestions are currently under consideration. In essence, U-Reserve stands as a well-developed system proficient in efficiently managing facility reservations for UCSI University while assisting university staff in handling lecturers' timetables.

5. Conclusion and Future Work

In conclusion, the development of U-Reserve, UCSI University's web-based facility reservation system, followed a prototype methodology involving planning, analysis, design, and implementation. This approach successfully engaged users early in the process, incorporating feedback from user acceptance tests to modify and enhance the system. Notable contributions include streamlining the reservation process for students, staff, lecturers, and administrators, reducing workloads, and improving facility management through data-driven analyses.

U-Reserve meets user preferences for colour tone and font type, fulfilling functional and non-functional requirements. It stands out from existing systems by integrating their features while introducing unique elements like tracking reservation details and offering announcement and contact information. Moreover, the comprehensive project report lays the groundwork for implementing the facility reservation system at UCSI University, signifying a substantial achievement for the institution and its users.

Future improvement suggestions for the facility reservation system at UCSI University include redesigning for formality and consistency, migrating to a cloud-based or hosted server, implementing approval periods for reservation requests, integrating with central databases and the IIS, expanding reservation options, and enhancing the password reset mechanism. These enhancements aim to boost the system's functionality, security, and user experience, aligning it with the evolving needs of the university community.

Ismail Ahmed Al-Qasem Al-Hadi, for their invaluable support, guidance, and mentorship throughout this project. Their patience and expertise have played a crucial role in shaping the direction and scope of this study, and I am truly thankful for their contributions. I also want to express my sincere gratitude to the lecturers, administrators, staff, and students of UCSI University for their cooperation in conducting the user acceptance tests (UATs) and providing valuable feedback. Special thanks to Assistant Professor Ts. Dr. Ghassan Saleh for helping me in finding more lecturers and staff to participate in my UATs. Additionally, I acknowledge the support of my friends, who assisted me in collecting details and images of the facilities.

References

1. Testech, "Effects of Poor Facilities Management - Integrated Facilities Management," Testech Group, Aug. 07, 2022.
2. M. Mosley, "Why is Word of Mouth Marketing so Important?," Social Media Today, Jun. 25, 2017.
3. B. Hogan, "Bookeo Review, Pricing & Features: Popular Online Appointment Booking Software for Small Businesses," SoftwarePundit, Feb. 26, 2018.
4. J. Beranek, "MRBS: Introduction," mrbs.sourceforge.io, 2008.
5. Joelmertanen et al., "Reservator," GitHub, Oct. 01, 2023.
6. Queen Margaret University, "Space and Timetabling BOOKING A ROOM VER. 1.1 Edinburgh College | MIS BOOKING A ROOM GUIDE 2 Contents," 2019. Accessed: Nov. 19, 2023.
7. J. Chen and T. Zhang, "Research and Implementation of Role-Based Access Control Model Based on Partition Number," 2009 Second International Symposium on Computational Intelligence and Design, vol. 2, Jan. 2009.
8. B. Mishra et al., "Privacy Protection Framework for Android," IEEE Access, vol. 10, pp. 7973–7988, 2022.
9. T. Zhang and W. Tan, "Role-based dynamic access control for Web services," 2010 International Conference on Computer Application and System Modeling (ICCASM 2010), vol. 4, Oct. 2010.
10. E. K. Zavadskas, A. Kaklauskas, M. Gikys, and N. Lepkova, "A multiple criteria decision support web-based system for facilities management," International Journal of Internet and Enterprise Management, vol. 2, no. 1, p. 30, 2004.
11. S. Casadei, A. Pierleoni, and M. Bellezza, "Sustainability of Water Withdrawals in the Tiber River Basin (Central Italy)," Sustainability, vol. 10, no. 2, p. 485, Feb. 2018.
12. F. Al-Hawari, M. Al-Zu'bi, H. Barham, and W. Sararhah, "The GJU Website Development Process and Best Practices," Journal of Cases on Information Technology, vol. 23, no. 1, pp. 21–48, Jan. 2021.
13. Das, "Effectiveness of Web-Based Decision Making to Deliver Decision-Support Information to Business Analyst using a 'Thin-Client,'" Technoarete Transactions on Advances in Data Science and Analytics, vol. 1, no. 1, pp. 1–7, Feb. 2022.
14. F. Provost and T. Fawcett, "Data Science and its Relationship to Big Data and Data-Driven Decision Making," Big Data, vol. 1, no. 1, pp. 51–59, Mar. 2020.
15. L. Hurbean, F. Militaru, M. Muntean, and D. Danaiata, "The Impact of Business Intelligence and Analytics Adoption on Decision Making Effectiveness and Managerial Work Performance," Scientific Annals of Economics and Business, vol. 70, no. SI, pp. 43–54, Feb. 2023.
16. K.-S. Joo and J.-W. Woo, "Development of Object-Oriented Analysis and Design Methodology for Secure Web Applications," International Journal of Security and Its Applications, vol. 8, no. 1, pp. 71–80, Jan. 2014.
17. Y.-M. Kim, "Factors Affecting University Library Website Design," Information Technology and Libraries, vol. 30, no. 3, Sep. 2011.
18. L. Khrouf and A. Frikha, "Web-surfers' conative reactions to the website's dominant hue: mental imagery's role," Internet Research, vol. 26, no. 5, pp. 1249–1268, Oct. 2016.
19. X. Zhang, W. Hu, and Q. Xiao, "Influences of Medical Crowdfunding Website Design Features on Trust and Intention to Donate: Controlled Laboratory Experiment," Journal of Medical Internet Research, vol. 23, no. 5, p. e25554, May 2021.
20. T.-C. Tung and H.-Y. Chen, "Integrating Conjoint Analysis with TOPSIS Algorithm to the Visual Effect of Icon Design Based on Multiple Users' Image Perceptions," Eurasia Journal of Mathematics, Science and Technology Education, vol. 13, no. 3, pp. 1025–1040, Dec. 2016.
21. D. Cyr, M. Head, H. Larios, and B. Pan, "Exploring Human Images in Website Design: A Multi-Method Approach," MIS Quarterly, vol. 33, no. 3, p. 539, 2009.
22. S. Djamasbi, M. Siegel, and T. Tullis, "Generation Y, web design, and eye tracking," International Journal of Human-Computer Studies, vol. 68, no. 5, pp. 307–323, May 2010.
23. K. Kirpitsas and T. P. Pachidis, "Evolution towards Hybrid Software Development Methods and Information Systems Audit Challenges," Software, vol. 1, no. 3, pp. 316–363, Sep. 2022.
24. H. M. A. Wahab, M. H. Hassan, N. Mohd, Z. Hanafi, and Hafizul, "WEB BASED INTELLIGENT APPOINTMENT SYSTEM," Jul. 2023.
25. R. Kimmons, "Color Theory in Experience Design," Learner and User Experience Research, pp. 103–125, 2020, Accessed: Nov. 14, 2023.
26. Banerjee and M. Bhattacharyya, "Selection of the optimum font type and size interface for on screen continuous reading by young adults: an ergonomic approach," Journal of Human Ergology, vol. 40, no. 1–2, pp. 47–62, Dec. 2011.
27. C. K. N. C. K. Mohd and F. Shahbodin, "Personalized Learning Environment: Alpha Testing, Beta Testing & User Acceptance Test," Procedia - Social and Behavioral Sciences, vol. 195, pp. 837–843, Jul. 2015.
28. M. Osman, M. Maghribi, M. Nurzaid, P. Zulfikri, M. Othman, and M. Shahrol, "Mobile cloud computing for m-learning application," vol. 10, pp. 1055–1068, May 2018.
29. Ozdemir and A. K. Kabakus, "User acceptance of cloud based hospital information system," typeset.io, vol. 6, no. 1, pp. 28–46, Mar. 2019.

Authors Introduction

Ms. Esther Chong Jun Lynn

She is currently a degree student at UCSI University, Kuala Lumpur, Malaysia, studying for a B.Sc. (Hons) in Business Information System.

Dr. Neesha Jothi

She received her PhD from the School of Computer Sciences, Universiti Sains Malaysia in 2020. She is currently an Assistant Professor in UCSI University, Malaysia. Her research interest areas are Data Mining in Healthcare and Health Informatics.

Dr. Ismail Ahmed Al-Qasem Al-Hadi

He is a highly qualified and experienced Assistant Professor who is currently working at UCSI University in Kuala Lumpur, Malaysia. He has a strong background in computer science, with a PhD in Intelligent Computing and a Master's degree in Computer Science. He also holds a Bachelor's degree in Comp Science.

Dr. Sook Fern Yeo

She is an Assistant Professor at the Faculty of Business, Multimedia University, Melaka. She is currently holding the position as of the Deputy Dean for Research & Industrial Collaborations. Her current research areas are service innovations, consumer behaviour, social media marketing, branding, supply chain management, healthcare management and tourism marketing..

The Smart Document Processing with Artificial Intelligence

Raenu Kolandaisamy, Heshalini Rajagopal
Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia

Indraah Kolandaisamy
School of Business Management, University Utara Malaysia 06010 Sintok, Kedah

Glaret Shirley Sinnappan
Tunku Abdul Rahman University of Management and Technology, Kuala Lumpur, Malaysia
E-mail: raenu@ucsiuniversity.edu.my

Abstract

This study focuses on the challenges and potential for Intelligent Document Processing (IDP) with Artificial Intelligence (AI) to manage unstructured data. A large amount of data in many different forms, such as information from the IoT, cybersecurity and more are produced during this modern Digital Age which has been widely distributed through a wide range of unformatted formats. IDP utilizes AI technologies like Machine Learning (ML), Natural Language Processing (NLP), and Computer Vision (CV), with the aim of converting unstructured data into structured, usable information. The IDP, are explored in the findings and discussion section that emphasizes its capability to automatically perform redundant tasks, reduce operating costs as well as improve employee productivity.

Keywords: Artificial Intelligence (AI), Natural Language Processing (NLP), Document Processing, Computer Vision (CV)

1. Introduction

In today's world, everything is connected to a data source, and digitization has made it possible to record everything. The current electronic era has an abundance of data available in various forms including the Internet of Things (IoT) data, cybersecurity data, smart city data, business data, smartphone data, social media data, health data, COVID-19 data, etc. This data can be structured, semi-structured, or unstructured [1].

With the emergence of new communication platforms and diverse applications such as social media, mobile applications, and digital marketing, the data delivered lacks a typical format or predefined schemas like the standard data and is unmanageable with the relational database models. Data is generated in various forms like text, audio, videos, emails, and images, which are categorized as unstructured data. Unfortunately, such data lacks structure and standardization, which leads to difficulty for most organizations to edit, search, and analyze [2].

Based on Forbes statistics, it reported that 95% of business organizations struggle to analyze unstructured data due to they do not have the required expertise to deal with unstructured data [2]. In a business company, 80% to 90% of data is in the form of unstructured format (Sukrutha, 2023). Business companies face significant challenges in effectively processing such unstructured data. In addition, the process of unstructured data analysis is always time-consuming and costly for any organization [2].

However, all organizations are required to extract valuable information from unstructured data to access functional information. Thus, automated analysis is crucial in helping the organization manage and access useful information in a structured manner. By automating the unstructured data kept in digital formats, organizations are allowed to quickly gain insight into their businesses, increase their competitive edge, improve productivity, and drive innovations. Artificial Intelligence-based (AI-based) technologies are critical for organizations that would like to adapt to automation solutions [2].

Intelligent Document Processing (IDP) is the conversion process of unstructured data such as email, images, and PDF documents to functional data. Intelligent Document Processing is one of the processes carried out by artificial intelligence (AI) and AI technologies are one of the best automation processes of the next generation. It involves Computer Vision (CV), Deep Learning (DL), Machine Learning (ML), etc. for data extraction [3].

The field of Artificial Intelligence (AI) is widely renowned as one of the most exciting and promising areas in the technology landscape. Analysts believe that AI will play a vital part in any business sector and industry. Besides that, the potential for value creation across the economy with machine learning (ML), a major subset of AI, is vast. It not only affects consumers, but it also transforms the way businesses operate aTsyapewtheoxlet . [4].

The increasing demand for the effective use of unstructured data necessitates the development of an AI-based unstructured document processing framework. This framework can aid organizations in extracting valuable insights from unstructured data automatically. Therefore, this research area is considered a significant and an important research topic [2].

Extracting insights from unstructured data can be challenging due to the lack of a predefined schema or format. Therefore, it is more challenging to analyze and process. Besides that, it also required the users to have a certain level of understanding to utilize the AI technologies and techniques. Thus, the main research objectives of this study are listed as below:

- To review the advantages and disadvantages of IDP with AI-based techniques for unstructured data.
- To understand the challenges with IDP to manage unstructured data.

Intelligent Document Processing (IDP) is a powerful technology that can transform unstructured and semi-structured data into a structured format. By leveraging advanced AI technologies like machine learning (ML) and natural language processing (NLP), IDP can capture, classify, and extract even the most complex and challenging data. This extracted data can then be validated and automatically entered into existing applications using Robotic Process Automation (RPA) technology [5].

According to [4], although advanced analytics and ML have been shown to provide a competitive advantage, many enterprises still struggle with their adoption. The lack of a clear AI strategy, talent, and data are major barriers, along with soft factors like leadership, decision- making, and company culture. According to

[6], it is shown that there are some potential issues and challenges with using machine learning, such as collecting sufficient and relevant training data. Besides that, another challenge is the development methods to automate document processing using machine learning, particularly for complex documents. Moreover, it requires efficient and effective technology and techniques to process and analyze fast streams of unstructured data. The complexity of unstructured data poses challenges for information extraction, especially for streaming data. Structural variation and differences in data formats are the critical obstacles associated with unstructured big data [7]. To fully leverage the IDP and gain access to valuable insights, organizations must go beyond simply knowing about it and actively incorporate it into their management and utilization of data.

Extracting insights from unstructured data can be challenging due to the lack of a predefined schema or format. Therefore, it is more challenging to analyze and process. Besides that, it also required the users to have a certain level of understanding to utilize the AI technologies and techniques. Thus, the main research objectives of this study are listed as below:

- To review the advantages and disadvantages of IDP with AI-based techniques for unstructured data.
- To understand the challenges with IDP to manage unstructured data.

2. Literature Review

Intelligent Document Processing (IDP) is one of the most popular technologies in this modern era with a high demand focused on the automation of the way humans work with different types of documents on a daily basis. Besides, Digital transformation is also one of the trending topics nowadays as Robotic Process Automation and other AI technologies such as ML work toward automating the ordinary operations performed by users. These AI technologies work together to process documents intelligently with limited human involvement [8].

The possibilities of electronically exchanging structured information continue to rise significantly. Nevertheless, a large amount of business processes continue to center on the sending, receiving, and processing of unstructured documents. According to Prasad (2020), only 20 % of data in enterprises is structured and the rest of it is unstructured data with most of it locked in documents. Companies with the adaptation of an automated document processing solution reported a reduction of 50% in their staffing and labor cost [9], while processing time is reduced to less than 20 % compared to manual processing time. It shows that

automation of business document processing is the key element for businesses to be more cost-effective and cost efficient.

Robotic Process Automation (RPA) refers to the automation of service tasks to perform repetitive human tasks precisely with the assistance of AI. The developer set the task instruction with the use of some form of screen recording and defining variables. Logging into applications, copying and pasting data, opening emails, filling forms are some of the examples of tasks. Different from the traditional automation method, RPA mainly emphasizes element identification instead of screen coordinates or XPath selections which results in better intelligent interaction with the user interface [9].

An alternative to the manual processing of documents is necessary to limit human involvement, minimize error rates and manage the time and cost efficiently and effectively. In this case, the research of automatic document processing has become one of the trending topics in recent years. RPA, also known as digital employee, has become a powerful tool to facilitate the automatic processing of business documents. RPA shows its effectiveness by enabling simple configuration of robots to handle high-volume and repetitive tasks automatically based on predefined rules which not only limits the human errors, but also improves the efficiency and cost reduction [10].

AI-based approaches are capable of extracting useful information from unstructured documents automatically. AI-based approaches are used in numerous applications such as invoice digitization, health record extraction, metadata extraction, machine translation, text summarization, insurance claims processing, and contract analysis [2].

Natural Language Processing (NLP) is one of the subfields of Artificial Intelligence that focuses on empowering computers to understand and process human languages, bringing them closer to human-like language understanding. Named Entity Recognition (NER) is an important component of NLP systems, which involves the process of automatic identification of named entities in a given text or document. Named entities refer to real-world objects such as people, location, date, and time. Recognizing and extracting real named entities is vital for various research areas, including Question Answering and Summarization Systems, Information Extraction, Machine Learning, Semantic Web Search, Video Annotation, and etc [11] According to [12], significant improvements have been made in extracting information from free text using NLP techniques based on AI. NLP uses syntactic rules to automatically scan and extract named entities like names, locations, organizations, dates, and invoice numbers from unstructured text using NER. The main steps in NER involve text pre-processing, extracting

NER features, then training and classification. In text pre-processing, the given text will be formatted to be understandable for machine learning algorithms. Then, features are extracted to create a numerical representation because the NLP model could not work on the text data. Lastly, the NER model will classify and categorize words and phrases. [2].

Computer Vision (CV) is also a component of AI that enables systems to identify and extract useful information from scanned document images and other visual inputs. The goal of computer vision is to extract meaningful information from images with the development of technology that enables computers or systems to understand and recognize digital images [13]. Machine Learning (ML) is a type of data-driven AI that provides the ability to learn about a system without explicit programming [14]. AI Algorithms and Machine Learning (ML) approaches have been effectively used in real-world scenarios which include the digital services, industry and commerce throughout the years. ML acts as the teacher of the machines which "teach" the efficient data handling skills by simulating the learning concept of rational beings. It can be achieved with AI algorithms that reflect diverse rational paradigms including connectionist, genetics, statistics and probabilities, based on cases, etc. The integration of AI algorithms and the ML approach provide a possibility for the exploration and extraction of information to perform tasks such as classification, association, optimization, grouping, prediction and pattern identification. [9]

Since a large amount of enterprise data these days is usually unstructured data which has no predefined schema or format, it is more complicated and challenging for users to collect, process and analyze the data without proper structure [2]. In this case, Intelligent Document Processing with AI is proven to be essential and necessary for the processing of unstructured documents.

However, there are several challenges and limitations that might affect the performance of which need to be addressed. There are major challenges regarding the use of Unstructured Data. Inaccurate text data is added by OCR text extraction in the form of noisy data sometimes. This is a key problem related to data. Unstructured data originating from multiple sources are not standard and have different formats of data. The lack of a sufficient supply of data and insufficient quality is also another challenge in data processing with unstructured data. Moreover, it is difficult to obtain information in a language that is complex such as Arabic, without the creation of dictionaries. It is difficult with information extraction methods to describe the semantic and contextual relationships among named entities in this ambiguous language.

Raenu Kolandaisamy, Heshalini Rajagopal, Indraah Kolandaisamy, Glaret Shirley Sinnappan

Another challenge is posed by domain specific entities. For example, the biomedical datasets are different from any other dataset in terms of domain specific entities [2]. As the business process requires handling and processing different types of unstructured documents from the clients or suppliers such as invoices, ID cards and application forms, it poses a challenge as these documents differ in type, form and layout which require classification based on their type and nature. Therefore, the AI technologies used for document processing should be able to extract the relevant fields from the particular documents with the application of either template-based or template-free approach. Moreover, AI technologies should be able to process and improve the quality of scanned documents, which is often submitted with low quality scanners or mobile devices. The collection of relevant target data from documents is complicated when it comes to managing multi-pages of unstructured documents which consist of tables with data that span across different pages [2].

3. Method

A case study method will be used for this study due to various data and information gathering have been undertaken to study this topic. A case study is a research methodology that assists in phenomenon exploration in a certain context via various data sources. Besides, it embarks on the exploration of phenomena from different perspectives to uncover various aspects. In the case study, the study is conducted in real-time within its naturally occurring context, with the consideration that context will influence the outcome [15].

According to [16], case study research is a qualitative approach in which the investigator explores real-life, recent cases over time involving detailed data collection from numerous sources of information, and reports a case description. To present the complexity of the issue being examined, various methodological approaches will be used. A case study will include a case study database, a clear evidence chain, and multiple sources of data from research. The depth and richness of the case study description allow readers to have a better understanding of the case and to determine the applicability of findings beyond the setting. The evidence sources for case studies include questionnaires, interviews, documentation, direct observations, participant observations, archival records etc. Therefore, case studies involve trustworthy information and act as a reliable research method since case study research is an in-depth and intensive research methodology [15].

4. Result and discussion

A questionnaire is conducted to review and understand about this study and research objectives. The questionnaire results from the respondents are shown in the Fig. 1.

Fig. 1 Results from the questionnaire

RPA is the automation technology that strives to automate human tasks including business processes by utilizing software robots that interact with systems through their user interface. It can help to increase efficiency, decrease costs and efforts that humans spend on repetitive tasks [17]. One of the advantages of RPA is that it can automate numerous processes smoothly and decrease operational costs significantly due to it can handle repetitive tasks and time-saving as well as resource-saving. Besides that, prior programming knowledge is not a must to set up and implement RPA, leading to accessibility to employees who are without any programming knowledge and experience. Moreover, RPA supports most of the regular business processes with error-free automation and reduces the need for human intervention. Furthermore, software tools are scalable and do not get exhausted [2]. RPA can handle tedious and redundant tasks, thus employees can concentrate on more complicated and valuable tasks that contribute more value to the company. The employees

with more meaningful tasks can invest more time to develop new skills to be more qualified for their job [18].

On the other hand, RPA might be limited when it comes to reading unstructured documents with various layouts and formats. The documents need to be in the same format for software bots to read accurately. In addition, even minor changes in the automation application also require reconfiguration of the RPA software bot [2]. Moreover, it can be disruptive and less effective if without the appropriate tools although it seems like a good idea for a company to shift to a more advanced business system [19].

Even though it has been proven that a competitive advantage can be achieved with the use of advanced analytics and ML, enterprises nowadays are still facing significant challenges in the adoption of ML due to lack of clear strategy for AI, insufficient time and limited data available. Soft aspects which include leadership, decision making, and company culture are also some of the challenges which affect the adoption of ML [4].

It is possible to achieve adoption through the creation of a tangible business value. However, as part of product or process innovation, the main challenge is to find the right purpose and to deliver the technology in a compliant manner. In these circumstances, for the development of business value it is imperative to have deep experience in processes and a good understanding of commercial requirements. Efficiency can be achieved by automation and an increased pace of implementation of business processes, saving costs at the same time. Effectiveness refers to higher business process quality which consists of fewer errors and compliant process execution [4].

ML is expected to address the clear business needs and integrate them into the business process so that it can deliver efficiency, preferably in the form of an end-to-end automation, simplifying the access of business users to technology. In this respect, 'everyday AI' which is capable of functioning without long data preprocessing and training activities is expected. In order to achieve this simplification, it is possible to integrate the technology into the process where the data is stored and improve the user experience [4].

In addition, access to data and its quality are two of the key elements that contribute to the success of ML. Firstly, it is necessary to consider the aspects of data gathering, access and transfer, anonymisation, annotation, curation, secure storage and deletion. Nevertheless, strict observance is required in the areas of legal frameworks and general regulations. When more data is created or collected, AI must be trained and evolved over time. The algorithms are expected to learn directly from user behavior, which is then used to internalize their knowledge. ML is currently still

considered as an assistance system which coexists with humans, and processes the data provided by humans. This is expected to happen automatically, in a way that human correction is used to improve algorithms and that this knowledge is not forgotten by the machine. This can be described as continuous learning. As a result, users will be able to anticipate increasing accuracy over time, build trust and depend on the results.

Furthermore, culture also plays an essential role in the introduction of machine learning and therefore needs to be adequately climate governed by innovation principles, resources, time, etc. The leadership's mindset and employees' willingness to accept change are the main reasons for this. It will have to be derived from and integrated with the proper strategy. It seems that enterprises with pragmatic and hands on attitudes tend to be more agile. They are aware of business needs, have well defined use cases, and provide the necessary environment for implementation such as fewer organizational obstacles, empowerment of staff, no fear of losing control or jobs. In fact, there seems to be a reluctance and much slower adoption by customers who are deeply entrenched in bureaucracy and hierarchies of managers without an explicit innovation culture [4].

5. Conclusion

The study focused on the potential and challenges of Information Document Processing with Artificial Intelligence techniques to manage unstructured data or documents. The prior study conducted by the researchers on IDP with AI techniques indicated that there are numerous benefits and challenges for AI-based techniques, but it is still a very popular and trending technique that most companies are adapting and utilising. According to the findings of this study, AI techniques such as RPA and ML have various benefits to companies, such as increased efficiency, reduced costs, and support for error-free automation of business processes. In addition, they are also facing some challenges, such as a lack of AI strategy, insufficient data, and culture, in the adoption process for the IDP with AI techniques. However, it can be overcome through value creation, integration, data considerations, and fostering an innovation-driven culture. IDP with AI techniques has been growing and increasingly important in each industry. Therefore, it is vital for companies to emphasise, adopt and utilise the current technologies to improve their business performance.

References

1. Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. SN computer science, 2(3), 160.

https://link.springer.com/article/10.1007/s42979-021-00592-x

2. Baviskar, D., Ahirrao, S., Potdar, V., & Kotecha, K. (2021). Efficient automated processing of the unstructured documents using artificial intelligence: A systematic literature review and future directions. IEEE Access, 9, 72894-72936. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9402739

3. Sukrutha, K. S., Harini, S., & Kusuma, M. V. (2023). A Study on Intelligent Document Processing Using AWS. International Journal for Multidisciplinary Research (IJFMR), 5(4), 1-5. https://www.ijfmr.com/papers/2023/4/4308.pdf

4. Janasz, T., Mortensen, P., Reisswig, C., Weller, T., Herrmann, M., Crnoja, I., & Höhne, J. (2021). Advancements in ML-enabled intelligent document processing and how to overcome adoption challenges in enterprises. Die Unternehmung, 75(3), 340-358. https://www.nomos-elibrary.de/10.5771/0042-059X-2021-3-340/advancements-in-ml-enabled-intelligent-document-processing-and-how-to-overcome-adoption-challenges-in-enterprises-jahrgang-75-2021-heft-3?page=1

5. Cutting, G. A., & Cutting-Decelle, A. F. (2021). Intelligent Document Processing--Methods and Tools in the real world. arXiv preprint arXiv:2112.14070. https://arxiv.org/ftp/arxiv/papers/2112/2112.14070.pdf

6. Sambetbayeva, M., Kuspanova, I., Yerimbetova, A., Serikbayeva, S., & Bauyrzhanova, S. (2022). Development of Intelligent Electronic Document Management System Model Based on Machine Learning Methods. Eastern-European Journal of Enterprise Technologies, 1(2), 115.https://www.researchgate.net/publication/359049293_Development_of_intelligent_electronic_document_management_system_model_based_on_machine_learning_methods

7. Adnan, K., & Akbar, R. (2019). Limitations of information extraction methods and techniques for heterogeneous unstructured big data. International Journal of Engineering Business Management, 11, 1847979019890771.https://journals.sagepub.com/doi/epub/10.1177/1847979019890771

8. Fernando, L. (2023). Intelligent Document Processing: A Guide for Building RPA Solutions (Edition 2023.1.). Notion Press.

9. Ribeiro, J., Lima, R., Eckhardt, T., & Paiva, S. (2021). Robotic process automation and artificial intelligence in industry 4.0–a literature review. Procedia Computer Science, 181, 51-58. https://www.sciencedirect.com/science/article/pii/S1877050921001393

10. Ling, X., Gao, M., & Wang, D. (2020). Intelligent document processing based on RPA and machine learning. In 2020 Chinese Automation Congress (CAC) (pp. 1349-1353).

11. Anandika, A., & Mishra, S. P. (2019). A study on machine learning approaches for named entity recognition. In 2019 International Conference on Applied Machine Learning (ICAML) (pp. 153-159). IEEE. https://ieeexplore.ieee.org/document/8989189

12. Macri, C. Z., Teoh, S. C., Bacchi, S., Tan, I., Casson, R., Sun, M. T., ... & Chan, W. (2023). A case study in applying artificial intelligence-based named entity recognition to develop an automated ophthalmic disease registry. Graefe's Archive for Clinical and Experimental Ophthalmology, 1-10. https://link.springer.com/article/10.1007/s00417-023-06190-2

13. Paneru, S., & Jeelani, I. (2021). Computer vision applications in construction: Current state, opportunities & challenges. Automation in Construction, 132, 103940. https://www.sciencedirect.com/science/article/pii/S0926580521003915

14. Bose, P., Srinivasan, S., Sleeman IV, W. C., Palta, J., Kapoor, R., & Ghosh, P. (2021). A survey on recent named entity recognition and relationship extraction techniques on clinical texts. Applied Sciences, 11(18), 8319. https://www.mdpi.com/2076-3417/11/18/8319

15. Rashid, Y., Rashid, A., Warraich, M. A., Sabir, S. S., & Waseem, A. (2019). Case study method: A step-by-step guide for business researchers. International journal of qualitative methods, 18, 1609406919862424. https://journals.sagepub.com/doi/full/10.1177/1609406919862424

16. Alpi, K. M., & Evans, J. J. (2019). Distinguishing case study as a research method from case reports as a publication type. Journal of the Medical Library Association: JMLA, 107(1),1. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6300237/

17. Santos, F., Pereira, R. and Vasconcelos, J.B. (2020), Toward robotic process automation implementation: an end-to-end perspective, Business Process Management Journal, 26 (2), 405-420. https://doi.org/10.1108/BPMJ-12-2018-0380

18. Costa, D., São Mamede, H., & Silva, M. M. D. (2022). Robotic Process Automation (RPA) adoption: a systematic literature review. Engineering Management in Production and Services, 14(2), 1-12. https://repositorioaberto.uab.pt/handle/10400.2/14058

19. Ansari, W. A., Diya, P., Patil, S., & Patil, S. (2019). A review on robotic process automation-the future of business organizations. In 2nd International conference on advances in science & technology (ICAST). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3372171

Authors Introduction

Dr. Raenu Kolandaisamy

He received his PhD from the Faculty of Computer Science & Information Technology, University Malaya in 2020. He is currently an Assistant Professor in UCSI University, Malaysia. His research interest areas are Wireless Networking, Security, VANET and IoT.

Dr. Heshalini Rajagopal

She received her PhD and Master's degree from the Department of Electrical Engineering, University of Malaya, Malaysia in 2021 and 2016, respectively. She received the B.E (Electrical) in 2013 Currently, she is an Assistant Professor in UCSI University, Kuala Lumpur, Malaysia. Her research interest includes image processing, artificial intelligence and machine learning.

Dr. Indraah Kolandaisamy

She is a senior lecturer in School of Business Management, College of Business, Universiti Utara Malaysia. Indraah A/P Kolandaisamy holds a doctorate in Management from Universiti Kebangsaan Malaysia in 2015. Her D.B.A work is on organizational citizenship behavior among public sector in Malaysia. She obtained her MSc (Management) and Bachelor in International Business Management (Hons) from Universiti Utara Malaysia respectively in 2007 and 2005.

Dr. Glaret Shirley Sinnappan

She holds the position of Assistant Professor in the Department of Information and Communication Technology at Tunku Abdul Rahman University College.

Digital Security Challenges Faced by Business Organizations

Raenu Kolandaisamy, Heshalini Rajagopal
Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia

Indraah Kolandaisamy
School of Business Management, University Utara Malaysia 06010 Sintok, Kedah

Glaret Shirley Sinnappan
Tunku Abdul Rahman University of Management and Technology, Kuala Lumpur, Malaysia
E-mail: raenu@ucsiuniversity.edu.my

Abstract

Over the decades, the forms of cyber-attacks have evolved from disruption level, cybercriminal, followed by cyber espionage and lastly threatening level. Digital security has played a significant role in protecting enterprises from any form of cyber-attacks, especially in today's era of digitalization. Aligning with the global effort in emerging the concept of Industrial Revolution 4.0 (IR4.0) in organizations, where sensitive data and confidential information can be accessed at a fingertip of an employee. This paper discusses the difficulties of implementing digital security solutions in an enterprise in terms of external potential cyber threats, internal cyber security roadblocks within the organization and how Covid-19 pandemic imposed challenges towards cyber security in the organizations.

Keywords: digital security, cyber security, cyber-attacks, cybercriminal, cyber espionage, business organizations

1. Introduction

The Information System has changed our lifestyle over the past decades among banking systems, education, transportation, business, tourism, shopping. The rise of internet consumption has dramatically increased by organizations and individuals with massive global international trade and changes of human interaction in the young generation. Digital economy can refer to integration of new business models, goods, services and markets established on a basic business infrastructure with digital technologies. Digital technologies can be defined as consolidation of communication, computing, information, and connectivity technologies that are driven by storage, software applications, sensors and bandwidth to the next generation of digital economy through cloud computing [1].

The rapid development of science and technology is far more than that such as the internet of things (IoT), artificial Intelligence (AI), augmented reality (AR), drones, additive manufacturing (3D printing), robotics and others. These latest technologies work from extracting data from sensor conditions of physical devices, spreading and storing it rapidly on cloud to analyse it in real time by using advanced big data analytics for generating useful information of integrating services, products and processes, and making decisive influence on existing business models. Meanwhile, approximately 90% of business managers in developed countries like the U.K. and U.S. forecast information technology and digital technologies contribute strategic values to the entire economy in the future. On the other aspect, information technologies will uncover a variety of new threats of cyber security challenges and risks when more companies foster innovation by increasingly popularizing new digital technologies, while security information technologies are changing and evolving to become a complicated threat (cyber security). Furthermore, information security cases might have caused a minor technical problem 10 or 15 years ago, but today advanced and intentional cyber attacks are challenging us that might cause large-scale events as a result of direct and indirect consequences affecting business's strategy and competitive advantage [2].

Browsing the internet, reading digital books, social networking, searching content, and online shopping are transforming our lives to better, efficient and productive

Raenu Kolandaisamy, Heshalini Rajagopal, Indraah Kolandaisamy, Glaret Shirley Sinnappan

ones. However, most of the netizens and organizations lack cyber security awareness and relevant knowledge of internet hazards. Individuals and business organisations have limited knowledge to protect their devices and servers from cyber hazards such as financial fraud, sensitive information theft, website spoofing, Internet of Things (IoT) hacking, malware, distributed denial of service attack (DDoS) as well as ransomware. Cyber hazards come from the work of hackers also known as "black hats" who committed cyber crime on their own or with an organized criminal group. Large-scale breach of cyber security has been rated by the World Economic Forum as one of the five most

critical threats confronting the world today where the scope of threat globally is spreading rapidly and cost an estimated US$6 trillion by 2021 [2]. Nevertheless, studies show that installing protective tools on the computers from factories does not maximize cyber security generally. Hence, this paper examines the challenges faced while implementing cyber security solutions into organizations' systems in various different perspectives.

2. Result and discussion

According to the Internet Crime Report published by the Cyber Division of the Federal Bureau of Investigation (FBI) in the year of 2020, losses exceeding USD 29 mil resulted from 2,474 ransomware attacks registered worldwide has been recorded [3]. The fact that cyber crimes against enterprises have increased tremendously over the years has shown that the digital security systems or practices within the organizations still remained susceptible and vulnerable towards cyber attacks in any form or even worse, could not keep up with the growing pace of these cyber threats. Thus, the roadblocks or difficulties to carry out the implementation of the cyber protection have to be discussed thoroughly in terms of external cyber hazards, internal cyber security challenges in organization and challenges of digital security during Covid-19 pandemic in order to improve, correct and make changes towards the current situation of cyber security support in the right direction.

2.1. External Potential Cyber Hazards

The Internet has improved the standard of living of IoT users, but every pillar has two sides. Everything can be done smoothly by one click as the same goes with cybercrime. That could so easily invade valuable information of individual and business organizations by

various cyber hazards such as malware, phishing, distributed denial-of-services attack (DDoS), man-in-the-middle attack (MMIT), ransomware and insider threats [4]. Every cyber attackers' intentions are different, but mostly demanding for money, financial information, personal information and even the business infrastructure. Cyber attacks target individuals and business organizations that have less cyber security awareness and protection tools because they always have key assets to be exploited by criminals. According to "PricewaterhouseCoopers survey", organizations that undergo cyber attack relevant matter vanished 2.1% of its value and lost more than 1.6 billion USD per case on average [2]. Understanding the motive of cyber attackers and hazards is essential for internet users to prevent financial loss and irreversible consequences. There are three aspects of motive to be proven as crime which are motive, means, and opportunity that can be treated as a conventional crime. Means refer to an attack tool or technical expertise to execute the attack, motive is the aims, intention, purpose or reason of the attacker selecting a target, opportunity can be seen as a timing, vulnerability and weakness in the network of the target [5].

Ransomware is a kind of malicious software that encrypts files of business organizations by demanding ransom in exchange for a decryption key [6]. It is very difficult to detect hijacked files and data before a cyber attacker initiates the attack. There are only 2 options for users to solve the problem which are to pay ransom in exchange for a decryption key to retrieve files and data or disconnect the internet and format the computer [4]. However, it is nearly impossible for internet-based business organizations to get rid of ransomware by formatting all data in digitized generation because it is the foundation and infrastructure of the company. In 2012, Reveton is one of the recognised ransomware trojans that ransom nearly $130,000 fines from victims who believed that they have been charged for participating in illicit online activities. Business organizations weren't too worried about the attack because it was affecting mostly individuals. By the end of 2016, an estimated $1 Billion was paid to the ransom hackers from Cryptorlocker and CryptoWall. It gave a huge profit trajectory to the hackers, something big known windows exploits were on the way from the National Security Agency (NSA). The worst and largest ransomware attack that ever happened before in 2017, "WannaCry" (Fig. 1) ransomware attacked more than 230,000 computers of various industries which were universities, government organization, hospital, bank, logistic, telecommunication companies, railways companies and business organizations across 150 countries [7]. According to research from "SonicWall Capture Labs", ransomware cases have risen by 20%,

with up to 121 million attacks recorded around the world in the first half of 2020 [8].

Working of ransomware (Fig. 2) will be introduced into 4 steps. First, the ransomware starts with an unsolicited email usually designed to trick the victim into clicking on an attachment or visiting a webpage. Second, ransomware leverages flaws in the computer's command-and-control server and forces it to download a public key for running ransomware code. Third, important data is encrypted on the system and demands a ransom payment using cryptocurrency by providing ransom fee instructions. Lastly, the decryption key will be given by the attacker once the ransom is paid [4].

Fig. 1 Activation of "WannaCry" on victim's computer
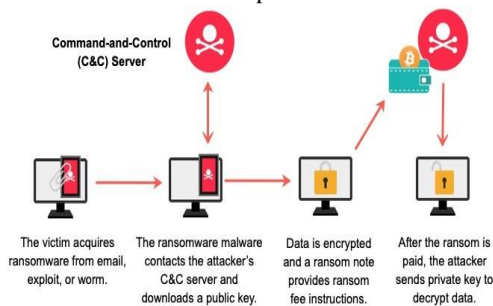
Fig. 2 Ransomware mechanism

Phishing attacks (Fig. 3) is a social engineering technique to bypass computer security and human defence by using social networks, emails, mobile apps, instant message, business email compromising or search engine poisoning to steal user's data, sensitive information, credit card number and code, as well as login credentials to make the victim perform an action. There are more than 30% of cybercrime caused by phishing attacks and cost at least 8 billion USD according to Verizon Data Breach Investigation (DBIR). There are some technical countermeasures to mitigate risk of phishing such as anti-phishing, email malware detection, data loss prevention and spamming tools.

However, it's not effective because these technical countermeasures are unable to analyze and differentiate between phishing and general email. Furthermore, business organizations provide phishing attack awareness training to their employees against phishing techniques, tactics, and procedures (TTPs) through email security guidelines and standard operating procedures to increase cyber awareness [9].
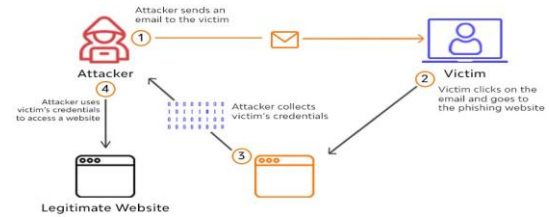
Fig. 3 Phishing attack mechanism

Distributed Denial of service attack (DDoS) attack is one of the most common cyber crimes faced by business organizations when a group of botnets, server or numerous connected internet devices flood the capacity limit and resources of a targeted system as a result in service degradation and unavailability by disrupted normal traffic. Attacker will control his resources to overwhelm the capacity limit of the server by sending numerous requests to the targeted IP address of the website in a short time that trigger a crash. DDoS attacks could stay as long as 24 hours and up to days to retain access to the unavailability of the company's server that lead to reputation damage and profit minimization. There are vulnerabilities in the system that are exploited by the hackers although various protections and prevention are applied.

DDoS attacks (Fig. 4) are conducted with private networks of internet-connected machines and any kinds of IoT devices. Attackers can remotely control the networks that are constituted by computers and devices that are malware infected. These infected devices can be called bots (or zombies), and a botnet where a group of bots is connected in a network. Attackers are allowed to send attack instructions to each bots when the botnet has been established. When each bot sends requests to the victim's IP address, possibly causing the server of the website to be down where limit traffic is triggered during a short period as a result of denial-of-service to normal traffic. Because the server is difficult to identify attack traffic and normal traffic from each bot is considered a legitimate internet device [10].
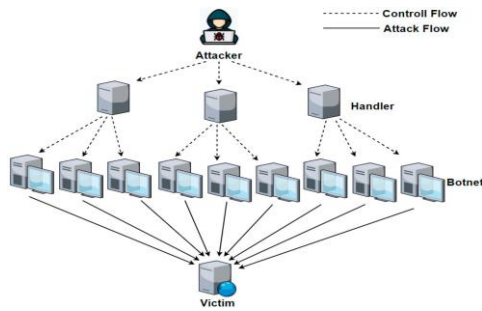
Fig. 4 DDos attack mechanism

Types of DDos Attack

- Ping of Death (POD) attack
  POD attack exploits the particular characteristic based on the maximum size of the packet of transmission control protocol (TCP)/ internet protocol (IP) can be up to 65,535 bytes in normal cases. Generally, in the case of a large IP packet, it will be splitted into numerous fragments and reassembled into a complete packet to the recipient host. However, POD attack exactly exploits the reassembling packet size of 65,535 bytes greater than its maximum to maliciously manipulate overflow of memory allocated as a result of denial of service[11].

- Hypertext Transfer Protocol (HTTP) Flood Attack
  HTTP flood attack is another method of resource consuming by manipulating the HTTP GET and HTTP POST requests while connecting with the victim's machine [12]. HTTP works as a request-response protocol and designs to allow communications between servers and clients. HTTP GET and POST are the two most common methods where GET requests data from a particular resource and POST sends data to a server to update or create a resource [13].

- Domain name system (DNS) Flood Attack
  DNS is the "contact list" of the internet which allows devices able to view the content of websites. It sends a huge amount of DNS requests to the target device in order to overload it and reduce the speed of traffic. Attack will be done by the botnet to generate a large volume of traffic [12].

Cloud computing attack (CCA) is a malicious activity that involves data breaches, manipulating, account hijacking or eavesdropping from a user's virtual machine on the physical server by injecting malware from the hacker. Once the cloud system is deceived, hacker can gain unauthorized access to the user's cloud storage without permission and awareness. CCA could be conducted by various methods such as structured query language (SQL) injection attack, DDoS attack, malicious insider, side channel attack, man-in-the-middle (MITM) attack and abuse of cloud services. Cloud computing attack is the most economically significant threat in the world of virtualization technology, confidential data breach is one of the key diminish potential revenue loss and customer trust [11]. In fact, a massive cyber attack in Sweden in July 2017, a high confidential data stored at cloud system was breached, causing irreversible consequences that exposed national secrets, national security, international wrongdoing, and resignation of ministers. In fact, a massive cyber attack in Sweden in July 2017, a high confidential data stored at cloud system was breached, causing irreversible consequences that exposed national secrets, national security, international wrongdoing, and resignation of ministers. In June 2017, British Airways cloud system was being disrupted by hackers as a result of financial loss at 114 million EUR due to over 100 flights being cancelled from London airports [2].

SQL injection (Fig. 5) attacks is a common attack where an attacker inserts malicious code to unauthorized access to a backend database to gain sensitive information that was not intended to be disclosed or private company data. Generally, SQL injection attacks require an attacker to figure out the vulnerability of a webpage or web application to perform unauthorized access without an official account. For example, a logical user's account coupled with usernames and passwords that are registered at the database, attacker could input malicious code like username = '1' or 2>1 -- ' and password = '111' to bypass source code of the web application [14].
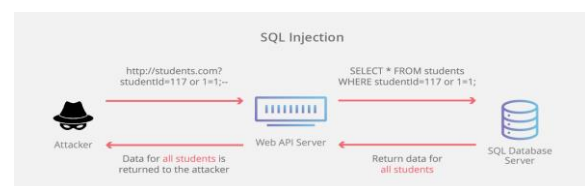
Fig. 5 SQL Injection mechanism

MITM (Fig. 6) occurs when an attacker establishes independent connection with separate ongoing conversations to different third parties by misleading them with the aim to inject, steal or exchange information to manipulate the result. Third parties are not aware that the attacker has control over the ongoing conversation between both parties. All of the message sent by one party to another will go through the attacker and it allows the attacker to tamper the information to mislead both parties [14].
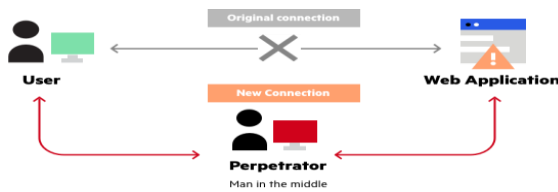
Fig. 6 MITM mechanism

2.2. Internal Cyber Security Challenges within the Organization

Other than external cyber security challenges in the form of various cyber threats, there are seven internal challenges in implementing or improving the cyber security in an organization where these challenges could be grouped into three different categories or foundations for the organizations which are the three pillars of People, Process and Technology [6]. Indeed, human resources or employees are important assets to the organization, however, also one of the internal factors that drive the organization to the risks of exposing them to the emerging and evolving cyber threats when the employees are lacking cybersecurity skills, demonstrating irresponsibility and conducting human error while dealing with sensitive data and processes [6]. Firstly, the lacking of cybersecurity skills is said to be one of the challenges of cybersecurity in an organization and getting worse four years in a row since 2017 where this crisis has impacted almost 70% of the organizations globally according to the fourth annual global study by both the cybersecurity professionals, the Information System Security Association (ISSA) and independent industry analyst firm, the Enterprise Strategy Group (ESG) [15]. In fact, there are numerous factors that lead to such cybersecurity expert shortage crisis which could impose significant cybersecurity risk to the organizations, they are mainly due to global worker shortages, limited hands on skills as well as experience in handling cyber attacks and threats, the increasing workload of employee resulting in lesser time to acquire the relevant skill set, the incompetence of the employees to learn or utilize the cybersecurity technologies to the fullest potential, etc [16]. Thus, it is clearly evident that lack of skills in cybersecurity is one of the most challenging concerns in implementing cybersecurity measures under the pillar of 'People' within the organizations. Besides cybersecurity talent shortage, the irresponsible behaviour of the employees towards cybersecurity is also one of the challenges faced by the organizations in implementing the relevant measures. Most of the employees in an organization would intend to push the responsibility of the IT staff only, whereby the cybersecurity implementation is not solely IT departments responsibility, but all employees from operational level to management level within the

ecosystem of the organization [17]. Thus, with such irresponsible behaviour, the employees tend to put lesser attention onto the data handling which leads to the ignorance attitude towards the basic cyber security practices and possibly violating the compliances of cyber security. Not only that, employees' careless work behaviour such as inappropriate sharing confidential information with third parties without permission, inappropriate use of IT resources, losing their own devices storing company data, etc. could lead to a significant financial loss to the organization [18]. This could be evident when the result from the IT Security Risks Survey 2017 has indicated that almost 11% of the cyber attack incidents are due to careless and uninformed employees as shown in Fig. 7 below [18]. Hence, proven that the irresponsible act of the employees could definitely lead to hard time for the organization to implement the cybersecurity measures.

Fig.7 Most serious cyber attack vectors over the 12
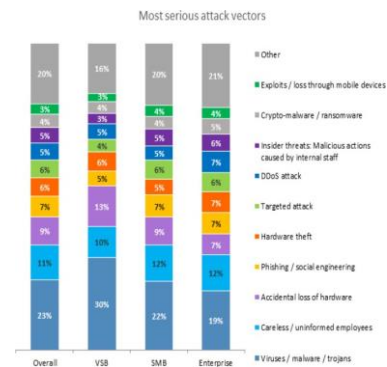
months in 2017.

Furthermore, human error has also contributed as an obstacle towards the implementation of digital security in an organization [6]. This could be proven from a survey conducted by IBM in 2018, indicating that the misconfiguration of cloud servers due to human error is the biggest risk in cyber security in cloud computing with 62% of survey participants made up of IT and security professionals, while 55% of them noting misuse of employee credentials and improper access as one of the human error contributing to challenges while implementing cyber security measures, followed by insecure interfaces at 50% [19]. In fact, misconfiguration of systems usually allows cyber criminals to access large data platforms consisting of usernames, passwords, credit card data, health data, national identification numbers, email addresses, etc. which could possibly be used to request for ransom from the organizations. For example, one of the significant cyber attack incidents happened due to misconfiguration of system and human error in 2018, where a well-known marketing firm in the US, Exactis was reported leaking approximately 340 million files to the public including sensitive personal details ranged

from national identification number to hobbsies by making them available through a server accessible to anyone [20]. This incident has definitely shown that human error within an organization especially involving data handling and system configuration will definitely lead to big loss to the organizations regardless of financial or reputation.

The processes, which is the second pillar of an organization, has played an important role in cybersecurity implementation. There are three cyber security challenges identified under the pillar of 'Process' in a recent research, which include the lack of implementation plans for cyber security measures execution, the misplaced human resources and the lack of budget allocated for cyber security efforts within the organization [6]. The readily compliance and regulations governed nationally such as cyber security guidelines used in Malaysia including Information Technology Instructions 2007, ICT Security Policy 2010, Electronic Government Act 1987, Public Sector Cyber Security Framework 2016, etc., would need a good implementation plan within the organization in order to apply the guidelines into processes and enforce the cyber security strength of the organization [6]. However, most of the organizations do not emphasize cyber security planning which ultimately leads to a higher risk of monetary loss and reputational damage [21]. The mentioned phenomenon could be proven by the results of the survey conducted by The Department for Digital, Culture, Media & Sport (DCMS) in the UK on 1,500 businesses from October to December 2017, where only 38% of the businesses is aware of the new incoming data protection law, while only 27% from these businesses have made the changes on the organization processes and less than half of these businesses made changes to the cyber security practices within the organization, leading to a total of 43% of these businesses in the UK experienced cyber threats in 2017 [22]. The lack of implementation plan could involve cyber security governance, protection measures, hardware maintenance plan, incident response plan, evaluation plan, documentation including policies, Standard of Procedures (SoP), Work Instruction (WI), etc. where the effectiveness of the cyber security implementation is directly affected by the quality of these elements [21]. Therefore, detailed construction of the implementation plan is indeed important to ensure the cyber security measures are being implemented and executed accordingly in order to reduce the risks of exposing the cyber physical asset of the organization towards potential cyber hazards.

The next digital security challenge in an organization under the pillar of 'Process' is poor human resource management. For instance, new arrangements for cyber security talents in the workplace due to promotion to managerial positions, assignments to new departments or even the failure to retain cyber security talent would lead to the reduction of trained and experienced talents that could counteract cyber threats in an organization [6]. Due to talent loss, more time would need to be consumed to acquire capable talents, to train and nurture the replacement, etc which would cost a knowledge gap in terms of cyber security in the organization, leading to the risks of cyber invasion [23]. In fact, a survey conducted on cyber resilience by IBM Security and the Ponemon Institute has shown that 75% of the survey respondents found it difficult to hire and retain skilled cyber security professionals in respective organizations, while 48% of them has also pointed out the complexity in operation due to various type of security tools being implemented, causing the weakening in cyber resilience due to the skill and abilities gap between existing talents and tools deployed within the organization [24].

The sixth challenge in digital security within an organization is the limited budget allocated for the implementation and enforcement [6]. A detailed cyber security budget allocation should include risk analysis cost, technology acquisition or software licensing costs, training charges, insider threat reduction cost, etc. to reduce the potential risk of cyber internal breach and external threats [25]. However, most of the top management in the organizations especially small and medium enterprises do not prioritize cyber security because they would be willing to maximize the profitability instead of spending on cyber protection measures for incidents that may not happen, but they did not consider the possibility of cyber security failures which would cost the organizations even more [26]. For example, RiskIQ estimated financial loss of $17,700 per minute will be imposed on the victims due to phishing attack while Accenture estimated Malware rates up to $2.6 million at most and Ransomware rates up to $646,000 on average [27]. Not only that, organizations with improper cyber security could also be fined as it would induce insecurity within users, for instance, Google was fined at a cost of $57 million in 2019 due to the failure in complying with General Data Protection Regulation (GDPR) in France [28]. Thus, limited budget allocation as a roadblock in cyber security of an organization would definitely cause larger amounts of financial and reputational losses in the future as the risks from organizations internally or cyber threats externally are not being mitigated sufficiently.

Lastly, business organizations are facing cyber security challenges when the technology advancement is evolving too fast, so does the advancement in cyber threat vectors [6]. In fact, the cyber threats are characterized as asymmetrical and multi directional, making the organizations being reactive instead of proactive while facing the new form of cyber attacks[6].

Also, as the world is currently accelerating towards IR4.0, the organization especially in manufacturing industry which adopting technologies such as Robotics and Automation (R&A), Artificial Intelligence (AI) and Machine Learning (ML), Big Data Analysis (BDA) and Cloud Computing (CC), etc, where these technologies adapted are actually associated with the risks of being attacked by cyber criminals [29]. For example, Booz Allen Hamilton has pointed out the largest potential risk of AI and ML algorithms that worry the most is the possibility of leaking the training data to cyber criminals as this training data is sufficient enough to rebuild the original AI model via reverse engineering [30]. Thus, it is clearly shown that the benefits from ever evolving technologies aligning with IR4.0 are always associated with greater risk of cyber hazards invasion.

Recommendations are proposed in terms of the three main pillars which are people, process and technology, in order to help overcome the internal challenges faced by the organizations while implementing cyber security measures.

Firstly, human resources (HR) directors should play a key role or even a leader to keep the organizations safe from internal cyber breaches and employee-imposed cyber threat risks. As a HR director, a HR team should be formed to work with the IT or cyber security department to align the expectations ensuring the digital security measures and policies to be implemented is practical, ethical and appropriate. Besides, the HR director and the team should also take ownership of the security risk posed by employees by providing sufficient education regarding the impact of the employees behaviours towards the cyber security of the organizations as well as identifying employees who may be insiders allowing risks of external cyber threats to be imposed on the company. Overall, the HR team plays a vital role in shaping the workplace culture where the employees are following the reviewed cyber security measures and policies to help reduce the risks of cyber attacks. [31]

Secondly, in terms of processes, the management systems within the organization are vital in combating cyber security breaches. A proper management system must be adopted to ensure all the employees are aware of their duties and responsibilities as well as to adequately manage the level of competence and interest of the employees in cyber security. Besides that, a series of enterprise cyber security governance activities should be conducted from time to time to continuously improve the strength of the cyber security, where the activities include adopting enterprise risk management (ERM) to prevent data loss, and also, conducting threat,

vulnerability and risk analysis tests within the organization to access the sustainment of the cyber security governance. Moreover, the policies, SOP and WI related to cyber security measures and practices should be comprehensive enough and made available to every employee to ensure these guidelines or documents are clear and easy to understand to avoid miscommunication as well as guiding the employees in their decision making process when facing issues

like cyber attacks attempts. Vendor management is also one of the important processes to be focused on regarding cyber security because the vendors may potentially serve as a medium for external cyber threats invading the organizations, thus, it is important to ensure the vendors engaged are equipped with the same level of security. [32]

Lastly, in the aspect of technological advancements, organizations are suggested to adopt AI-based cyber security management systems where the systems allow to leverage up-to-date global cyber security knowledge and industry specific threats to make significant prioritization decisions as well as prompt incident response when exposing to cyber threat, keeping track of all organizations' IT assets in a domain to perform breach risk prediction and generating comprehensive reports on recommendations and analysis of the decisions proposed to all the involved stakeholders. [33]

3. Conclusion

In this paper, cyber security challenges faced by the business organizations have been discussed and analyzed in two different perspectives including the external potential cyber threats and internal factors within the organizations. Notable external cyber hazards such as the Ransomware, Phishing Attacks, DDoS Attacks as well as Cloud Computing Attacks comprises SQL Injection Attacks and Man-in-the-middle have been introduced and their mechanism also have been discussed in this paper. Seven internal cyber security challenges within the organizations have also been highlighted in terms of three different elements which are people, process and technology, where by the challenges discussed include the lack of cyber security skills, irresponsible behaviour of employees, human error, lack of digital security implementation plan, poor human resource management, constraint of budget allocation for cyber security implementation and lastly technology advancement. Indeed, both the internal and external cyber security obstacles faced by the organization have allowed many digital security breaches and invasions to happen, which leads to critical negative impacts to the organizations, corporate

and national such as breach of trust, data, privacy, financial or economical losses as well as the damage of reputation and public image. Therefore, further broad researches and detailed studies on cyber security especially the method of combating cyber threats and enhancing the strength of cyber security in organizations are recommended so that the readers could be provided with useful insights and ideas as well as emphasizing the importance of everyone's role in ensuring the success of implementing effective and efficient digital security measures in the organizations to instill the supportive behaviours and foster the cyber security culture within the companies.

References

1. M. Zwilling, G. Klien, D. Lesjak, Ł. Wiechetek, F. Cetin, and H. N. Basim, Cyber Security Awareness, Knowledge and Behavior: A Comparative Study, Journal of Computer Information Systems, pp. 1–16, Feb. 2020.

2. M. Spremić, and A.Šimunic. Cyber security challenges in digital economy, Proceedings of the World Congress on Engineering, vol.1, pp. 4-6, Jul. 2018.

3. Internet Crime Complaint Center (iC3) Internet Crime Report 2020, Federal Bureau of Investigation. Accessed: Aug. 26, 2021. [Online]. Available: https://www.ic3.gov/Media/PDF/AnnualReport/2020_IC3Report.pdf

4. S. Mohurle and M. Patil, A brief study of wannacry threat: Ransomware attack 2017, International Journal of Advanced Research in Computer Science, vol. 8, no. 5, pp. 1938-1940, 2017.

5. A. Abhishta, M. Junger, R. Joosten, and L. J. M. Nieuwenhuis, A Note on Analysing the Attacker Aims Behind DDoS Attacks, Intelligent Distributed Computing XIII, Springer International Publishing, pp. 255–265, Jan. 2020.

6. C. S. Teoh, A. Kamil Mahmood, and S. Dzazali, Cyber Security Challenges in Organisations: A Case Study in Malaysia, 2018 4th International Conference on Computer and Information Sciences (ICCOINS), Aug. 2018

7. C. Adams, Learning the lessons of WannaCry, Computer Fraud & Security, vol. 2018, no. 9, pp. 6–9, Sep. 2018.

8. B. Hellard, 121 million ransomware attacks recorded in the first half of 2020, Itpro.co.uk, Accessed: Aug. 26, 2021. [Online]. Available: https://www.itpro.co.uk/security/ransomware/356567/1212-million-ransomware-attacks-in-the-first-half-of-2020

9. H. Shahbaznezhad, F. Kolini, and M. Rashidirad, Employees' Behavior in Phishing Attacks: What Individual, Organizational, and Technological Factors Matter?, Journal of Computer Information Systems, pp. 1–12, Oct. 2020

10. M. Aamir and S. M. A. Zaidi, DDoS attack detection with feature engineering and machine learning: the framework and performance evaluation," Int. J. Inf. Secur., vol. 18, no. 6, pp. 761–785, Apr. 2019.

11. G. Levitin, L. Xing, and H. Huang, "Security of Separated Data in Cloud Systems with Competing Attack Detection and Data Theft Processes, Risk Analysis, vol. 39, no. 4, pp. 846–858, Oct. 2018.

12. L. Visalatchi, P.Yazhini, and M. Scholar, The survey DDoS attack prevention and defense technique, International Journal of Innovative Science and Research Technology, vol. 5, no. 2, pp. 65-68, Feb. 2020.

13. HTTP Methods GET vs POST, W3schools.com, Accessed: Aug. 28, 2021. [Online]. Available: https://www.w3schools.com/tags/ref_httpmethods.asp

14. N. Amara, H. Zhiqui, and A. Ali, Cloud computing security threats and attacks with their mitigation techniques, 2017 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC), pp. 244-251, Oct. 2017.

15. Help Net Security, Cybersecurity skills shortage still the root cause of rising security incidents, Accessed: Aug. 28, 2021. [Online]. Available: https://www.helpnetsecurity.com/2019/05/14/cybersecurity-skills-shortage-causes-security-incidents/

16. J. Oltsik, ESG Research Report - The Life and Times of Cybersecurity Professionals 2020, ISSA, Accessed: Aug. 28, 2021. [Online]. Available: https://2ll3s9303aos3ya6kr1rrsd7-wpengine.netdna-ssl.com/wp-content/uploads/2020/07/ESG-ISSA-Research-Report-Cybersecurity-Professionals-Jul-2020.pdf

17. Cybersecurity regained: preparing to face cyber attacks, Ernst and Young (E&Y). Accessed: Aug. 28, 2021. [Online]. Available: https://pdf4pro.com/view/cybersecurity-regained-preparing-to-face-cyber-attacks-ey-559d48.html

18. "The Human Factor in IT Security: How Employees are Making Businesses Vulnerable from Within," Kaspersky Daily. Accessed:Aug. 28, 2021. [Online]. Available: https://www.kaspersky.com/blog/the-human-factor-in-it-security/

19. Kolandaisamy, R., Noor, R. M., Zaba, M. R., Ahmedy, I., & Kolandaisamy, I. (2019, July). Markov chain based ant colony approach for mitigating DDoS attacks using integrated vehicle mode analysis in VANET. In 2019 IEEE 1st International Conference on Energy, Systems and Information Processing (ICESIP) (pp. 1-5). IEEE.

20. T. Collins, Exactis leaks the private details of 340 MILLION people to cyber criminals, including their phone numbers and home addressees, in a one of the biggest security breaches of its kind, Mail Online, Accessed: Aug. 28, 2021. [Online]. Available: https://www.dailymail.co.uk/sciencetech/article-5900071/Marketing-firm-Exactis-leaks-340-million-files-containing-private-data.html

21. S. Ursillo and J. C. Arnold, Cybersecurity Is Critical for all Organizations – Large and Small, IFAC, Accessed: Aug. 28, 2021. [Online]. Available: https://www.ifac.org/knowledge-gateway/preparing-future-ready-professionals/discussion/cybersecurity-critical-all-organizations-large-and-small

22. The Department of Digital, Culture, Media & Sport (DCMS), "New figures show large numbers of

businesses and charities suffer at least one cyber attack in the past year," Gov.uk, Accessed: Aug. 28, 2021. [Online]. Available: https://www.gov.uk/government/news/new-figures-show-large-numbers-of-businesses-and-charities-suffer-at-least-one-cyber-attack-i n-the-past-year

23. Cipher, What Causes Cyber Security Projects to Fail?, Accessed: Aug. 28, 2021. [Online]. Available: https://cipher.com/blog/what-causes-cyber-security-projects-to-fail/

24. Help Net Security, 77% of orgs lack a cybersecurity incident response plan, Accessed: Aug. 28, 2021. [Online]. Available: https://www.helpnetsecurity.com/2019/04/12/cybersecurity-incident-response-plan/

25. Kolandaisamy, R., Subaramaniam, K., & Jalil, A. B. (2021, March). A Study on Comprehensive Risk Level Analysis of IoT Attacks. In 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS) (pp. 1391-1396). IEEE.

26. S. Kabanda, M. Tanner, and C. Kent, "Exploring SME cybersecurity practices in developing countries," Journal of Organizational Computing and Electronic Commerce, vol. 28, no. 3, pp. 269–282, Jul. 2018.

27. J. Fruhlinger, "Top cybersecurity facts, figures and statistics," CSO Online, Accessed: Aug. 28, 2021. [Online]. Available: https://www.csoonline.com/article/3153707/top-cybersecurity-facts-figures-and-statistics.html

28. R. Dillet, French data protection watchdog fines Google $57 million under the GDPR, Tech Crunch, Accessed: Aug. 28, 2021. [Online]. Available: https://techcrunch.com/2019/01/21/french-data-protection-watchdog-fines-google-57-million-under-the-gdpr/

29. Baytamouny, M., Kolandaisamy, R., & ALDharhani, G. S. (2022, April). AI-based Home Security System with Face Recognition. In 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI) (pp. 1038-1042). IEEE. [30]. M. Korolov, How secure are your AI and machine learning projects?,CSO Online. Accessed: Aug. 28, 2021. [Online]. Available: https://www.csoonline.com/article/3434610/how-secure-are-your-ai-and-machine-learning-projects.html

30. Alt, R. (2021). Digital transformation in the restaurant industry: Current developments and implications. Journal of smart tourism, 1(1), 69-74..

31. Putting human resources at the heart of cyber security, PA Consulting. Accessed: Aug. 28, 2021. [Online]. Available: https://www.paconsulting.com/insights/putting-human-resources-at-the-heart-of-cyber-security/

32. The three-pillar approach to cyber security: processes are crucial, Det Norske Veritas. Accessed: Aug. 28, 2021. [Online]. Available: https://www.dnv.com/article/the-three-pillar-approach-to-cyber-security-processes-are-crucial-162890

33. Using Artificial Intelligence in Cybersecurity, Balbix. Accessed: Aug. 28, 2021. [Online]. Available: https://www.balbix.com/insights/artificial-intelligence-in-cybersecurity/

Authors Introduction

Dr. Raenu Kolandaisamy

He received his PhD from the Faculty of Computer Science & Information Technology, University Malaya in 2020. He is currently an Assistant Professor in UCSI University, Malaysia. His research interest areas are Wireless Networking, Security, VANET and IoT.

Dr. Heshalini Rajagopal

She received her PhD and Master's degree from the Department of Electrical Engineering, University of Malaya, Malaysia in 2021 and 2016, respectively. Her research interest includes image processing, artificial intelligence and machine learning.

Dr. Indraah Kolandaisamy

Dr. Indraah A/P Kolandaisamy is a senior lecturer in School of Business Management, College of Business, Universiti Utara Malaysia.

Dr. Glaret Shirley Sinnappan

Dr. Dr. Glaret Shirley A/P Sinnappan holds the position of Assistant Professor in the Department of Information and Communication Technology at Tunku Abdul Rahman University College.

The Study on Perception on E-Waste Among the People

Raenu Kolandaisamy, Heshalini Rajagopal
Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia

Indraah Kolandaisamy
School of Business Management, University Utara Malaysia 06010 Sintok, Kedah
E-mail: raenu@ucsiuniversity.edu.my

Abstract

E-waste is piling up at an alarming rate, which results in our environment's pollution, while only a few are being disposed of correctly. This research aims to discover people's awareness of e-waste and how to handle it. An online survey was done among 75 respondents, who were asked about their general knowledge of e-waste. Most respondents are aware of the effects of e-waste but lack knowledge on how to tackle them, including questions concerning e-waste management and its causes. In conclusion, the general public must be educated on e-waste to reduce and dispose of e-waste efficiently and effectively.

Keywords: E-waste, Technology, Online Survey, Environment

1. Introduction

The 20th century was rich with pivotal technological advancement, from the birth of the television to the creation of personal computers. Technological advancement kickstarted the tech industries towards more innovations that are rapidly improving, which led to inventions such as the iPhone that paved the way for modern-day smartphones. Technology these days has become mainstream. Statistics show an uphill trend regarding the adoption of technologies in a household. For example, colour TV use increased exponentially, recording a 10 per cent adoption in 1966 to 96 per cent in 1990 and a fairly stable adoption rate moving towards the 21st century, staying around 90 per cent. The adoption of computers has also shared the same prosperity, skyrocketing from 20.70 per cent in 1992 to 89.30 per cent in 2016. The biggest winner of all this is the smartphone. Since its introduction around the 21st century, its adoption has rapidly increased, from 35 per cent in 2011 to 73 per cent in 2016. Its predecessor, the cellular phone, is the only factor stopping the adoption of smartphones from growing further since it also has an increase of adoption from 10 per cent in 1994 to 93 per cent in 2016 [1].

There are a multitude of factors feeding the growth of the usage of technology. The role of the government and human nature to favour efficiency are the pivot of this swift growth of technology usage. The government have been emphasizing Vision 2020 lately, which includes an objective to establish a scientific and progressive society that is not only a consumer of technology but also a contributor [2]. Another thing to note is the term Industry 4.0, which aims to create a smart factory environment. Alternatively, technology is known to ease day-to-day tasks. Due to this, technology has infused into our lives and become an ongoing trend. To make it clear, if the five big tech companies are combined, they would have a higher revenue (up to 802 billion) compared to Saudi Arabia, which has the 19th highest GDP at 684 billion [3].

Unfortunately, there is always a catch. The ever-increasing demand for technology causes tech companies to manufacture more and people to upgrade their technologies. All of this causes a domino effect, which leads to e-waste. Electronic waste or e-waste could be defined as any electrical and electronic equipment disposed of without considering it being repurposed. The term covers any objects with screens, temperature-related equipment like refrigerators, lamps, large or small electronic appliances such as a vacuum or dishwashers, and portable electronics ranging from Global Positioning System (GPS) to calculators [4]. As with any production, the manufacturing of technologies is not proportionate to its recycling work. As of 2016, 44.7 million metric tons of e-waste were produced,

comparable to 6.1 kilograms per inhabitant (kg/inh), which increased compared to the 5.8 kg/inh amount in 2014 [5]. In other words, the perception of people regarding e-waste remains uncertain. Therefore, this study aims to investigate the general public's perception of e-waste in Klang Valley.

2. Background

2.1. Overview

Electronic waste and electronic waste identify mechanical and digital products that have been discarded. These wastes are also regarded as appliances deemed refurbishable, recyclable, disposable and saleable. The processes involved when managing e-waste are burdensome, as they cause pollution and negatively impact human conditions, especially in developing countries.

2.2. Definition

E-waste or digital waste is generated after completing its useful life when an electronic item is discarded. The rapid expansion of innovation and the economy powered by usage contributes to producing a large amount of e-waste every minute. The European WEEE Directive classifies electronic and electrical equipment for waste management purposes into ten categories namely large household appliances (e.g., refrigerators, washing machines), small household appliances (e.g., toasters, coffee makers), IT and telecommunications equipment (e.g., computers, phones), consumer equipment (e.g., TVs, radios), lighting equipment (e.g., lamps, fluorescent tubes), electrical and electronic tools (e.g., drills, saws), toys, leisure, and sports equipment (e.g., video game consoles), medical devices (e.g., X-ray machines, pacemakers), monitoring and control instruments (e.g., thermostats, smoke detectors), and automatic dispensers (e.g. vending machines).

2.3. Current Situation of E-Waste

Electronic waste keeps growing. Today, people are buying more electronic devices, which are being quieted faster. For example, cell phones generally have a short lifespan of one to two years. People use most of their incomes to buy electronics yearly. According to a study, more than 20 million tons of e-waste are produced annually.

The environmental impact of e-waste is a growing concern in today's society. With the widespread use of electronic devices and rapidly advancing technology,

people constantly upgrade to more efficient electronics, contributing to the ever-increasing e-waste problem.
The old electronic devices are thrown away without considering the toxic materials from the electronic devices released into the environment. The groundwater, soil and air affect land and sea animals, causing harmful substances to seep into groundwater. When e-waste is released, toxic chemicals are released into the air, damaging the atmosphere.

Numerous electronic devices contain hazardous materials like lead and barium. In particular, the release of lead into the environment can result in adverse health effects, including damage to the human blood, kidneys, and peripheral nervous system.

2.4. Amount of Electronic Waste Worldwide

Increasing technological changes, media developments (tapes, computers, MP3), affordable market prices and the deliberate design of short life span technologies have culminated in a rapid buildup of accumulation of e-waste internationally. An achievable countermeasure has been developed, but in most situations, it is necessary to implement it to work legally in the real world, which is time-consuming even in the most advanced nations, especially due to the bureaucracy.
Different models and versions of TVs, processors, storage units, and audio modules have varying lifespans. Among these, screen systems and processors are commonly affected by electronic waste. Screen systems are often discarded without attempts at repair, whereas processors see a continuous influx of new and improved versions to meet the demands of ever-advancing video gaming graphics.

2.5. Environmental Impact

The methods used to manage e-waste in developing countries are often questionable. It has caused many issues that have an unfortunate effect on Mother Nature. These range from airborne or liquid-borne material leakage into the atmosphere, soils, water and foods, affecting animals, humans and plants.

2.6. Electronic Waste Substances

Many electronic parts can be recycled in manufacturing new computer devices, while others are reduced to materials that can be reused for industries as diverse as building, flatware, and jewelry. Large-scale materials include fiberglass, PVC, fabrics, gold, iron, copper, quartz, beryllium and oil. Almost all electronics contain lead-tin (as solder) and copper.

2.7. Effects

It is known that 13 percent of e-waste is processed in developing countries. Yearly, e-waste is generated at a large sum (around 40 million metric tons) worldwide. According to the United Nations Environment Program (UNEP), the European Union contributed up to 22.5 percent of the e-waste. Meanwhile, in China, the Philippines, Pakistan, India and Vietnam, unconventional ways of the procession of e-waste are largely practiced, with 50-80 percent of the e-waste procession. The processes involve dismantling, burning and shredding the items in backyards. These actions cause bi-products that have been proven impactful (in a bad manner) to mother nature and humans alike. Their action, however, is also due to the large amount of e-waste flowing in since they are handling e-waste domestically and internationally.

2.8. Methods Used to Reduce E-Waste

The production of devices and the use of the materials that go into their manufacturing represent a major way of embodied energy. Reducing the e-waste helps protect the resources and minimizes the required total energy taken from the earth.

- Recycling
  Recycling the highly valuable materials of old electronics rather than making or mining new materials for new devices.
- Stem the Spread
  Before buying a new electronic device, think twice. We can use the old electronic as an MP3 player or GPS device. Please do not throw it away.
- Sell and donate old electronics.
  The best and easiest way to minimize the electronic waste footprint is to sell or give it away to the people who need it. Giving out an electronic device to poor people is a good example of reducing e-waste; not only that, but it also makes us feel good.

- Educate and organize
  We may also educate our neighbors and local businesses on recycling. Returning these electronics to the recycling flow will lower the pressure on the mining, refining, manufacturing, and transportation industries and reduce pollution. Make a much more affordable trip to the local recycling centre to organize a group project to clean up old electronics. For example, Green Citizen.com organizes pickups for those in the San Francisco Bay area and provides mail-in services to those who cannot make it.

- Recycle and dispose of e-waste properly.
  Inappropriate throwing of e-waste is a more hostile problem as we speak, as its volume is increasing every year. For this reason, major business brands provide their customers with a platform to recycle their old electronics.

- Maintain the electronics.
  Effectively taking care of electric gadgets and appliances is a way of killing two birds with one stone since it could halt the production of e-waste and, at the same time, save money since it could function longer.

- Store data online
  Cloud services are effective in minimizing environmental effects. By keeping data online, we can access the data from any place all over the globe. There is no need to take other storage devices.

3. Methodology

This study is based on an online survey. The questionnaire was prepared with multiple-choice questions that comprised general information about the respondents (age and gender) and knowledge in regards to the topic, such as the proper e-waste disposal technique, country produces the most e-waste, e-waste hackable after being thrown away, country hosts the biggest digital dumping ground, the main causes of e-waste, its effect on children within the proximity of the dumping ground, environmental pollution and the main victim of e-waste. The data obtained from the survey was computed in Microsoft Excel for analysis. The statistics of each question were presented in the form of tables, pie charts and bar charts, as shown in the next section.

4. Result and discussion

In total, 75 respondents took this survey. 69.3 percent of the respondents were male, and 30.7 percent were female. Therefore, most of the respondents are males. The survey was taken by people ranging from the age of 19 and below or 20 to 49. None of the respondents over 50 years old participated in this survey. About 74.7 percent of people aged 19 and below took this survey, while 25.3 percent of respondents aged 20 to 49 took this survey. Seventy-six per cent of respondents choose to give to a reliable e-waste company. Moreover, 18.7 percent of respondents choose to throw e-waste in the

Raenu Kolandaisamy, Heshalini Rajagopal, Indraah Kolandaisamy

rubbish bin, while 4 percent throw e-waste by throwing it in the streets, and 0.3 percent of respondents choose to burn e-waste. The right answer is to give it to an e-waste company, as shown in Fig. 1. Based on a study, Switzerland and Germany tackled the e-waste problem efficiently. All these countries have a specialized company that handles it, such as the Public Waste Management Authorities (PuWaMa) of Germany or the Producer Responsibility Organization (PROs) of Switzerland [6].

Fig. 1: Respondents way of disposing E-waste
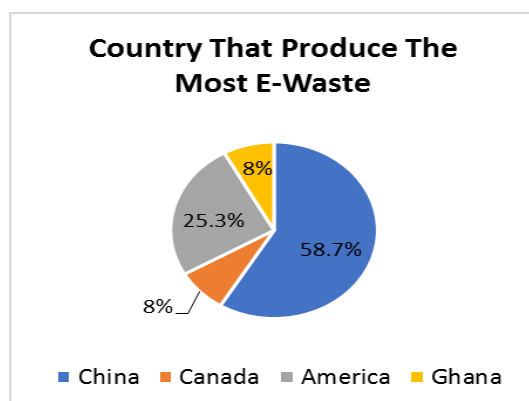
Fig. 2: The Country that produces the most E-Waste

Based on Fig. 2, respondents think China produces the most E-waste, with 58.7 percent of respondents, USA at 25.3 percent, and Canada and Ghana, gaining 8 percent of respondent choices. Unsurprisingly, China produces the most e-waste, amounting to 7.2 tons per annum. However, it is also worth noting that Norway tops it all in kg per capita, at 28.3 [7].
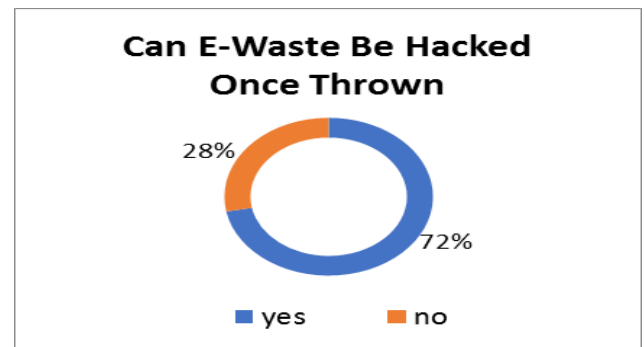
Fig. 3: Result of respondent regarding the prospect of E-Waste being hackable

Fig. 3 represents respondents' view on whether they think e-waste are possibly hackable once thrown away. As we can see, 72 percent of respondents said yes while 28 percent of respondents said no. Fig. 4 illustrates the percentage of people's predictions on which country would have the largest digital dumping ground. Brazil was the most chosen, at 41.3 percent, followed by Ghana, which had 25.3 percent of the respondents' choices. Mexico came in third, sharing 18.7 percent, while Kenya received 14.7 percent. Most respondents
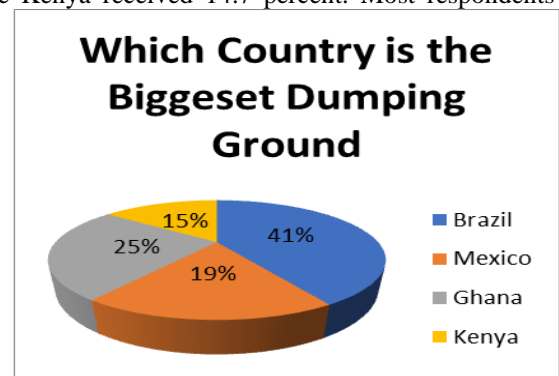
chose Brazil as the world's biggest digital dumping ground.

Fig. 4: People's view on who shares the biggest E-Waste dumping ground

Fig. 4 illustrates the percentage of people's predictions on which country would have the largest digital dumping ground. Brazil was the most chosen, at 41.3 percent, followed by Ghana, which had 25.3 percent of the respondents' choices. Mexico came in third, sharing 18.7 percent, while Kenya received 14.7 percent. Most respondents chose Brazil as the world's biggest digital dumping ground. However, statistics showed Ghana is the worlds' largest digital dumping ground [8].
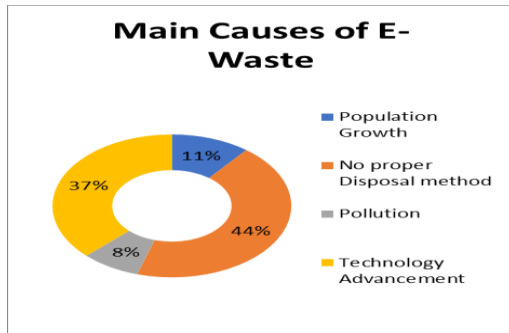
Fig. 5: Main causes of E-Waste

Based on the above Fig. 5, our respondents regarded improper disposal as the main cause of e-waste, gaining 44 percent of the responses. On the other hand, 11 percent of respondents chose population growth, 8 percent chose pollution, and 37 percent assumed due to technological advancement. In truth, the main contributor to e-waste is the failure to dispose of it effectively. Countries usually lack appropriate infrastructure, legislation, quantification measures and funding, resulting in a mass accumulation of e-waste [9].
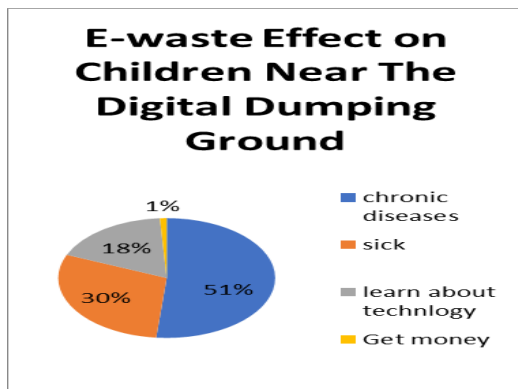
Fig. 6: Effect of E-Waste on children near the digital dumping ground

The pie chart in Fig. 6 shows respondents' views on the effect of e-waste on children near digital dumping grounds. 29.7 percent of the respondents chose the answer of the possibility of the children getting sick. 51.4 percent chose the answer of children getting chronic diseases, 17.6 percent of children would get a chance to learn technology, and 1 per cent would get money. Children exposed to e-waste are highly likely to get a chronic illness or disease. It could harm the nervous, immune, reproductive and digestive systems. Due to this, they would have permanent, irreversible damage to their body (World Health Organization [10]. The pie chart in Fig. 7 below shows the possible types of pollution e-waste can do to the environment. 50.7

percent of respondents chose soil pollution,17.3 per cent chose water pollution, and 32 per cent chose air pollution, concluding that soil pollution is the main pollution mostly caused by waste. E-waste produces all types of pollution. It affects the air from the open burning and dust emitting into the air, while water is polluted when chemicals used are improperly treated and let out into the normal world. However, soil pollution is the most harmful form of pollution caused by e-waste [11], [12]. Toxic chemicals would be absorbed into the soil, which later would be used for farming. This would then grow harmful foods that humans would consume. In other words, it slowly and silently kills people who ate the food produced.
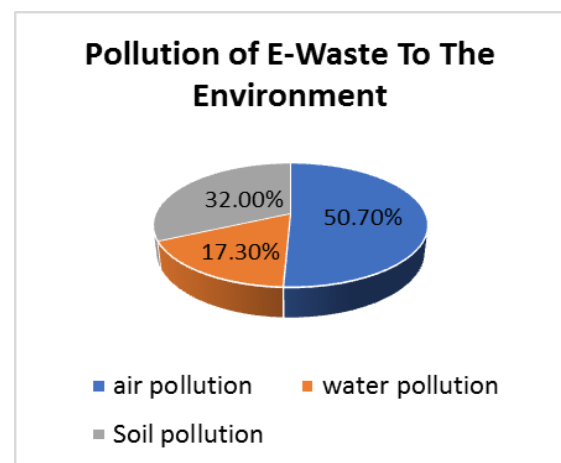
Fig. 7: Respondents' view on the possible pollution due to E-Waste
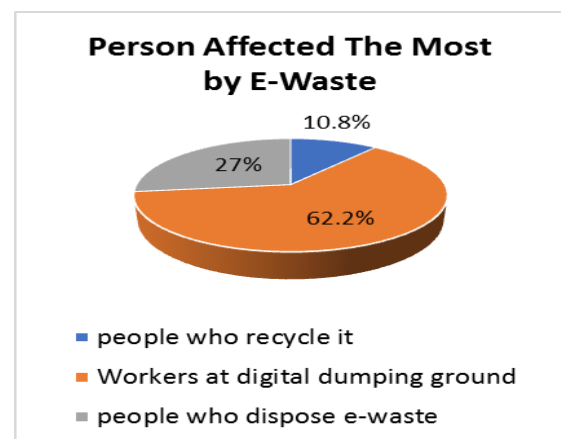
Fig. 8: People affected by E-Waste

The types of people affected by e-waste are shown in Fig. 8. 62.2 percent of respondents chose that workers working at digital dumping grounds are affected most

27 percent of respondents chose people who dispose of their e-waste, and 10.8 percent of respondents chose people who recycle e-waste. In conclusion, workers working at digital dumping grounds are affected the most by e-waste. The people that are most affected are those who work near the dumping ground. As stated before, e-waste causes air, water and soil pollution. Therefore, anything they do has a chance of being contaminated, thus harming them, especially in the long term.

5. Conclusion

E-waste is one of the rapidly growing environmental problems around the world. This is because of the lack of awareness and appropriate information on disposing of these wastes. Another reason for the rapid growth of E-waste is technological advancement. As technology advances, the old technologies will be left out because they are outdated. Furthermore, because of the ever-growing rate of technological advancement, we are creating more waste than ever every day. Most of these wastes are hazardous to the environment, and due to the growing amount of waste, we are damaging the ecosystem more and more every day [13]. Moreover, very few authorities are making any necessary moves yet. We hope the government can create more campaigns and talk about e-waste, its destructive properties and what we can do to stop this environmental crisis. The school authorities must also create awareness in kids and young adults so they would not even destroy the future. We hope the government and school authorities take immediate action to reduce this environmental crisis.

References

1. Ritchie, H., Roser, M. & Mathieu, E. (2019). Technological change. https://ourworldindata.org/technology-adoption
2. Prime Minister's office of Malaysia. (n.d.). https://www.pmo.gov.my/vision-2020/malaysia-as-a-fully-developed-country/
3. Desjardins, J. (2019, April 3). How the tech giants make their billions. https://www.visualcapitalist.com/how-tech-giants-make-billions/
4. California Department of Resources Recycling and Recovery (Cal recycle). (2019). https://www.calrecycle.gov/electronics/whatisewaste
5. Balde, C.P., Forti, V., Gray, V., Kuehr, R. & Stegmann, P. (2020). The global e-waste monitor 2020: Quantities, flows and the circular economy potential. https://collections.umu.edu/eserv/unu:6341/Global-e-waste
6. Chaudhary, K. & Vrat, P. (2018). Case study analysis of e-waste management systems in Germany, Switzerland, Japan and India: A RADAR chart approach. Benchmarking: An International Journal, 25(9), 3519-3540. https://doi.org/10.1108/BIJ-07-2017-0168
7. Kinhal, V. (2017). Highest e-waste generating nations in the world. https://www.worldatlas.com/highest-e-waste-generating-nations-in-the-world.html
8. Biello, D. (2014, Jan 1). E-waste dump among top 10 most polluted sites.
9. Chowrimootoo, D.J.M (2011). E-Waste causes, hazards, barriers and approaches to effective management. https://www.academis.edu/8958150/e-waste
10. https://www.who.int/ceh/risks/ewaste/en
11. Wilson, D. (2016). Impacts of WEEE (e-waste). https://ewaste.ece.uw.edu./students/impacts-of-e-waste-on-the-environment/
12. Kolandaisamy, R., Subaramaniam, K., & Jalil, A. B. (2021, March). A Study on Comprehensive Risk Level Analysis of IoT Attacks. In 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS) (pp. 1391-1396). IEEE.
13. Kolandaisamy, R., Noor, R. M., Zaba, M. R., Ahmedy, I., & Kolandaisamy, I. (2019, July). Markov chain based ant colony approach for mitigating DDoS attacks using integrated vehicle mode analysis in VANET. In 2019 IEEE 1st International Conference on Energy, Systems and Information Processing (ICESIP) (pp. 1-5). IEEE.

Authors Introduction

Dr. Raenu Kolandaisamy

He received his PhD from the Faculty of Computer Science & Information Technology, University Malaya in 2020. He is currently an Assistant Professor in UCSI University, Malaysia. His research interest areas are Wireless Networking, Security, VANET and IoT.

Dr. Heshalini Rajagopal

She received her PhD and Master's degree from the Department of Electrical Engineering, University of Malaya, Malaysia in 2021 and 2016, respectively. Her research interest includes image processing, etc.

Dr. Indraah Kolandaisamy

She is a senior lecturer in School of Business Management, College of Business, Universiti Utara Malaysia.

Emergence of Cybercrimes in Online Social Networks

Raenu Kolandaisamy, Heshalini Rajagopal
Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia

Indraah Kolandaisamy
School of Business Management, University Utara Malaysia 06010 Sintok, Kedah
E-mail: raenu@ucsiuniversity.edu.my

Abstract

The rise of social networking websites has been seen in recent years. Everyone will be spending most of their time on social networking websites such as Facebook, Instagram and Whatsapp. The great advantage that this social networking website offers benefited many of the users. It can help people to promote themselves or their business to gain more popularity and also customers through these social networking websites. There are many cybercrimes that can be identified which are identity theft, hacking, fraud and so on. The emergence of cybercrimes has created an awareness so that the users will know what the common attacks are and how they can be prevented from being lure and being a victim of this attack. This research will discuss about the attacks and how these attacks can be prevented by the users.

Keywords: Social Network, Cybercrimes, Digital World

1. Introduction

Social network has been growing in the 21st century and has become one of the trend that people use to communicate everyday. Social network is connecting with different individuals using the internet through social media websites [1]. It allows many individuals to be connected to each other and thus allowing many people to have communication with each other around the world. Social networking has been so popular because of the great ability and the features that it offers. The number of social networking sites has been growing from time to time. Some of the popular social media application and websites are Facebook, Whatsapp, WeChat, Instagram, Twitter, Google + and Skype. An example will be Facebook that have over 2 billions of active users each month based on Fig. 1 as of June 2017 [2].

Each of the social media platform have different features and design in it. There are many reasons why social networking has gain so much popularity over the years. The first reason will be the opportunity to meet new people. We live in a world that is full of different culture and environment thus meeting new people from different countries all over the world is indeed a great joy for many of us.
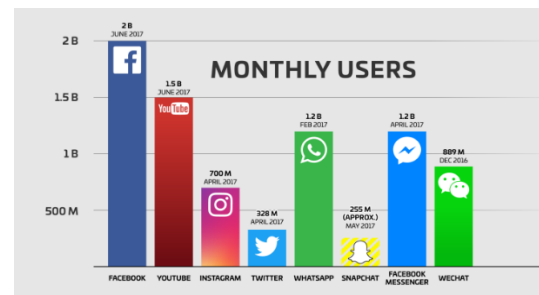
Fig. 1: Monthly users of social media applications

The second reason will social networking sites are user friendly and thus is easy to use. Most sites are developed with a very user friendly interface and is easy to navigate which require a little knowledge to learn. The third reason is that it is available to use on many platforms such as desktop computers, tablets and also smartphones with different operating system. The fourth reason is that social networking sites are free of charge to use and therefore many users will certainly want to try and use it [3]. Through gaining popularity to this sites, many also get the opportunity to advertise their product and services thus making social networking a good place to start and grow your business. The fifth reason is that social networking sites create the opportunity for graduates, businessmen and also

Raenu Kolandaisamy, Heshalini Rajagopal, Indraah Kolandaisamy

professionals to find their ideal jobs and also to meet the people that have the same interest as theirs.

The rise of social media certainly has a great impact on the community and the great benefits it has to offer but there also dangers in using social networking because of the large community it possesses and thus making user vulnerable to security threat and attacks. It can be shown that in recent years through social media that many have faced attacks such as scam and conned by users on social networking sites trying to impersonate another person seeking for donation and other personal information such as bank details and so on. There are very dangerous in the society and unaware users might be lured so this studies to help to increase the awareness of cybercrime in social network.

Cybercrime can be any activities that compromise on the security of the users and try to exploit another user using the internet. It is an illegal activity that involves the use of computer and the internet through a platform like social networking to gain or to cause destruction to another user on the that network. Cybercrime is the greatest attack based on the Fig. 2 that shows the motivation of the attackers [4].
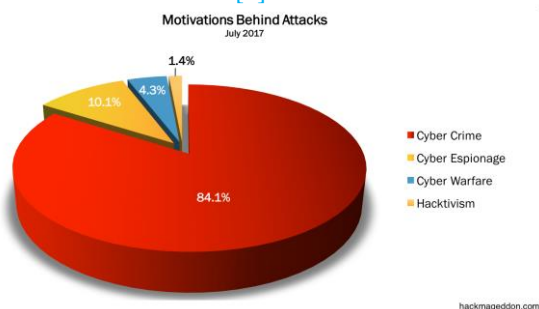
Fig. 2 Motivation behind attacks

Cyber criminal can be categorized into many aspect. It can be broken down into a few different parts which are

- A website lure a users to click on an image or a link to visit another website that can spread malware to the users computer to take control and gain information from the users computer.
- A scammer try to claim to impersonate someone to be a bank staff to gain personal information such as password or even donation by creating a fake account.
- The hackers hack your social networking account to use as a way to get details from close people like family.

The danger to the victim here is that, they can lose their information in a matter of minutes. It will be a big problem to many society and what's more dangerous is that one victim can lead another victim to be a victim. For example, a victim receive an email regarding a free gift that can be obtained if they follow the steps to complete it on a website, but he or she sends it to her many friends on Facebook thus making them to be a victim as well because of the person doing things unknowingly.

2. Literature Review

2.1. Detecting Compromised Accounts on Social Networks

Social networking is one of the most used website in today's internet world. Social networking is a website that uses the internet to bring people together to share information, ideas and also make new friends. This can also be known as social media. Everyone loves to use social media because it is convenient and free. Social media websites such as Facebook, Instagram, Twitter and Whatsapp can all be accessed on various platform such as smartphones, tablets and computers. This has been a great advantage to many users around the world. Social networking has been very demanding and popular but the ever increasing of cyber criminals activities are also increasing as well making it very dangerous.

The study on detecting compromised accounts on social networks by Manuel Egele, Gianluca Stringhini, Christopher Kruegel and Giovanni Vigna will be discussed. The act of compromising on social network accounts has been a great and profitable job that has benefit the cyber criminals. The reasons here is because the speed and reliability of social medias account. By targeting a social media account of a famous artist that have many followers and by sending out fake messages can be able to scam many thousands of users that have follow that famous artist. It is very dangerous because once an attacker managed to compromise a social network account, the person can use it for many purposes such as sending out spam messages to people or even phishing web sites link. It is a higher benefit for attackers to target large and popular company that has large amount of followers like news company.

Fake accounts and legitimated hacked accounts have different characteristic so the system is harder to detect because it is not the fake accounts that is created by a user, it has been a good standing account but it is being compromised. This research introduced COMPA which is basically a detection system that is designed to identify compromised social network accounts. COMPA is use to monitor the habits of the user in a social network account [5]. A person who uses social media like Facebook usually have a fairly stable behaviour from the messages that is being send or the amount of time that the person is active. How this works that COMPA will actually build a behavioural

profile for social network accounts and then it will start to monitor and compare the new messages that is being sent by the account, if it looks different from what it usually is then Compa will flag it as being compromise. There are two common use to detect compromise account and the first is call suspicious group. Suspicious group are those that send and do more than what the accounts normally do. For example, account A usually send out 50 messages a day and averages about 45 per month, but that particular month itself, it exceeded the regular threshold by sending over 500 messages a day. It can be easily be detected especially if the amount raise too significantly.The second is call bulk application. Bulk application are the one that send out large amount of messages with the same amount of text and words.

The studies have evaluated both social media websites which are Facebook and Twitter. The data collection method was to collect real time tweets from Twitter and ran Facebook experiments on a large dataset crawled in 2009. The advantages of COMPA is able to identify groups of compromised accounts that are to distribute malicious messages on this accounts.

The limitation of this studies is that if the attacker is aware of COMPA, the attacker learn about the victim's behaviour before compromising so that COMPA cannot detect the unusual behaviour of the account because of the similarity of the action that is being possessed by the attacker that imitate the action of the victim. Besides that, social networking sites are harder to gather data.

2.2. Cyberbullying detection in the Twitter network

The rise of social networking websites has been such a great impact to many users in the world. Many use this opportunity to gain new friends, establish new relationship and also to connected with family and friends. Although it has been seen by many as a benefit but the use of social networking websites has been so popular that many have use it in the wrong manner to abuse it and gain advantages over it but causing harm to another. The discussion in this literature review is by Mohammed Ali Al-garadi , Kasturi Dewi Varathan and Sri Devi Ravana which is about the experimental case of cyberbullying detection in the Twitter social network[6].

Firstly cyberbullying can be defined as the use of information and communication technology on the internet and it can on any social media platform for example Facebook, Twitter, Google+ and also Instagram to harass another user. This research and studies is to focus on how cyberbullying happens on Twitter. The study is to improve the detection rate on

detecting cyberbullying on Twitter that has been done by detecting the use of offensive language on Twitter. The detection of of abusive language can include the use of acronyms and words that can increase the cyberbullying classifier. Cyberbullying can be also associated with the aggressive behaviour of the users on Twitter. As certain words can be harder to detect because the words needs to be read as a whole sentence to know whether it is offensive or not. In this case, plenty of information that needs to be utilized to achieve and accurate detection to says that the particular tweet is dangerous and abusive. The detection of vulgar words can also be used as this also a signal of offensive behaviour. The cyberbullies can also use many ways to tweet and of way is to use abbreviation or acronyms to shorten the sentence. The studies uses machine learning algorithms to detect the words with various method. They are four different ways that has been used which are NB, LibSVM, random forest and KNN. This are all the method that is used the various results and conclusions has showed that the study of abusive language and cyberbullying will required the machine to adapt and new learn words as time goes on because new vulgar words with abbreviation may be form and thus making it difficult to detect. The studies suggest that using other social medias can also improved the learning process of the system so that cyberbullying can be resolved and reduced in the community of social networking users.

2.3. Organised crime goes online: realities and challenges

The Internet is like a vast galaxy of interconnected networks that allows many transactions and events taking place virtually from the whole world - including cyber crimes and also Internet-facilitated organized crime. The study of the realities and challenges of organised crime online by the author Anita Lavorgna will be discussed.

Cyber crimes are known as a group of criminals that uses ICT infrastructure to commit crimes, such as phishing, malware injection, Denial of Service, and so more. However, in this world we are living in, organized crime groups (OCGs) have existed for long time and have conducted crimes offline. This paper mostly discussing on how these OCGs make use of Internet or online presence to evolve themselves in today's situation where everyone is more Internet-dependant.

The data for this paper are from interviews and also data collected from law-enforcement officials and experts researchers in the United Kingdom, United States, and the Netherlands. 120 case studies regarding crime with

the usage of Internet are also taken into the analysis of this paper.

This paper focuses on two types of OCGs, one is traditional business-style groups, the other is mafia-style groups. The type crimes committed by these business-style OCGs are identified - mostly on trafficking activities such as wildlife, human, psychoactive substances (drugs), and also fake medicines. These criminal activities are more to businesses motive and thus using the Internet to facilitate online. With the use of online social forums, instant messaging feature for example, it is easier for these crime groups to conduct and also expand their businesses ties with the client.

Mafia-style groups, however, were found to be still grasping on the use of Internet technologies for their criminal activities. Trafficking and gambling activities were their main focus. However, mafia-style groups as are mostly depend on a territorial and political trust, hence face-to-face interactions are needed. The highest hierarchy of this group were found to be mostly leaded by a leader who are not fancy of the use of the Internet. However, as these groups keeps evolving in the future, the usage of Internet in their criminal activities are to be expected as new generations will lead in the future.

The previous points discusses the reality of these OGCs. The challenges however, observed on legislation and policies for these Internet-mediated crimes. It is found that the regulatory policies regarding these crime activities via the Internet are not enough to control them. The instantaneous and universality of the Internet makes these inhibits these regulatory policies to fully take effect. Other challenges found for legislation and policies are such as different approaches applied in different countries as some of these countries have difference in experience in regards to tackling online crimes and carrying out investigations. [7]

2.4. Privacy and Security Issues in Social Online Networks

Social networking has becoming a necessity in today's society as more people in the world are having easier access to the Internet. However, this new necessity can bring a potential danger especially to new and also "naive" users of online social networks, especially teenagers. In this paper by Radhika Bhagat, Rajvi Modi, Palaumi Patel, and Harshil Joshi discusses about the potential privacy and security threats in social online networks today.

Users of online social networks were discussed in this paper. Most of the users using social network sites (SNS) to get informations and happenings all around them instantaneously. Personal information were propagated by them using SNS. Users of SNS tend to

take security lightly and this poses security and privacy threat upon their account.

Among the security issues discussed in this paper are, firstly, identity theft. It is defined as an action of acquiring someone's identity without permission and misusing it for malicious purposes, such as fraud. The author(s) suggested that users of SNS should not simply accept unknown friend requests and check the requesting account beforehand for safety.

Secondly, phishing is also a security issue found in SNS. Preparators of phishing uses 'mock-up' websites to lure users into logging in and thus exposing their sensitive credentials such as login credentials. Author(s) suggested users to always use trusted devices to log in to accounts and be wary of unknown websites. Other security issues found in SNS are such as brute force attack, Sybil attack, and Trojan attack.

Privacy issues in SNS were also discussed in this paper. Privacy settings is said to be always overlooked by most SNS users therefore this poses a huge privacy risk to them as social networks get to hold users sensitive data. It is found that some users even share their sensitive data, whether aware or unaware.

Among privacy threats discussed are such as cyber stalking, whereby most SNS has made it easy for these lurkers to stalk users' profiles without adding them as friends. Author(s) suggested users to keep their profiles privacy mode hidden and only to be shown to trusted parties only. Secondly, cyberbullying also a privacy threat found in SNS. Cyberbullying example are such as commenting illicitly on people's photos or posts and also tarnishes one's reputation online, mostly anonymously. Author(s) suggested users to keep things shared on SNS away from third party users and also be wary of them.

Other privacy threat includes profile cloning, whereby preparators create another account with similar information as the users', and therefore performing malicious actions online to the victim's friends.[8]

2.5. Cybercriminal Networks, Social Ties and Online Forums: Social Ties Versus Digital Ties Within Phishing And Malware Networks

As we are living in a digital age, everything is adapting to it. Not just people, and organizations as a whole, but even cyber criminals too. The usage of digital means such as Internet, and its platforms such as forums, social networking services, has made it easier for individuals to get together and communicate anywhere and anytime. This paper, authored by E. Rutger Leukfeldt, Edward R. Kleemans and Wouter P. Stol, discusses about

cybercriminal networks applying social ties and/or digital ties for their activities and recruiting.

A social network usually formed by actor to another actor with ties between them that defining their connection and relationship. A social tie are usually formed via social contact, for example, meeting in a specific place such as cafes, to engage social activity. Digital tie can be formed virtually online without social contact such as using forums.

Social relationships are crucial in criminal networks. However, social relationships are highly clustered with existing relationships such as contacts from former workplace, friends, family, and these relationships are said to yield less opportunity. With Internet, the boundaries of traditional social network are blurred and "specific offender convergence setting" can take place via online forums where these cybercriminals can get together and share information online from different places and also recruit new members or enablers.

The origins and growth of cybercriminal groups were mostly investigated in this paper. The data for this paper is taken from 18 Dutch police files on cybercrimes that provides them with knowledge cybercriminal networks and its members.

Four models of growth were developed in this paper. It is found that cybercriminal networks' growth are associated with recruitments of enablers - professional and/or recruited. These enablers provides services to cybercriminals to give them access to useful and relevant information such as credit card information and postal information details. The four models of growth consists of the first stage, that is "entirely through social contacts", second stage "based on social contacts; forums for recruiting specialists", third stage "based on forums, with social contacts used to recruit local criminals" and fourth stage that is "entirely through forums".

It is found that most of the cybercriminal networks uses social contacts entirely (first stage) to grow their network and does not use online forums. However, in future, forums may become a more common place for the growth of cybercriminal networks in this digital era. [9]

3. Methodology

The methodology used in this paper are are as follows:
- Questionnaire
  A set of questionnaire created in Google Forms is distributed online for respondents to fill in. The respondents sample size is taken among students of UCSI University as a representative of the public society. The questionnaire would help to analyze the approaches and actions

taken by the society in dealing with online social networks, especially in situations that may cause cyber crimes to take place.
- Social network visualization
  A Facebook dataset from Stanford Network Analysis Project (SNAP) are used to demonstrate and visualize how actors are connected in social network such as Facebook. [10] Social network analysis visualization tool used is Gephi.

4. Result and discussion

We have found that the problem with cybercrimes will continue to increase as more users increase in the use of social media. This is because the great number of users can be a good target for the hackers and attackers because it can be very lucrative if the hacker is being able to compromise millions of users at once. That is why online social media has also been a dangerous platform to know friends because of the use of a device and a face to face conversation has not been done and that makes it very difficult to know the real person's identity. The person can be a boy on Facebook talking to another person, but in reality she is a girl. That is really a very dangerous act and it can cause many unwanted things to happen on social media websites. Nowadays, because people are also attracted by free stuff and gifts only, it can be such an easy way to cheat people because of the lack of knowledge the users possess and thus it is very dangerous for users to click on links that link to malicious website that can harm the user's computers.

To find out how actors connected to each other in a social network, we have used a social network analysis tool software Gephi to visualize the network based on a dataset collection from Stanford Network Analysis Project (SNAP). This dataset is particularly from Facebook. As shown in Fig. 3 this is the result of the visualization projected from the dataset. The type of the dataset is undirected, it has 4039 nodes, and 88234 edges.

Raenu Kolandaisamy, Heshalini Rajagopal, Indraah Kolandaisamy
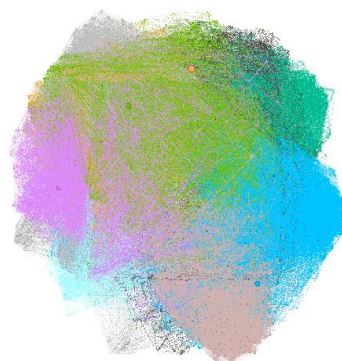
Fig. 3 Result of the visualization projected from the SNAP dataset

Fig. 4 shows our simulation on how a group of "rogue actors" can affect the social network. These "rogue actors" nodes are represented in color red. This shows that a small group of them can have quite a spread impact on the network. Note that this dataset is not a full dataset from Facebook - it is collected from SNAP survey participants on Facebook. However, a small dataset can be observed quite complex.
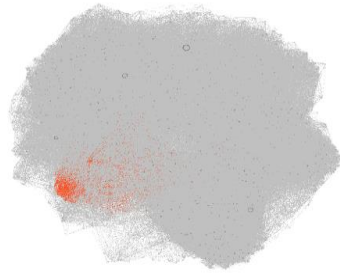
Fig. 4 Simulation on how a group of "rogue actors" can affect the social network

The questionnaire was distributed online and conducted via Google Forms. Over 35 respondents responded to the questionnaire. Questions asked were about their online social network service they are using, how do they use, how many connections they have, how do they react with unknown person on social network, and so on. The results are shown in the charts and graphs below. Based on the chart above in Fig. 5 asking about social network that is being used by the respondents and Facebook has the highest amount of users based on the survey done. It clearly shows that Facebook is still the most popular social network website that is use among the people.
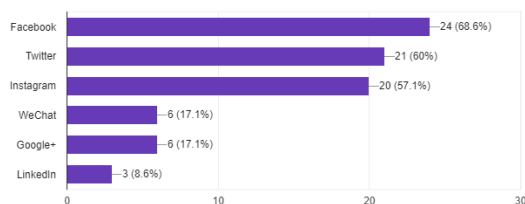
Fig. 5 Social Media used

The pie chart above in Fig. 6 shows that many people have many friend on social networking website and it can be clearly seen that many spend time on social networking site making new friends.
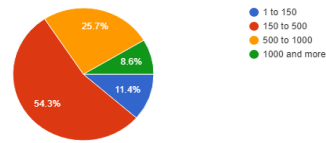
Fig. 6 Number of friends on social networking site

Fig. 7 shows that most of the respondents accepts friend requests even if they don't know the person well. Although the difference between those who answered Yes and No is quite minimal, this shows that there are people still doing that practice.
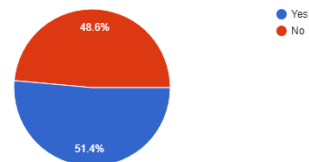
Fig. 7 Friend request acceptance

Fig. 8 shows that most of the respondents do get a private message from unknown person on online social networks.
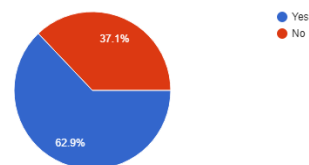
Fig. 8 Receiving private messages from unknown person on online social networks

Fig. 9 shows that many messages that is being sent to the user are mostly an advertisement with a link attached to lure the users to go to another website. The next is the attempt to get to know you more which is very common in social networking sites because is a platform to get to know more new friends. The last is just a simple link without any description.

If yes, how is the message typically be?
28 responses

- An advertisement (with a link attached)
- An attempt to get to know you more
- Just a link without any description
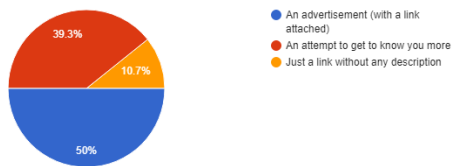
Fig. 9 The typical messages

Fig. 10 shows that many users will still click on the link even is a banking website. It shows that many users will click if the link looks trustworthy. Which the yes has the highest user.

Would you click links that appear to be a website that you familiar of such as banking website, even if its coming from your friend?
35 responses

- Yes
- No
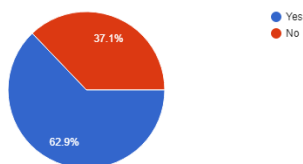
Fig. 10 Clicking links on website

Fig. 11 shows that many users are being added into unknown groups by other people that they do not even know.

Have you ever encountered a situation where you have been added into a group chat by unknown person without knowing?
35 responses

- Yes
- No
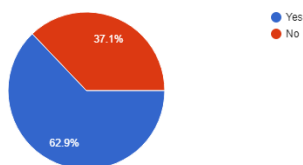
Fig. 11 Added to group chat by unknown person

Fig. 12 shows that there are people who are being tagged by others into a picture which their faces are not even in. This shows that many people are being by others as a way of advertising their product.

Does anyone ever tagged you in pictures even though you are not in the picture
35 responses

- Yes
- No
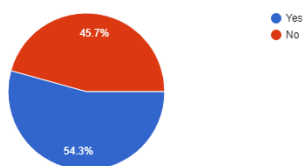
Fig. 12 Anonymous tagging

Fig. 13 shows that they are many people who use Facebook do not set their post to public because and prefer to be private. This is a good measure of the users to prevent unwanted use for comments by unwanted users.

If you are using Facebook, do you set your privacy settings on your posts and profile details visible to Public?
35 responses

- Yes
- No
- Not sure
- I am not using Facebook
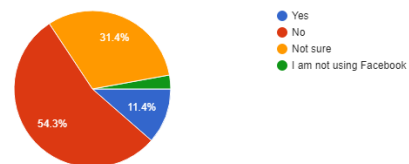
Fig. 13 Privacy settings

Fig. 14 shows that many people do not check and see who is the publisher of the content creator before allowing others to access their Facebook page and content.

If you are using an app on Facebook (such as quizzes or games), do you take look at the app publisher details and read its Privacy Policy?
35 responses

- Yes
- No
- I am not using Facebook
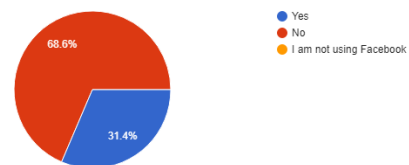
Fig. 14 Publisher details and privacy policy

Fig. 15 shows that many have not being cyber bullied on social network sites and that is a good thing.

Have you ever get cyberbullied on social network site? (Illicit comments on your posts, harsh words, etc)
35 responses

- Yes
- No
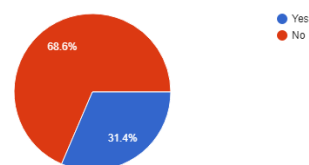
Fig. 15 Cyberbullied

Fig. 16 shows that they are many users who will just accept friend request without thinking twice whether the person is real or not. This is pretty dangerous as this can allow hackers to compromise other users.

Raenu Kolandaisamy, Heshalini Rajagopal, Indraah Kolandaisamy

If you get a friend request from your friend that claims it is his/her second account, would you still accept it?
35 responses
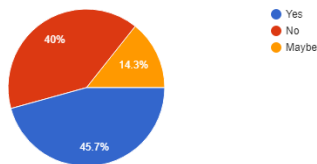
Fig. 16 Accepting anonymous friend request

Fig. 17 shows that many users are still very unaware that cyber crime is dangerous and this users will require need to have more knowledge if they want to keep themselves safe.

Do you aware about cyber crime incidents happening in the cyberspace?
35 responses
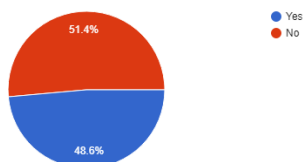
Fig. 17 Awareness on cyber crime

Fig. 18 shows that many are still in the middle and some still think that social networking websites is still unsafe and is dangerous to many people.

How confident do you feel about security measures you have taken when using online social networks?
35 responses
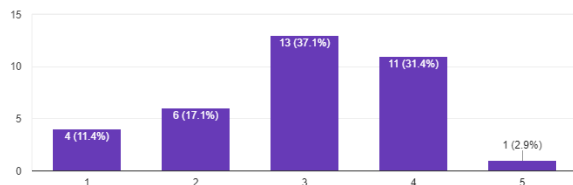
Fig. 18 Safety of social networking sites

5. Conclusion

Social networking has been really a good way to communicate with friends and family and thus make life happier and more joyful. This is because of the convenience that is offer and many users love using it and have spend hours and hours a day just to go through social media and looking at post posted by others. The security of this social medias websites has improved over the years but there are still many loopholes that be compromised by the hackers because of the lack of knowledge that the users have and thus making users more vulnerable to attacks. One common mistakes that

most users make is to create a password that is similar to their birthday date or telephone number and that makes easy for many to easily compromise his or her account. Besides that, talking to an unknown stranger can also be very dangerous especially those who are seeking for new friends and are desperate, they tend to do things without thinking of the danger.

From the findings gathered in the previous section, it is apparent that because of the way on how online social networks users uses the platform, it contributes to the emergence of the cyber crime in online social networks.

References

1. What is social networking? - Definition from WhatIs.com", WhatIs.com, 2018. [Online]. Available: http://whatis.techtarget.com/definition/social-networking. [Accessed: 07- Mar- 2018].
2. J. Constine, Facebook now has 2 billion monthly users… and responsibility, TechCrunch, 2018. [Online]. Available: https://techcrunch.com/2017/06/27/facebook-2-billion-users/. [Accessed: 07- Mar- 2018].
3. M. Fita, M. Fita and M. Fita, 6 Reasons Why Social Networking is So Popular These Days |, Brandignity.com, 2018. [Online]. Available: https://www.brandignity.com/2012/11/6-reasons-why-social-networking-is-so-popular-these-days/. [Accessed: 07- Mar- 2018].
4. July 2017 Cyber Attacks Statistics, HACKMAGEDDON, 2018. [Online]. Available: https://www.hackmageddon.com/2017/08/24/july-2017-cyber-attacks-statistics/. [Accessed: 07- Mar- 2018].
5. Manuel Egele, Gianluca Stringhini, Christopher Kruegel, and Giovanni Vigna, Towards Detecting Compromised Accounts on Social Networks.
6. Kolandaisamy, R., Noor, R. M., Zaba, M. R., Ahmedy, I., & Kolandaisamy, I. (2019, July). Markov chain based ant colony approach for mitigating DDoS attacks using integrated vehicle mode analysis in VANET. In 2019 IEEE 1st International Conference on Energy, Systems and Information Processing (ICESIP) (pp. 1-5). IEEE.
7. A. Lavorgna, "Organised crime goes online: realities and challenges", Journal of Money Laundering Control, vol. 18, no. 2, pp. 153-168, 2015.
8. R. B. R. M. P. Patel and M. H. Joshi, "Privacy and Security Issues in Social Online Networks,", National Conference on Latest Trends in Networking and Cyber Security, 2017.
9. E. Leukfeldt, E. Kleemans and W. Stol, "Cybercriminal Networks, Social Ties and Online Forums: Social Ties Versus Digital Ties within Phishing and Malware Networks", British Journal of Criminology, p. azw009, 2016.
10. J. Leskovec and A. Krevl, SNAP Datasets: Stanford Large Network Dataset Collection, 2014.

Authors Introduction

Dr. Raenu Kolandaisamy

He received his PhD from the Faculty of Computer Science & Information Technology, University Malaya in 2020. He is currently an Assistant Professor in UCSI University, Malaysia. His research interest areas are Wireless Networking, Security, VANET and IoT.

Dr. Heshalini Rajagopal

She received her PhD and Master's degree from the Department of Electrical Engineering, University of Malaya, Malaysia in 2021 and 2016, respectively. She received the B.E (Electrical) in 2013 Currently, she is an Assistant Professor in UCSI University, Kuala Lumpur, Malaysia. Her research interest includes image processing, artificial intelligence and machine learning.

Dr. Indraah Kolandaisamy

She is a senior lecturer in School of Business Management, College of Business, Universiti Utara Malaysia. Indraah A/P Kolandaisamy holds a doctorate in Management from Universiti Kebangsaan Malaysia in 2015. Her D.B.A work is on organizational citizenship behavior among public sector in Malaysia. She obtained her MSc (Management) and Bachelor in International Business Management (Hons) from Universiti Utara Malaysia respectively in 2007 and 2005.

Development of a music recommendation application by using facial emotion recognition

Shengke Xie, **Raenu Kolandaisamy, Ghassan Saleh**
Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia

Heshalini Rajagopal
Department of Electrical and Electronics Engineering, MILA University, 71800 Nilai, Negeri Sembilan, Malaysia
E-mail: raenu@ucsiuniversity.edu.my

Abstract

Music is an important part of human life and culture, and it can affect people's emotions and moods. However, choosing music from a large library can be a challenging and time-consuming task. In this paper, we propose a facial expression recognition-based music recommendation system that can recommend suitable music which matches the user's current mood. The system uses a camera to capture the user's face and a convolutional neural network model which trained facial emotion recognition database to recognize seven basic emotions: anger, disgust, fear, happiness, sadness, surprise and neutral. The paper contributes to the research of facial emotion recognition and music recommendation and provides a convenient way for people to enjoy music.

Keywords: Facial Emotion Recognition, Music Recommend Systems, Convolutional Neural Network

1. Introduction

With the development of the Internet of Things (IoT) technology, various smart homes are entering the public eye, and AI voice assistants used to unify the management of various smart home appliances have also emerged. People control the use of various electrical appliances and furniture in the same network by giving commands to the AI voice assistants with their words, while many AI voice assistants have added many other functions in addition to the basic home control functions in order to increase the diversity of functions, such as music recommendations. When people are looking for music to listen to, they are often confused about the choice of music and are looking for some suggestions to help them choose, since it is often difficult to choose the most suitable music from a large library of thousands of songs, and music recommendations are one such feature that can reduce the difficulties users have in choosing music [1]. However, the music recommendation function sometimes does not meet the user's needs, and there are often situations where the recommended music conflicts with the user's current mood, such as recommending sad songs when the user is happy, or recommending songs with a strong rhythm when the user needs to calm down [2]. Therefore, how to make the computer quickly and accurately recommend the right song to the user is a problem that needs to be solved. It should be noted that facial expressions account for two-thirds of human communication and are one of the most important means of expressing human emotions. In the absence of facial expression recognition, AI is sometimes unable to give correct and rapid feedback through speech alone. For example, the same phrase what's up may have different meanings depending on the emotion, it may be a greeting between acquaintances, or it may be a concern to ask what difficulties you are experiencing, so people hope that computers will be able to understand human emotions and give correct feedback, bringing people a better experience with more intelligent human-computer interaction [3].

For this problem, we will use a camera to capture the user's image, use python deep learning based facial expression recognition technology to identify the user's current mood, and select the music that matches the user's current mood from the music categories in the database that apply to different moods to recommend music to the user, thus trying to reduce the user's anxiety in choosing music and improve the user's experience

2. Literature Review

Communication is a bridge to build interpersonal relationships, people send or obtain information through communication, and facial expressions are an important means and factor to help people understand the information conveyed by others in communication, according to the survey [4] the non-verbal component of communication reaches about 55% of interpersonal communication, and as an important category of non-verbal component, the study of human facial expressions is not only in medicine and psychology, its

wide application prospects have also led to its widespread interest in computer science [5].

As early as 1971, two American psychologists, Ekman and Friesen, had systematically studied facial expressions and in 1978 developed and defined the Facial Action Coding System (FACS), a system for recognizing various human emotions by representing different facial expressions in terms of different facial muscle changes. The system encodes specific facial muscle changes, called Action Units (AUs). Fig. 1 shows the main types of Aus.

Action Unit	Description	Facial Muscle
AU0	Neutral Face	
AU1	Inner Brow Raiser	frontalis (pars medialis)
AU2	Outer Brow Raiser	frontalis (pars lateralis)
AU4	Brow Lowerer	depressor glabellae, depressor supercilii, corrugator supercilii
AU5	Upper Lid Raiser	levator palpebrae superioris, superior tarsal muscle
AU6	Cheek Raiser	orbicularis oculi (pars orbitalis)
AU7	Lid Tightener	orbicularis oculi (pars palpebralis)
AU8	Lips Toward Each Other	orbicularis oris
AU9	Nose Wrinkler	levator labii superioris alaeque nasi
AU10	Upper Lip Raiser	levator labii superioris, caput infraorbitalis
AU11	Nasolabial Deepener	zygomaticus minor
AU12	Lip Corner Puller	zygomaticus major
AU13	Sharp Lip Puller	levator anguli oris (also known as caninus)
AU14	Dimpler	buccinator
AU15	Lip Corner Depressor	depressor anguli oris (also known as triangularis)
AU16	Lower Lip Depressor	depressor labii inferioris
AU17	Chin Raiser	mentalis
AU18	Lip Pucker	incisivii labii superioris and incisivii labii inferioris
AU19	Tongue Show	
AU20	Lip Stretcher	risorius
AU22	Lip Funneler	orbicularis oris
AU23	Lip Tightener	orbicularis oris
AU24	Lip Pressor	orbicularis oris
AU25	Lips Part	depressor labii inferioris, or relaxation of mentalis or orbicularis oris
AU26	Jaw Drop	masseter; relaxed temporalis and internal pterygoid
AU27	Mouth Stretch	pterygoids, digastric
AU28	Lip Suck	orbicularis oris
AU41	Lid Droop	relaxation of levator palpebrae superioris
AU42	Slit	orbicularis oculi
AU43	Eyes Closed	relaxation of levator palpebrae superioris
AU44	Squint	orbicularis oculi, pars palpebralis
AU45	Blink	relaxation of levator palpebrae and contraction of orbicularis oculi, pars palpebralis.
AU46	Wink	levator palpebrae superioris; orbicularis oculi, pars palpebralis

Fig. 1. The main types of Aus

It is well-known that human beings have six basic emotions (BEs), namely happiness, surprise, anger, sadness, fear and disgust, on the basis of which Du et al [6] have proposed 22 compound emotions (CEs) made up of combinations of basic emotions with each other, and different combinations of AU classify these facial emotions to facilitate identification, e.g. AU combinations 1, 4, 20, 25 can be identified as fear

emotions. Fig. 2 shows the main types of human facial emotions represented by the various AUs combinations.

Category	AUs	Category	AUs
Happy	12, 25	Sadly disgusted	4, 10
Sad	4, 15	Fearfully angry	4, 20, 25
Fearful	1, 4, 20, 25	Fearfully surprised	1, 2, 5, 20, 25
Angry	4, 7, 24	Fearfully disgusted	1, 4, 10, 20, 25
Surprised	1, 2, 25, 26	Angrily surprised	4, 25, 26
Disgusted	9, 10, 17	Disgusted surprised	1, 2, 5, 10
Happily sad	4, 6, 12, 25	Happily fearful	1, 2, 12, 25, 26
Happily surprised	1, 2, 12, 25	Angrily disgusted	4, 10, 17
Happily disgusted	10, 12, 25	Awed	1, 2, 5, 25
Sadly fearful	1, 4, 15, 25	Appalled	4, 9, 10
Sadly angry	4, 7, 15	Hatred	4, 7, 10
Sadly surprised	1, 4, 25, 26	-	-

Fig.2. The AUs combinations that can be observed in different human facial emotions [5]

3. Method

This system implements a music recommendation system based on facial expression recognition on the PC. Fig. 3 shows the steps for facial recognition.

Fig.3. Steps for facial emotion recognition

The aim of this system is to use the computer's powerful computing resources to quickly and accurately understand human emotions and recommend more appropriate songs to the user, bringing a smarter experience to people.

The system was developed on a PC with Windows 10, 64-bit operating system, Intel-i7-10510U CPU, 16G Byte RAM, PyCharm 2022.2.2, OpenCV-3.4.4.19 compilation environment and python as the main programming language. Tests of the system were also carried out on this machine.

The process of using this system can be simply summarized as capturing the user's image using the local camera, using deep learning based facial expression recognition technology to identify the user's current mood, and recommending music that matches the user's current mood from the downloaded music that has been categorized into different mood folders, thus trying to reduce the user's anxiety in choosing music and improve the user's experience. Fig.4 shows the workflow of this system.
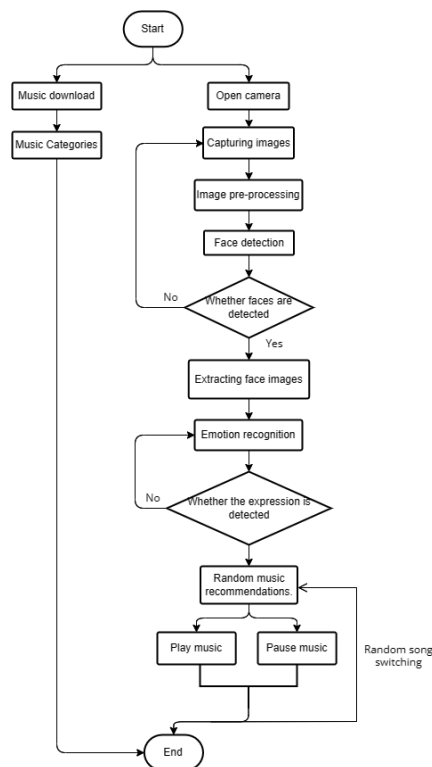
Fig. 4. System workflow diagram

According to the need, the design of this system as shown in Fig. 5 is mainly divided into seven modules such as Image pre-processing, Face detection, Emotion recognition, Music download, Music classification, Music recommendation and Music playing:

1) Image pre-processing module: the image captured by the local camera is pre-processed, the pre-processing process is mainly to convert the colour image into a grey scale image, which is convenient for face detection.

2) Face detection module: face detection is performed on the grayscale image, after the face is detected, a rectangular box is drawn on the original colour image to frame out the face, in preparation for expression recognition.

3) Emotion recognition module: The detected face image is fed into the expression recognition model for expression recognition. The recognition model is a convolutional neural network model with parameters trained to recognize the face expression and return the result.

4) Music download module: The crawler crawls the MP3 file of the searched music from the music playing website and downloads it to the local area, while crawling the corresponding lyrics and saving them in txt format, so as to prepare for music classification.

5) Music classification module: count the emotion value of each emotion word in the text according to the emotion dictionary, and classify the music files corresponding to the lyrics according to the emotion value.

6) Music recommendation module: according to the expression recognition result, a song is randomly loaded from the local music folder of the corresponding emotion

7) Music playing module: play or pause the loaded song.

Among them, the core function of the system is Emotion recognition, the recognition result will be directly related to the quality of the system, the appropriateness of the recommended music.
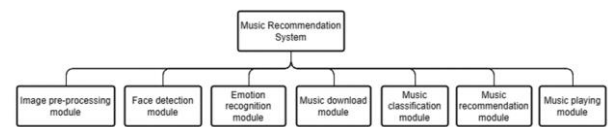
Fig.5. Structure diagram of the system modules

3.1. User Interface Design

A clean and aesthetically pleasing user interface (UI) provides a good user experience and helps users to interact easily with the software system. Here the UI is designed by using PyQt5, a GUI framework for the Python programming language, which makes it easy to implement powerful user interfaces using a variety of UI controls. The main controls used in this system are Label, Button, Line Edit and Timer.

Fig. 6 show the design blueprint of the UI and screenshots of the actual UI interface respectively, with the functions of each UI control as follows:

1) close_btn: Button control, used to close the whole system.

2) title_label: Label control for scrolling the name of the currently playing music file.

3) result_label: Label control, used to display the result of face expression recognition.

4) camera_btn: Button control for turning the local camera on and off.

5) play_btn: Button control to control the playing and pausing of the currently loaded music.

6) next_btn: Button control for randomly loading a piece of music from the folder matching the expression recognition result.

7) detect_btn: Button control that detects the user's current expression, returns the result and displays it in the result_label.

8) searchMusic_LEdit: Line Edit control for entering music names in preparation for music downloads.

9) download_btn: Button control for crawling and downloading mp3 files and lyric content of the music entered by the user.

10) camera_label: Label control to display the image captured by the camera.
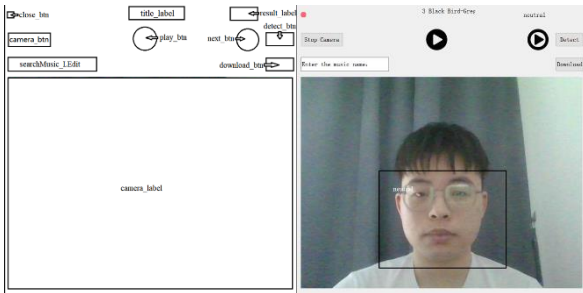
Fig.6. Design of the UI

3.2. *Design of Convolutional Neural Networks*

The design of the network structure affects the model performance and the accuracy of the classification. A good network structure can extract features more efficiently and reduce the training time. This system uses the convolutional neural network structure designed by Zhou [7] as the basis for classification. It achieved good recognition results on both the FER-2013 test set and the dataset [8]. The structure of this convolutional neural network consists of a 48×48 input layer and a 1×1 convolutional layer to form the first group of layer sets, and two consecutive convolutional layers with filter sizes of 3×3 and 5×5 followed by a 2×2 maximum pooling layer to form the second and third group of layer sets. Each convolutional layer uses ReLU as the activation function. The second group of layers has one more convolutional layer than the first group, which means that further refinement of features can be learned based on the first group of layers. The detailed structure of the network parameters is shown in Fig. 7.

Layer	Number of filters	Filter size	Stride	Padding	Feature image size
input	0	0	None	None	(48,48,1)
conv1-1	32	1×1	1	0	(48,48,32)
conv2-1	64	3×3	1	1	(48,48,64)
conv2-2	64	5×5	1	2	(48,48,64)
pool2	0	2×2	2	0	(24,24,64)
conv3-1	64	3×3	1	1	(24,24,64)
conv3-2	64	5×5	1	2	(24,24,64)
pool3	0	2×2	2	0	(12,12,64)
fc1	None	None	None	0	(1,1,2048)
fc2	None	None	None	0	(1,1,1024)
output	None	None	None	0	(1,1,8)

Fig.7. Network structure parameters [7]

3.3. Design and implementation of the main functions of the system

The music recommendation system based on facial expression recognition is divided into seven modules, and the main functions of each module are briefly introduced above.

Image pre-processing module

In order to ensure consistency in face size and position as well as face image quality in face images and to improve the efficiency and accuracy of face recognition, the images acquired by the camera need to be pre-processed prior to face detection, which can significantly reduce the computational effort of the device. The main task of pre-processing is image greyscale, which means that the image is unified into a grey-scale map of a specified size. In actual use will meet the image acquisition effect is not good, such as lighting and other factors cause image results are not good, and ultimately affect the recognition results, and the grayscale process can effectively remove these noise influence, after the grayscale, the influence of the noise is reduced to a minimum.

The method first normalises the input raw face image, scaling the pixel value range from 0-255 to between 0-1, then scales the raw face image to a specified image size of 48×48 and generates a set of face images of different sizes and orientations, providing more information for subsequent tasks such as face recognition and expression recognition, thus improving the accuracy and robustness of the algorithm.

The greyscale of the images is achieved using the COLOR_BGR2GRAY method in OpenCV.

Face detection module

The prerequisite for expression recognition is that the recognized image contains a face component, so face detection is a powerful guarantee for expression recognition. This module uses the face_detection_front.tflite and face_detection_back.tflite models from [8] for face detection.

The method first initialises the face detector, uses the detectFaces() method which from detector to detect faces in the input image, and returns a detection result containing the bounding box, key points and scores for each detected face. Finally, the returned detection results are used to obtain and return the absolute coordinates of the face bounding boxes.

Emotion recognition module

Emotion recognition is the core function of this system, image pre-processing, face detection modules are all paving the way for the implementation of this module.

In this method we first load a pre-trained facial emotion recognition model, read the video frames from the local camera using the capture.read() method in the OpenCV computer vision library, resize the video frames and greyscale them. Next, the faces returned by the face detection module are recognized for their emotions using the predict() method, and the predicted results are annotated in the original image and displayed in the control camera_label.

Music download module

The music download module mainly uses python's powerful crawler function to crawl music from music websites. The module consists of three main parts: music search, music crawling and music saving. Firstly, the searchSong() method is used to send an HTTP GET request to the url with the search content, and return a list of songs in JSON format to be prepared for the music download.

Then the MusicDownload() method uses the returned JSON list of songs to extract the song information for the first song in the list, including the song name, artist and song ID. Then the downloadSong() method is called and passed the extracted song information.

Finally, in the downloadSong() method, a request is sent to the url consisting of the song ID and a response is received. The music content is found in the JSON format response, read and saved as an MP3 file in the local download folder, thus completing the download of the music and also preparing the music for classification.

Music classification module

In the music classification module, we will classify the downloaded music according to the corresponding emotion and transfer the files to the corresponding emotion folder. For this function we rely on a sentiment dictionary from [9], which assigns sentiment values to each English word and the corresponding Chinese word for eight basic emotions (anger, fear, expectation, trust,

surprise, sadness, joy and disgust) and two emotions (negative and positive), and saves them in a csv file for easy searching. We used this sentiment dictionary to analyse the lyrics of the song to find the emotion with the largest sentiment value among all the emotions in the text of the lyrics as the final classified emotion, thus determining the sentiment classification of the song.

The code first read the sentiment dictionary and map the words to their corresponding sentiment values. Then the classify_sentiment() method below counts the number of each sentiment word in the input text and the total sentiment value based on the sentiment value of the word in the dictionary, thus returning the sentiment with the highest sentiment value as the result.

Music recommendation module

The music recommendation module is mainly based on the result of face emotion recognition, the music in the emotion folder corresponding to the recognition result will be played randomly using the method in the random library in python. The next_song() method in this module is also bound to the control next_btn, which is also a method for randomly switching music.

The code first opens the path to the folder where the emotional music of the category is stored based on the results of the face emotion recognition, searches the folder for files ending in '.mp3', '.wav', '.ogg ' or '.m4a' extension, saves the path to one of the files at random, and finally loads and plays the music using the method in the pygame library.

Music playing module

The Music playing module is used to control the playing and pausing of loaded music, and is implemented by the methods provided in the pygame library.

This method uses the Boolean variable ‘playing’ to determine if the music is playing and to control the playing and pausing of the music. When the playing variable is false, the music is not playing and can be played via the button, when the playing variable is true, the music is already playing and can be paused via the button.

4.   Result and discussion

4.1. *Under different light conditions*

The tests were first carried out under different lighting conditions and the results are shown in Figs. 8

and 9. As can be seen from the above diagram, under different lighting conditions, Fig. 8 shows the condition of insufficient lighting and Fig. 9 shows the condition of sufficient lighting, it can be seen from the figures that the function of the system expression recognition is not affected by the lighting, but the accuracy of expression recognition may be affected and reduced as a result.
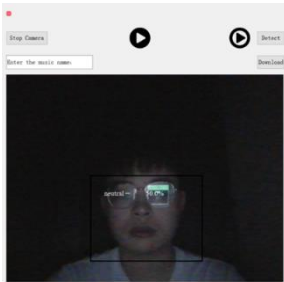
Fig.8. Insufficient light
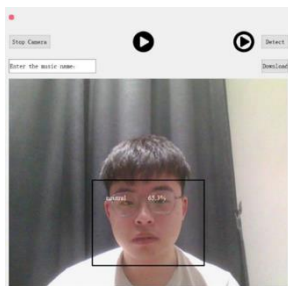
Fig.9. Sufficient light

4.2. *Testing at different distances*

The next test was conducted for different distances, with the face being closer to the camera and further away. The test results are shown in Figs.10 and 11.
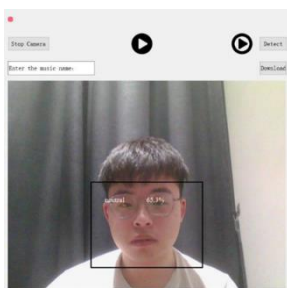
Fig.10. Distance from camera 0.3m
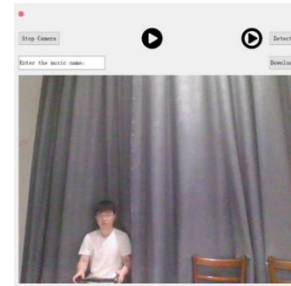
Fig.11. Distance from camera 2m

Fig. 10 shows the results of the test at a distance of 0.3m from the camera and Fig. 11 shows the results of the test at a distance of 2m from the camera. After testing, it was found that the recognition of faces within two metres from the camera would not be affected; however, once the face is more than two metres away from the camera, the recognition will be affected and the face may not be detected, so when using this system, the user needs to be as close to the camera as possible so that the face can be better captured for recognition.

4.3. *Recognition accuracy test*

In this study, the 7 categories of "neutral", "surprise", "happy", "angry", "sad", "fear" and "disgust" were made in turn as shown in Fig. 12. The first five expressions matched the actual expressions, while the last two did not match the actual expressions. The highest recognition accuracy was achieved for the "happy" and "surprise" expressions, with over 90%, and for the "neutral", "sad" and "angry" expressions, with over 60%. The intended "fear" expression was identified as "sad" and the intended "disgust" expression was identified as "fear". The reason for this may be, on the one hand, the low recognition rate of these two expressions by the model, and on the other hand, the ambiguity of these two expressions, which could also be identified as "sad" and "fear" if judged by the human eye.
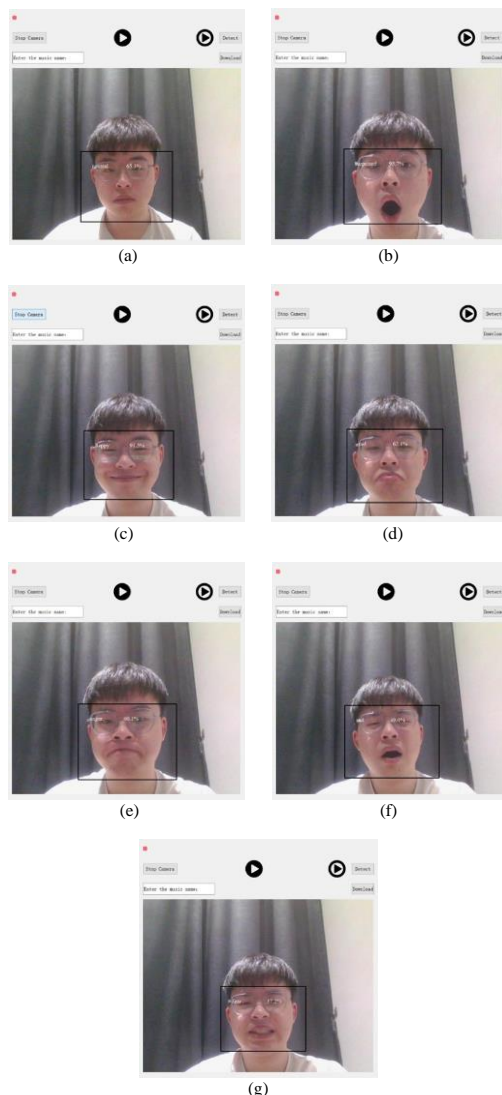
Fig. 12 (a) Neutral, (b) Surprised, (c) Happy, (d) Sad, (e) Anger, (f) Fear, (g) Disgusted expressions

5. Conclusion

Facial expression recognition has application needs in many scenarios, and today more and more fields have introduced expression recognition technology, such as the education field and the medical field. In order to meet the requirements of a facial expression recognition based music recommendation system, a convolutional neural network model based on the FER-2013, JAFFE, CK+ expression datasets was used in this project. A music recommendation system based on facial emotion recognition was successfully designed and implemented in the PyCharm development environment using the python programming language. It mainly includes seven

modules: image pre-processing, face detection, emotion recognition, music download, music classification, music recommendation and music playing. After the design was completed, the system was also tested in actual use. The results showed that the system has a high accuracy rate of expression recognition and is highly practical. However, the expression recognition function in this system is still inadequate for the occlusion, side face and distance cases. At a later stage, it is considered to add the expression images for the occlusion and side face cases to the existing dataset to improve the face detection rate and thus the expression recognition rate. In addition, the music classification function of the system designed in this project is limited to classifying music with lyrical content. In the future, the authors will consider improving the music classification algorithm and adding the classification of pure music without lyrics to the music classification module of this system.

References

1. S. Metilda Florence and M. Uma, 'Emotional Detection and Music Recommendation System based on User Facial Expression', IOP Conf. Ser.: Mater. Sci. Eng., vol. 912, no. 6, p. 062007, Aug. 2020, doi: 10.1088/1757-899X/912/6/062007.

2. Song, Y., Dixon, S., & Pearce, M. (2012, June). A survey of music recommendation systems and future perspectives. In 9th international symposium on computer music modeling and retrieval (Vol. 4, pp. 395-410).

3. D. O. Melinte and L. Vladareanu, 'Facial Expressions Recognition for Human–Robot Interaction Using Deep Convolutional Neural Networks with Rectified Adam Optimizer', Sensors, vol. 20, no. 8, Art. no. 8, Jan. 2020, doi: 10.3390/s20082393.

4. K. Kaulard, D. W. Cunningham, H. H. Bülthoff, and C. Wallraven, 'The MPI Facial Expression Database — A Validated Database of Emotional and Conversational Facial Expressions', PLOS ONE, vol. 7, no. 3, p. e32321, Mar. 2012, doi: 10.1371/journal.pone.0032321.

5. B. C. Ko, 'A Brief Review of Facial Emotion Recognition Based on Visual Information', Sensors, vol. 18, no. 2, Art. no. 2, Feb. 2018, doi: 10.3390/s18020401.

6. S. Du, Y. Tao, and A. M. Martinez, 'Compound facial expressions of emotion', Proceedings of the National Academy of Sciences, vol. 111, no. 15, pp. E1454–E1462, Apr. 2014, doi: 10.1073/pnas.1322355111.

7. Kolandaisamy, R., Subaramaniam, K., & Jalil, A. B. (2021, March). A Study on Comprehensive Risk Level Analysis of IoT Attacks. In 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS) (pp. 1391-1396). IEEE.

8. M. Lyons, S. Akamatsu, M. Kamachi, and J. Gyoba, 'Coding facial expressions with Gabor wavelets', in Proceedings Third IEEE International Conference on

Automatic Face and Gesture Recognition, Apr. 1998, pp. 200–205. doi: 10.1109/AFGR.1998.670949.

9. https://github.com/luanshiyinyang/FacialExpressionReco gnition

---

Authors Introduction

Shengke Xie
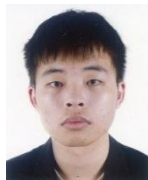
He has completed his Bachelor of Computer Science (Hons) Mobile Computing and Networking, UCSI University, Malaysia.

Dr. Raenu Kolandaisamy

He received his PhD from the Faculty of Computer Science & Information Technology, University Malaya in 2020. He is currently an Assistant Professor in UCSI University, Malaysia. His research interest areas are Wireless Networking, Security, VANET and IoT.

Dr. Ghassan Saleh Hussein Al-Dharhani

He received his Ph.D. degree in Computer Science from Universiti Kebangsaan Malaysia (UKM). He is currently an Assistant Professor with the Institute of Computer Science and Digital Innovation, UCSI University, Malaysia. His research interests include Artificial Intelligence, Data Mining and Knowledge Discovery, etc.

Dr. Heshalini Rajagopal

She received her PhD and Master's degree from the Department of Electrical Engineering, University of Malaya, Malaysia in 2021 and 2016, respectively. She received the B.E (Electrical) in 2013 Currently, she is an Assistant Professor in UCSI University, Kuala Lumpur, Malaysia. Her research interest includes image processing, artificial intelligence and machine learning.

A Model of Reaction-diffusion phenomena with Multiset Processing

Yasuhiro Suzuki

Nagoya University, Furocho, Chikusa-ku, Nagoya, 464-8601, Japan

Abstract

We propose a model of reaction-diffusion phenomena using Abstract Rewriting System on Multisets ARMS, which is a model of Multiset Processing. Although proposed model is simple, computer simulations confirm that the Turing pattern is generated.

Keywords: Reaction-Diffusion, Activator-Inhibitor, Multiset Processing , Abstract Rewriting System on Multisets, ARMS

1. Introduction

A reaction-diffusion system is a mathematical model of how the concentration of one or more substances distributed in space changes under the influence of two processes: local chemical reactions, in which substances change each other, and diffusion, in which substances spread in space. One of the most famous reaction-diffusion systems is that proposed by Alan Turing. Turing showed a system that is locally stable but destabilized by diffusion. The spatio-temporal pattern that occurs in this system is called the Turing pattern [1].

The reaction diffusion model, which Alan Turing proposed is one in which two oscillators are coupled by diffusion (Eq. 1), where the two oscillators mutually activate and inhibit each other.

$$\frac{\partial u}{\partial t} = f(u, v) + D_u \nabla^2 u,$$

$$\frac{\partial v}{\partial t} = g(u, v) + D_v \nabla^2 v, \tag{1}$$

In the equation (1), $u \equiv u(\boldsymbol{r}, t), v \equiv v(\boldsymbol{r}, t)$, f, g are reaction terms, D is a diffusion coefficient and ∇^2 is Laplacian.

2. Methodology

We model a system that performs activation and inhibition by diffusion coupling with two oscillators by a multiset rewriting system, Abstract Rewriting System on Multisets, ARMS.

2.1. Abstract Rewriting System on Multisets, ARMS

Abstract Rewriting System on Multisets, ARMS is a multiset rewriting system [2]. A multiset is defined as a simple set and a map, which returns the duplication of element. We denote the duplication (multiplicity) of an element as M (a), for a $\in$ A and in case c $/\in$ A, M (c) = 0; for example M (a) and M (b) of {a, a, b, b} are 2, and M (c) = 0; in the mathematical description, a multiset is described as; < sup, M () >, in which sup is a simple set of elements, in this paper we describe a multiset by denoting the same alphabet in its number of multiplicity such as {a, a, b, b} or a vector w =(M (a1) M (a2) . . . M (an)).

The union of two multisets M1, M2 is the same as the union of simple set and in vector description, the union of multisets is addition of vectors w1 and w2. And inclusion of sets is also the same as the simple set, when M1(a) $\le$ M2(a) for all a $\in$ A, the multiset M1 is included in M2 and we write M1 $\subseteq$ M2.

A reaction rule is a pair of multiset, we denote A# as a set of all combinations of multisets over A and in the combinations, an empty multiset is included. A reaction rule l $\rightarrow$ r, l, r $\in$ A# is described as a pair of multiset likewise chemical equations or a pair of its vector expression; and in some case, we can describe a reaction

rule as a vector r, r = −l + r, it is simple and good for examining the dynamics of an ARMS, but this description can not illustrate when there are the same species of element in the left-hand side and right-hand side such as a, b → a, c; in this case l = (1, 1, 0) and r = (1, 0, 1) and r = −(1, 1, 0) + (1, 0, 1) = (0, −1, 1).

A reaction is described as the rewriting of a multiset, if the left-hand side of a reaction rule is included in a multiset, these elements in the multiset are excluded and the right-hand side of the rule is merged to the multiset; the case when the multiset is a, a, b, b and the reaction rule is a, b → c, d, the left-hand side of the rule is included in the multiset, {a, b} ⊆ {a, a, b, b} so the {a, b} is excluded from the multiset and it is transformed to {a, b} and the left-hand side of the rule {c, d} is merged to the set and we obtain {a, b, c, d} by this reaction. By using vector ex- pression, a reaction is the addition of vectors as w − l + r,

in which l and r are vector expression of a reaction rule; the case when M = {a, a, b, b}, w = (2, 2, 0, 0) and the reaction rule is a, b → c, d, ({1, 1, 0, 0}, {0, 0, 1, 1}); and the reaction is denoted as {2, 2, 0, 0} − {1, 1, 0, 0} ={1, 1, 0, 0} and {1, 1, 0, 0} + {0, 0, 1, 1} = {1, 1, 1, 1}.

2.2. Activator-Inhibitor model

In the activator-inhibitor model, activator X increases themselves by autocatalytic reaction (Eq. 2); X → X, X. The reaction of activator X produces inhibitor Y is, X -> Y, X. X is put in both reaction equations because X does not increase by this reaction, while only Y increases. X, Y -> Y describes the inhibition of increasing X by Y, where X decreases while Y does not change. Y → ∅ denotes the decreases of Y, ∅ illustrates an empty set. We give reaction coefficients for each rule. We give reaction coefficients for each rule as k_1, k_2, k_3, k4. Each coefficient gives the probability of firing of the rule.

$$X \xrightarrow{k_1} X, X : \text{r\_1},$$

$$X \xrightarrow{k_2} Y, X : \text{r\_2},$$

$$X, Y \xrightarrow{k_3} Y : \text{r\_3},$$ \hspace{1cm} (2)

$$Y \xrightarrow{k_1} \emptyset : \text{r\_4}.$$
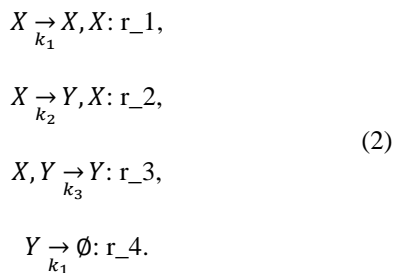
Modelling diffusion :

The Eq. 2 above does not include diffusion. So, we need to consider diffusion separately. Why does diffusion occur? If we put a drop of ink on clean water and watch it for a long time, the ink spreading on the water's surface

will eventually dissolve into the water. No matter how long we keep looking at the surface of the water where the ink has dissolved, there is no return of the drop of ink when the dissolved ink is dropped on the surface of the water again. In other words, diffusion is the process of homogenizing things.

If the state quantities of cells i and j are x_i and x_j, and diffusion occurs between these cells, the amount of diffusion from cell j to cell i can be expressed as d=(x_j - x_i)×D. D is called the diffusion coefficient. For example, let us consider a sequence of cells [2,3,1,2]. (The sequence of cells is indicated by enclosing the amount of state in each cell). For the cell with state 1, the state to the left of this cell is 3, and the state to the right is 2. The state of this cell is less voluminous than the left and right states. Therefore, an inflow occurs from the left and right cells to this cell, changing its state. The degree of inflow is determined by d=(x_j - x_i)×D. If the diffusion coefficient D is 0.1, the inflow from the left neighbour d_left=(3-1)×0.1=0.2 and the inflow from the right neighbour d_right=(2-1)×0.1=0.1. Thus, the sequence of cells [2,3,1,2] becomes [2,2.8,1.3,1.9] due to this diffusion. (Here, only the cell with state one is shown, but in reality, other cells change in parallel as well.) In this way, the state of each cell is averaged by diffusion.

2.3 Diffusion in a two-variable system

The active/inhibited system of Eq (2) consists of two variables, X and Y. Therefore, we will use a one-dimensional sequence of X-only cells and a one-dimensional sequence of Y-only cells. For example, if the sequence of cells representing the state of X is [2,3,1,2] and the sequence of cells representing the state of Y is [0,0,1,0], the states of X and Y in the cell located second from the leftmost are (1,1). When the diffusion coefficient of X is 0.1, and that of Y is 0.3, if diffusion occurs in cell (1,1), the state of the cell changes (Again, only cell (1,1) change is shown).

3. Results and Discussion

We assume that the rightmost cell is connected to the leftmost cell; the ten cells are considered to be connected in a ring. We show the result of simulation whose initial states are;

X=[0,10,0,10,0,10,0,10,0,10,0,10],

Y=[0,10,0,10,0,10,0,10,0,10,0,10].

We set the reaction coefficients are k_1=k_2=k_3=k_4=0.01. The state quantity is first updated by diffusion, as described in the previous section. Next, reaction rules are applied in parallel to update the state quantities. The update by diffusion and the update by reaction rules are repeated, and this process is repeated.

When the diffusion coefficients of X and Y are D_X=D_y=0.0, when there is no diffusion, there is no change from the initial state. Such a state is called an equilibrium state (equilibrium means "balanced"). As will be explained in detail in the next section, if X=Y (the amount of state of X is equal to the amount of state of Y), neither the amount of state of X nor the amount of state of Y will change in this reaction system. Therefore, since X=Y holds in the corresponding cell in the initial state, the system would remain in the initial state if there were no diffusion.

What would happen if diffusion were to occur in a state where the reaction is in equilibrium (X=Y)? First, assuming that X and Y diffusion coefficients are the same with D_X=D_y=0.1, X and Y change but always remain X=Y. In other words, the reaction is always in equilibrium. Eventually, X and Y are homogenized by diffusion, and the entire reaction reaches equilibrium. This is true even if the diffusion coefficient is changed as D_X=D_y=0.2,0.3···.

Next, when X's diffusion coefficient is more significant than Y's with D_x=0.3 and D_y=0.1, X and Y become homogenized and almost X=Y, and the reaction approaches equilibrium. On the other hand, when the diffusion coefficient of Y is more significant than X's with D_x=0.01 and D_y=0.3, Y becomes homogenized. However, a larger or smaller pattern appears in X (Fig. 1) shows an example of the results). In this case, X does not equal Y, and X and Y continue to change (increase).
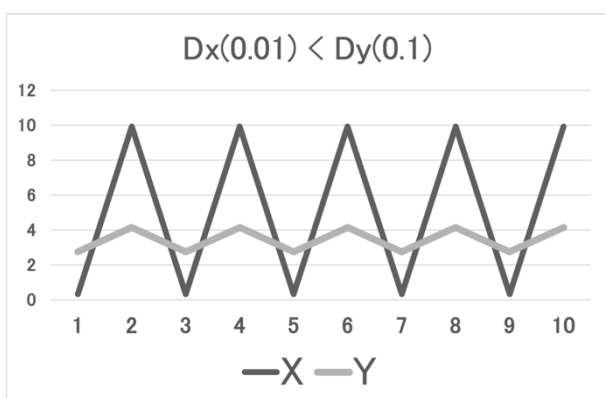
Fig.1 Turing Pattern like behavior, where the diffusion coefficient of X is 0.01, while Y, 0.1.

4. Conclusion

The reaction-diffusion phenomenon has been modelled and investigated as a partial differential equation by modelling the activator-inhibitor system. We used ARMS by modelling the activator-inhibitor system and confirmed that the ARMS model shows Turing pattern-like behavior.

Acknowledgements

References

1. Turing, Alan Mathison. "The chemical basis of morphogenesis." Bulletin of mathematical biology 52 (1990): 153-197.
2. Suzuki, Yasuhiro, and Hiroshi Tanaka. "Symbolic chemical system based on abstract rewriting system and its behavior pattern." Artificial Life and Robotics 1 (1997): 211-219.

Authors Introduction

Dr. Yasuhiro Suzuki

He received his Doctor degree in Informatics in 2001 from Graduate School of Informatics, Kyoto University, JAPAN. He is currently an associate professor of Graduate School of Informatics, Nagoya University, JAPAN

Extract tactile qualities from time series data

Yasuhiro Suzuki

Graduate School of Informatics, Nagoya University, Furocho,
Chikusa, Nagoya 464-8601, Japan

E-mail: ysuzuki@i.nagoya-u.ac.jp
https://ysuzuki.info

Abstract

We proposed the Tactile Score to describe time-varying tactile sensation by the time variation of vertical force. Tactile quality is essential in hardness/softness, roughness, and temperature. Hardness and softness can be extracted from the shape of the Tactile Score. Roughness can be extracted from the pattern of hardness and softness. The arbitrary time series data can be interpreted as a tactile score by considering the time variation of the vertical force, and the hardness and softness are extracted from the time series data interpreted as the tactile score. This method can extract different features from conventional data science methods.

Keywords: Data Science, Time series data, Tactile of Data, Tactile Score

1. Introduction

Sensitivity differs from person to person. The same greeting of "Good morning" may be perceived as cheerful by some and noisy by others. There is no correct answer to sensitivity, nor can it be generalized.

We can generalize if we take the average of many people's sensitivities. Sensitivities that deviate from the average should not be ignored or directed toward sensitivities closer to the average.

A general sensitivity search system is a system in which evaluation criteria for content are modelled for each individual through instructional learning, and each user's evaluation criteria model is used for searching.

The following algorithms have been used in a sensory search; color histogram A method to extract features of images and videos; impression analysis using the SD method A method to quantify the impression received from contents by assigning degrees to impression words, learning correspondence between impression words and contents, extracting correlation coefficients between contents, and Extracting correlation coefficients, the distance between contents Projecting the quantified impression words and features of contents onto the feature space and measuring the distance between them.

2. Tactile Score

Why do we feel "cheer" from the children's "Good morning"? What is the difference between them? The difference is "the way you say it. How we say it can be characterized by the pitch and volume of our voice, but let us look at the volume of our voice.

Yasuhiro Suzuki

When we think back to the children's "Good Morning," their voices gradually became louder and louder. So, let us describe the change in "loudness" by using the music notation. In musical notation, the higher the note is, the lower the note is, and the higher the note is, the lower the note is. The middle line (the third line) is the "normal" volume, and the lower the volume, the higher the volume, and the higher the volume, the lower the volume. The "rhythm of speech" is represented by the notes as they are. We call this notation to describe tactile sense as the Tactile Score [1] (Fig. 1).

Fig. 1 Tactile Score

2.1. *Tactile Quality*

There are several essential factors in the sense of touch, the most crucial being "hardness" and the second most important being "roughness. These properties are called "tactile qualities. Is it possible to describe these tactile qualities in tactile notation?

Think of two kinds of "springs": a stiff spring and a soft spring. The softer spring will expand and contract more when these springs are pushed and pulled with the same force (difference in elastic force). Therefore, we call a tactile spring with a slight change in size "hard" and a tactile spring with a significant change in size "soft". If we consider this as a way of saying "good morning," "hard" is a way of saying "no intonation," and "soft" is a way of saying "full of intonation.

Now we can describe the most critical tactile quality, "hardness," in tactile notation. What remains is roughness. Roughness is described as "smooth" when there is little change in hardness and "rough" when there is a significant change in hardness (Fig. 2). If we are walking on the grass in a park and the hardness of the grass is constant, we will feel "smooth".

3. Extract tactile qualities from time series data
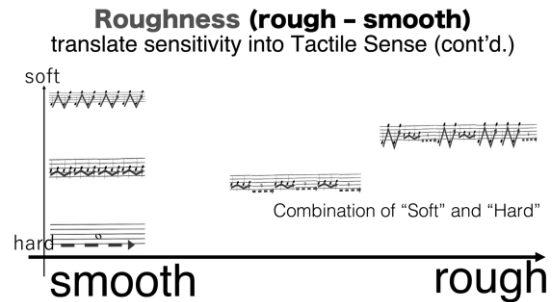
Fig. 2 Roughness, read from Tactile Score

When we press a key, the piano makes a sound as it is. It is impossible to adjust the sound later, as with a violin or trumpet. So, are all piano performances the same? Of course not. Even the same music can sound very different depending on the pianist. One crucial element that makes a difference in piano playing is the "sense of touch. It is often referred to as "the touch of the piano". The word "touch" includes the meaning of "tactile sensation". The sound of the piano, as this expression suggests, is the sound of the sense of touch.

We examine the time series data of piano play and compare Arthur Rubinstein [1887 - 1982] and Stanislav Stanislavovich Bunin [1966-] playing Chopin's Polonaisein A-flat major, Op.53 with WAV format file.

We split the data by 0.1 seconds each and used the mean of the differences in the data as the hardness and the standard deviation of the differences in the data as the roughness.

4. Discussion and Conclusion

Rubinstein is one of the foremost authorities on Chopin. On the other hand, Bunin's style of playing is extremely unique and different from Rubinstein's (Fig. 3).

Fig.3. Tactile qualities of piano play, above) R. Stein, bottom) S. S. Bunin, where the vertical axis illustrates softness (the graph origin shows the hardest and far more upper shows softer and the horizontal axis illustrates roughness, the origin is the smoothest and far from origin means rougher.

Tactile quality analysis of sounds distinguished the differences in play style. In original music analysis, for example, the power spectrum of sounds cannot clearly show the difference (Fig. 4).
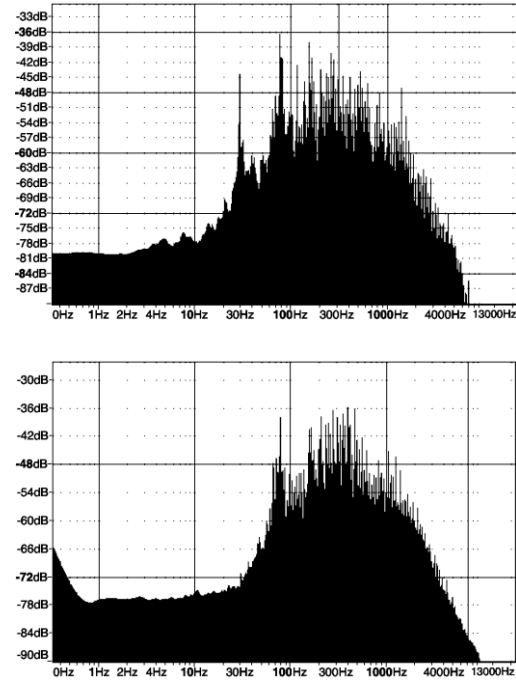
Fig.4. Spectrum distribution of piano play, above) R. Stein, bottom) S. S. Bunin, where the vertical axis illustrates frequency, and the horizontal axis illustrates power.

References

1. Y. Suzuki, R. Suzuki, Tactile score: a knowledge media for tactile sense. Springer Japan, 2014.
2. Kocyan, Wojciech. "The evolution of performance style in the history of the International Fryderyk Chopin Piano Competition in Warsaw." Chopin Review 1 (2018): 20-37.

Authors Introduction

Dr. Yasuhiro Suzuki

He received his Doctor degree in Informatics in 2001 from Graduate School of Informatics, Kyoto University, JAPAN. He is currently an associate professor of Graduate School of Informatics, Nagoya University, JAPAN

Wojciech Kocyan is a Polish pianist. He is most well known for his performances of Chopin, and for his 2001 album, which was named as one of the 50 best classical recordings ever made by Gramophone in 2007. He mentioned about the play style of Bunin.

"there are laureates of the First Prize whose performance style does not fit easily into what would otherwise seem to be a quite clear picture of the stylistic preferences of Chopin Competition juries. They are the Russian pianists Stanislav Bunin (11th Competition, 1985) and Yulianna Avdeeva (16th Competition, 2010)." (from [2] p32).

Healthcare applications of vibrotactile stimulation developed by Tactile Score

Yasuhiro Suzuki

Nagoya University, Furocho, Chikusa-ku, Nagoya, 464-8601, Japan

Abstract

We constructed a system to convert the tactile score into vibratory tactile sensation. We conduct basic biological experiments on electronic tactile stimuli and apply the results to medical and healthcare applications. We use two types of electronic tactile stimuli. One is direct vibrating tactile stimulation, and the other is airborne vibrating tactile stimulation. Verification experiments of this system have confirmed cosmetic effects such as skin collagen aggregation, improvement of hypertension and diabetes, and reduction of pain and nerve paralysis. In addition, a clinical study was conducted in which patients with dementia were exposed to vibrotactile sensation (low-frequency sound) via air vibration and an improvement in cognitive function was confirmed.

Keywords: Healthcare, vibrotactile, Tactile score

1. Introduction

Our hands have great power. There is a word in Japanese called, "Te-arte [1]". This word is the English term for "hand healing," or healing by laying one's hands on someone. In Japan, hand healing has been valued since ancient times.

It seems which is not limited in Japan. Dr. Abraham Varghese, a physician and professor at Stanford University Medical School, gave a TED Talk in 2011 titled "A doctor's touch," in which he emphasized that, the importance of healing through the doctor's touch. He addressed that:

"*the most important innovation in medicine, to come in the next 10 years, that is, the power of the human hand -- to touch, to comfort, to diagnose and to bring about treatment* [2]".

The sense of touch is the earliest of the five senses to develop, which is developmentally and biologically fully formed at birth. A baby's sense of touch at birth, is said to be superior to that of an adult.

However, the science and technology of the sense of touch is, still in its developmental stage compared to the other senses. For example, the senses of audiovisual, smell and taste have their own "products". What about touch? Can you think of some products based solely on the sense of touch? There is a product that people have used since time immemorial, across cultures and customs, that is based solely on the sense of touch. It is the massage. The mural paintings on the tombs of doctors from the Egypt period, depict a massage that is familiar to us today. We may have yet to develop massage techniques for more than 4000 years.

2. Methodology

We have studied massage and proposed a method for describing the sense of touch called the Tactile Score. In our previous research, we found that tactile sensation can only be expressed by the change in vertical force. So, we decided to describe only the change in vertical force over time. It is like describing music only in terms of pitch changes.

2.1. Tactile Score

Why do we feel "cheer" from the children's "Good morning"? What is the difference between them? The difference is "the way you say it. How we say it can be characterized by the pitch and volume of our voice, but let us look at the volume of our voice.

Yasuhiro Suzuki

When we think back to the children's "Good Morning," their voices gradually became louder and louder. So, let us describe the change in "loudness" by using the music notation. In musical notation, the higher the note is, the lower the note is, and the higher the note is, the lower the note is. The middle line (the third line) is the "normal" volume, and the lower the volume, the higher the volume, and the higher the volume, the lower the volume. The "rhythm of speech" is represented by the notes as they are. We call this notation to describe tactile sense as the Tactile Score [3] (Fig. 1)

Fig. 1 Tactile Score

2.2. Transform Tactile Score into Vibrotactile

Tactile score was for mainly hand massages. Hand massage has long history, and its effectiveness has been confirmed. However, its effectiveness would be affected by skill of technician and even if well trained and experienced technician, it is impossible to give massage to more than two persons at the same time. And also, it is also very difficult to confirm that the technician gives exactly same tactile stimulations to patient.
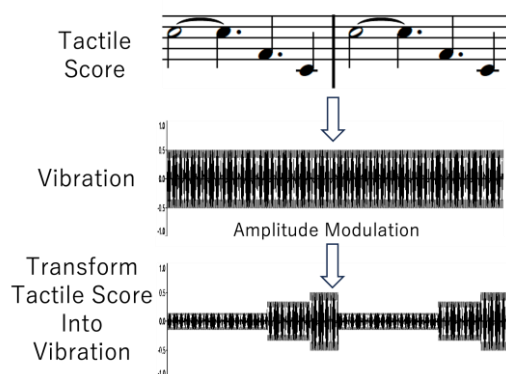
Fig. 2 Transform Tactile Score into vibrotactile

To solve the difficulties, we transform the Tactile score into vibrotactile. Since the Tactile score denotes the strength of force, we transform the force into the amplitude of vibration. Hence, we found that applying amplitude modification with the Tactile score can

transform a Tactile score into a vibrotactile. Since a vibrotactile is a very low-frequency sound, we can record it as a music file in WAV or MP3 format. Hence, we can exchange a vibrotactile translated from a Tactile score via the Internet of electrical files.

By this Tactile score - vibrotactile translation, we can transform tactile stimulation by human hands into electrical data; this technique gives us the way to the digital transformation of tactile sense, tactile DX.

We obtained a Tactile Score, which is converted to vibrations. In order to confirm that this transformation works correctly, the biological responses of the vibration transformed from a Tactile score were compared to a hand massage using the same Tactile Score. Then, we confirmed that the transformed vibration shows almost the same biological response. So, we obtained a method which allows a skilled technician's massage to be converted to a vibratory massage using the Tactile Score (Fig. 2).

We had a method to convert tactile sensations into vibrations but no general-purpose equipment to replay the vibrations. Therefore, together with electrical and audio equipment manufacturers, we developed a general-purpose, face-shaped vibration presentation device (Fig. 3). This device allows users to receive massages by distributing tactile sensations online. We are starting a pilot service of internet delivery of vibration massage around 2019, and more than 1,000 users are already using the service.
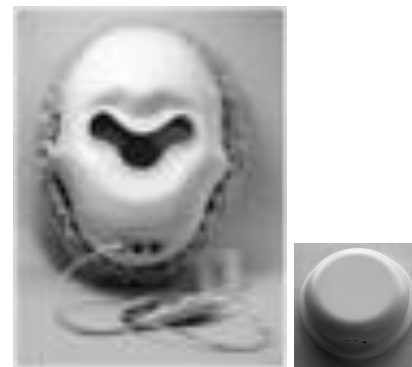
Fig. 3 General-purpose vibrotactile display players; left) for face, right) for body.

With using this equipment, in order to verify the effectiveness of the vibrating massage, we compared the skin condition before and after the treatment. We examined that collagen fibers had agglomerated on the skin surface. It has known that the collagen fibers sink from the surface of the skin due to ageing and other factors, but by applying transformed vibration, they rose to the surface of the skin.

3. Preliminary Results

We are now analyzing test data from users of our Internet massage delivery system who visit hospitals for hypertension and diabetes before and after receiving the vibration massage service. We have just started the analysis and are surprised at the results. One user who was taking medication for hypertension, had stopped taking the user's hypertension medication. Because the user's blood pressure became lower and felt terrible after the user started the vibration massage. Now, the user's blood pressure has stabilized in the normal range without medication. We have confirmed several such cases. Since blood pressure is subject to seasonal fluctuations, we will continue to monitor it over a long period.

In addition, users who had high blood glucose levels and were scheduled to apply insulin, have stabilized their blood glucose levels at normal levels, after starting vibration massage. We have monitored this user's blood glucose levels regularly. We have observed that when the user skips the use of vibratory massage, the blood glucose levels rise again.

4. Conclusion

Our survey of past literature shows that when wounded soldiers were transported in horse-drawn carriages for long periods during the Greek period, their wounds healed faster. Also, in the 18th century, it was confirmed that patients with Parkinson's disease improved their motor functions when transported in a carriage for long periods.

From the survey, our findings in preliminary results would be a rediscovery of the wheel. Vibrotactiles had been utilized in medicine before World War II and were once forgotten. However, in the 21st century, vibrotactile has been rediscovered and applied in medical treatments; at Tronto University, a vibration chair has been used to treat patients with dementia and Parkinson's, and also vibroacoustic therapy has been used in northern Europe or Germany.

The main difference between our method and others is the Tactile score. Other methods have focused on sound or music, while we have been interested in tactile stimulation. Hence, we have focused on how to design the tactile stimulations.

Acknowledgements

References

1. https://tearte.or.jp/archives/1076/ (Japan Te-arte association web page, in Japanese)
2. https://www.ted.com/talks/abraham_verghese_a_doctor_s_touch?language=ja
3. Y. Suzuki and R. Suzuki, Tactile Score, Springer Verlag, 2014.

Authors Introduction

Dr. Yasuhiro Suzuki

He received his Doctor degree in Informatics in 2001 from Graduate School of Informatics, Kyoto University, JAPAN. He is currently an associate professor of Graduate School of Informatics, Nagoya University, JAPAN

Image Collection Experiments of a Handy AUV for Offshore Structure Inspection

Keisuke Watanabe
Tokai University, 3-20-1 Orido, Shimizu-ku, Shizuoka, 424-8610, Japan

Koki Amano
Tokai University, 3-20-1 Orido, Shimizu-ku, Shizuoka, 424-8610, Japan

Shingen Urano
Tokai University, 3-20-1 Orido, Shimizu-ku, Shizuoka, 424-8610, Japan

Yasutaka Taniguchi
Tokai University, 3-20-1 Orido, Shimizu-ku, Shizuoka, 424-8610, Japan

Konosuke Watanabe
Tokai University, 3-20-1 Orido, Shimizu-ku, Shizuoka, 424-8610, Japan
Email: keisuke@tsc.u-tokai.ac.jp, 3ckgm001@mail.u-tokai.ac.jp

Abstract

In recent years, in Japan, with the decline in the working population, there has been a noticeable shortage of labor at construction sites and offshore operations. On the other hand, with the aim of realizing a GX society, it is expected that offshore wind power generation platforms will rapidly increase and be deployed offshore. Since it is not easy to access offshore platforms, it is necessary to automate equipment inspection work, and it is expected that the use of AUVs will be particularly desirable for underwater inspection work. The most basic inspection work is assumed to be to collect images of underwater structures. Therefore, in this study, we created an AUV and conducted an image collection experiment using a pier owned by our university as an example to confirm its functionality.

Keywords: Offshore wind firm inspection, AUV, Underwater image

1. Introduction

In recent years, the worldwide development of wind-power generation has been expanding and the development of offshore wind power is booming in some areas because the construction site of onshore wind power is saturated. For example, in Europe, due to saturation of onshore wind power sites, it is expected that the installation capacity of offshore wind firms will continue to increase, under the condition that the European Union aims at reaching about 100 GW of offshore wind capacity by 2030 [1].

In Japan, the goal of either reducing greenhouse gas emissions to virtually zero or becoming carbon neutral by 2050 was set by the Cabinet recently. In 2020, the government released its "Green Growth Strategy for Carbon Neutrality by 2050", which sets 14 action plants that Japanese industries should start making efforts to realize the net zero society. Under this strategy, wind power is considered very important and set as one of the 14 action plans, with the goal of introducing 10 GW of offshore wind power by 2030 and 30–45 GW by 2040. Since Japan's topography has few open plains, to achieve this ambitious goal, accelerated development of offshore wind firm is required [2] and many offshore platforms will be installed off the coast of Japanese islands.

These offshore wind power platforms will be located far from the shore and accessing many platforms will not be easy. Generally, when maintaining offshore wind power facilities, some specialized crews have to be transferred to the platform using a CTV (Crew Transfer Vessel). If one platform can generate 10MW power, then around 4000 platforms must be installed to achieve the 40GW goal. In that case, it is not hard to imagine that maintenance costs will become huge because an operator must hire CTVs which goes around those platforms. So our concept is such that each platform should be equipped with a unmanned inspection robotic system which performs base inspection menus like taking images from underwater structures to up water windmill structures.

Image collection is one of the basic inspection menus for offshore structures such as harbor structures [3]. Underwater structures are damaged by long-term use and may be corroded by salty sea water. So, inspection process should be performed to check the safety of the structures. The regular safety inspection is very important to prevent collapse accidents. The conventional inspection of harbor structures used to be carried out by professional divers. However, recently, the aging of professional divers and the lack of successors have become a very critical issue in Japan. Also, while port structure is shallower than approximately 20m, structure of an offshore wind platform like a semi-sub column is likely to be deeper than 30m, which makes divers limit their activities in both time and space.

ROVs are widely used for offshore platform inspection [4], where it is assumed that an operator will operate the

Keisuke Watanabe, Koki Amano, Shingen Urano, Yasutaka Taniguchi, Konosuke Watanabe

system by monitoring underwater camera images. In out concept, the inspection procedure must be performed unmanned, so even if the vehicle has a tether cable, the system should be developed as autonomous.

2. AUV System Development

2.1. *System architecture*

An external view of the developed AUV is shown in Fig. 1. It has two pressure-tight vessels, one for the control system and the other for the battery. A total of 6 thrusters are installed, 2 each in the Surge, Sway, and Heave directions. The camera to capture images of the structure is installed inside the acrylic window in front of the upper hull and has a tilt mechanism. It is equipped with sonar to detect obstacles in front and to the sides. A DVL is attached to its bottom to measure altitude above the ocean floor and speed over the ground. The battery is stored in the lower hull, and buoyancy material is placed above to improve self-stability in roll and pitch. This AUV is a prototype for developing various control programs through experiments at sea, and was designed to have an air weight of approximately 120N to make it easier to handle on land. The dimensions of the AUV is shown in Fig. 2.

The architecture of the control system is shown in Fig. 3. This AUV is equipped with two microcomputers. One computer was designed to mainly control input and output of sensors and actuators, and the other computer was designed to control trajectory generation, control mode switching, and image acquisition. It has both wireless LAN and wired LAN, and when a tether cable is attached, it can be used so that it can be connected to a USV or land-based computer and operated in ROV mode. When surfacing and GNSS is available, it navigates using GNSS coordinates, and when it starts diving, it uses dead reckoning from the GNSS coordinates just before diving.

2.2. *Control algorithm*

To implement the control algorithm, we assumed that the motion in the horizontal plane and the vertical motion can be separated because when the velocity is small, the
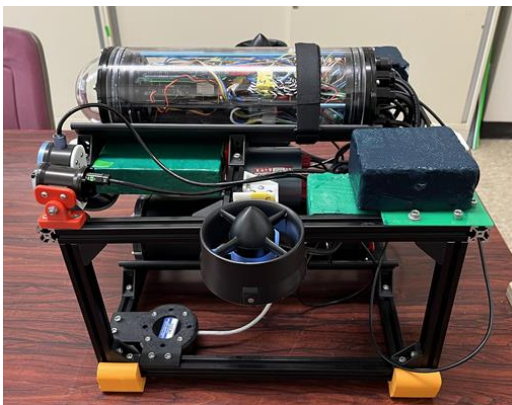
Fig. 1. External view of the AUV
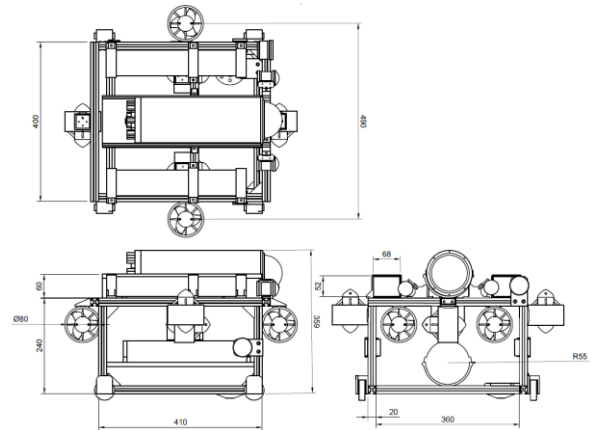
Fig. 2. Dimension of the AUV
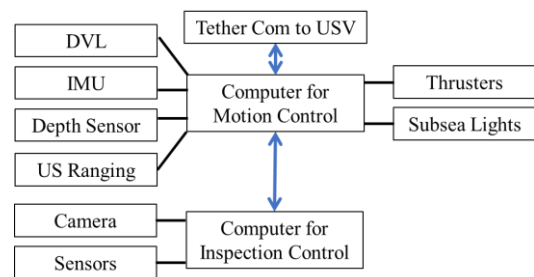
Fig. 3. Control system architecture
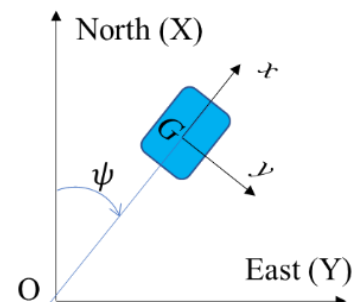
Fig. 4. The coordinate system for horizontal motion

motion in the heave direction has little effect on the motion in the horizontal direction. The coordinate system is set as shown in Fig.4. The coordinate system is a space-fixed coordinate system (O-XY) based on latitude and longitude, and a vehicle-fixed coordinate system (G-xy). G is the vehicle center of gravity. The longitudinal direction of the vehicle is the x-axis and the north is the x-axis. Let the angle formed by the X-axis and the x-axis be the azimuth angle ψ. ψ is positive when clockwise from the north. Space fixed Z axis and vehicle fixed z axis is set as vertically downward to form the right-handed system.

The AUV's equation of motion includes a second-order fluid force term and is nonlinear, but since the speed here is slow, the fluid force is linearized by a representative speed, and the linearly approximated equation of motion is used as the dynamics model. Based on the linearly approximated dynamics model, an optimal control

system of 1-type servo was designed and implemented. Here, the system matrix depends on the coordinate transformation matrix which consists of ψ. Since the coordinate transformation matrix changes according to the motion of the vehicle, it is controlled by solving the LQR problem for each step while updating the system matrix at each control step[5], [6].

2.3. *Sea experiment result*

To verify all functions we designed and implemented, we conducted a sea trial experiment at Tatsugo fishing port in Amami Oshima Island. The experiment site is shown in Fig.5. We set a trajectory along the quay of Tatsugo fishing port as shown in the blue line in Fig.5, and conducted an experiment in which we controlled the depth and azimuth while following this trajectory. The results of this experiment are shown from Figs.6 to Fig.9. The following control modes as shown in Table 1 were set for the trajectory tracking control, and the control mode was changed when the error from the control target value became less than the allowable value as shown in the column of the transition condition.

Table 1. Control modes set for the experiment.

Mode	Target	Transition condition
0	Azimuth ψ= 150 deg Depth = 0.5 m (X,Y)=(7.5cos ψ, 7.5sin ψ)	Error of (X,Y)<=1.0m Depth<=0.1m Then go to mode1
1	Azimuth ψ= 150 deg Depth = 0.5 m (X,Y)=(15cos ψ, 15sin ψ)	Error of (X,Y)<=1.0m Depth<=0.1m Then go to mode2
2	Azimuth ψ= -30 deg Depth = 0.0 m (X,Y)=(15cos ψ, 15sin ψ)	Error of (X,Y)<=1.0m Azimuth<=10 deg Then go to mode3
3	Azimuth ψ= -30 deg Depth = 0.5 m (X,Y)=(7.5cos ψ, 7.5sin ψ)	Error of (X,Y)<=1.0m Depth<=0.1m Then go to mode4
4	Azimuth ψ= -30 deg Depth = 0.5 m (X,Y)=(0,0)	Error of (X,Y)<=1.0m Depth<=0.1m Then go to mode5
5	Termination	Stop all thrusters
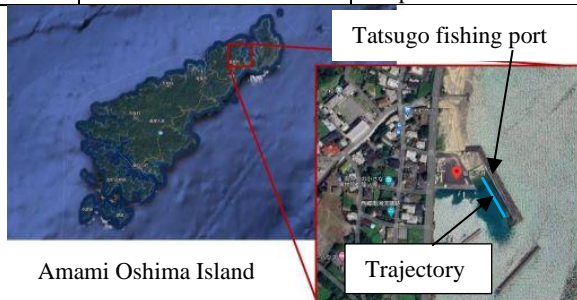
Fig.5. AUV sea experiment site
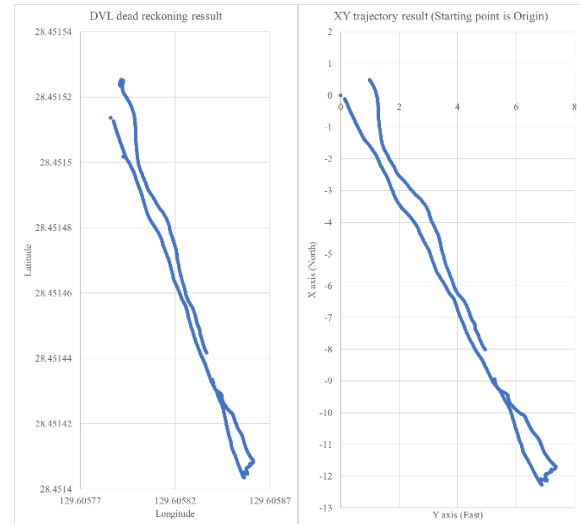
Fig.6. DVL dead reckoning coordinate result
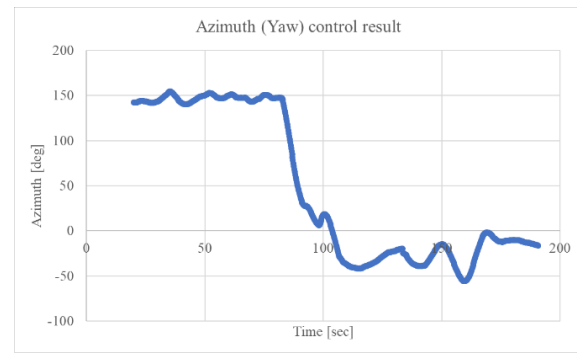
Fig.7. Azimuth angle control result
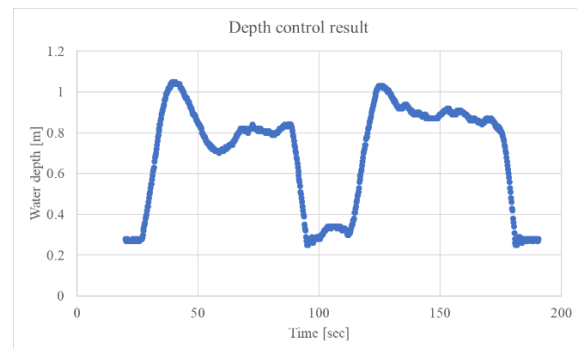
Fig.8. Depth control result (The depth sensor is attached in the middle of the hull cylinder and the initial depth when the AUV is floating is around 0.25m)
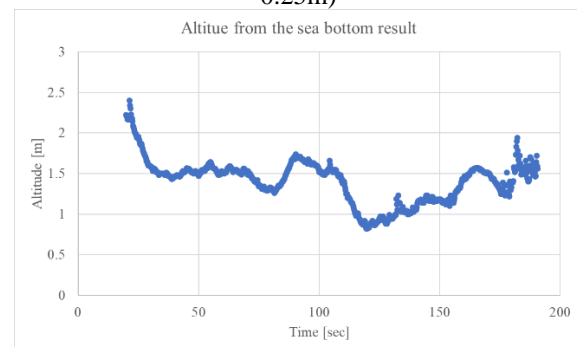
Fig.9. Altitude from the sea bottom

Keisuke Watanabe, Koki Amano, Shingen Urano, Yasutaka Taniguchi, Konosuke Watanabe

Fig.10 Concrete image captured by the AUV

As shown in those Figures above, despite model errors, all functions of the AUV operated smoothly, and the control mission set as shown in Table 1 was successful. Fig.10 shows an example of captured images of the harbor concrete structure. Because the sea was highly transparent, we were able to obtain relatively clear images. We plan to conduct further research and improve AUV functionality.

3. Conclusion

In this research, we developed an AUV platform and conducted trajectory tracking and image acquisition experiments with the goal of unmanned inspection of offshore wind power generation platforms. It was confirmed that trajectory tracking, azimuth, and depth control all functioned smoothly, and image acquisition surveys could be carried out by setting a trajectory.

References

1. E.P. Soares-Ramos, L. de Oliveira-Assis, R. Sarrias-Mena, L.M. Fernández-Ramírez, "Current status and future trends of offshore wind power in Europe", Energy, Vol.202, 2020, 117787.
2. K. Iwata, S. Kyoi, Y. Ushifusa, "Public attitudes of offshore wind energy in Japan: An empirical study using choice experiments", Cleaner Energy Systems, Vol.4, 2023, 100052.
3. J. Choi, Y. Lee, T. Kim, J. Jung and H. -T. Choi, "Development of a ROV for visual inspection of harbor structures," 2017 IEEE Underwater Technology (UT), Busan, Korea (South), 2017, pp. 1-4
4. Capocci R, Dooly G, Omerdić E, Coleman J, Newe T, Toal D. "Inspection-Class Remotely Operated Vehicles—A Review", Journal of Marine Science and Engineering. 2017; 5(1):13.
5. K. Watanabe, K. Amano, "Conceptual Design of USV/UUV Combined Autonomous Platform for Offshore Structure Inspection," 2023 IEEE International Conference on Marine Artificial Intelligence and Law (ICMAIL), Taipei, Taiwan, 2023, pp. 1-6.
6. K. Watanabe, M. Shimpo, "Examination of Sea Experimental System for Catamaran Type USV Development", The Journal of Japan Institute of Navigation, 2022, Vol.146, pp.30-29(in Japanese)

Authors Introduction

Dr. Keisuke Watanabe

He received his B.Engineering, M.E., ph.D in 1991,1993,1996,respectively from the Department of Naval Architecture and Ocean Engineering, the University of Tokyo. He is a Professor of the School of Marine Science and Technology, Tokai University in Japan. He is a member of JASNAOE, MTS, IEEE etc..

Mr. Koki Amano

He received his B.Oceanography in 2023 from the School of Marine Science and Technology, Tokai University in Japan. He is currently a master student in the Graduate School of Marine Science and Technology, Tokai University.

Mr. Shingen Urano
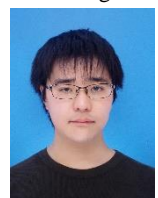
He received his B.Oceanography in 2023 from the School of Marine Science and Technology, Tokai University in Japan. He is currently a master student in the Graduate School of Marine Science and Technology, Tokai University.

Mr. Yasutaka Taniguchi

He is currently an undergraduate student in the School of Marine Science and Technology, Tokai University.

Mr. Konosuke Watanabe

He is currently an undergraduate student in the School of Marine Science and Technology, Tokai University.

Design of Disassembly-reassembly Type USV for Coral Reef Research

Keisuke Watanabe
Tokai University, 3-20-1 Orido, Shimizu-ku, Shizuoka, 424-8610, Japan

Koki Amano
Tokai University, 3-20-1 Orido, Shimizu-ku, Shizuoka, 424-8610, Japan

Gaku Minato
Tokai University, 3-20-1 Orido, Shimizu-ku, Shizuoka, 424-8610, Japan

Yasutaka Taniguchi
Tokai University, 3-20-1 Orido, Shimizu-ku, Shizuoka, 424-8610, Japan

Konosuke Watanabe
Tokai University, 3-20-1 Orido, Shimizu-ku, Shizuoka, 424-8610, Japan
Email: keisuke@tsc.u-tokai.ac.jp, 3ckgm001@mail.u-tokai.ac.jp

Abstract

Researchers are conducting physical surveys by diving and swimming to study the effects of microplastics and global warming on coral reefs. The area, time, and water depth that can be investigated by diving are extremely limited. In addition, to collect microplastics, it is necessary to tow the nets, but chartering a ship is expensive. Therefore, the authors are developing a system that simultaneously operates a USV and a UUV to simultaneously observe the sea surface and underwater. In this paper, we conducted a conceptual design, carried out fluid force measurements, trajectory tracking experiments, and image recognition using AI, with the aim of realizing a lightweight USV that can be divided during transportation to coral reef areas and reassembled on site for operation.

Keywords: USV/UUV combined, Disassembly-reassembly type USV, Coral reef monitoring, AI image processing

1. Introduction

In recent years, marine plastic waste has been widely recognized as a global environmental problem. Because most plastics undergo very slow chemical or biological degradation in the environment, debris can remain in the ocean for years, decades, or even longer [1]. The term microplastic refers to very small and ubiquitous plastic particles less than 5 mm in diameter. They have been separated into different fractions: large (1–5 mm) or small (1 μm–1 mm) microplastics and the fraction below 20 μm (20 μm–1 μm) [2], [3]. Microplastics can adsorb and carry hydrophobic chemicals that can have biological and toxicological effects on the environment [4]. Therefore, a clear understanding of the interactions of small microplastics with the environment, especially living organisms, is essential to assess potential health hazards. To this end, it is important to have accurate methods to quantify the mount of such particles in natural environments. By monitoring the amount and characteristics of particles over time, along with the coordinates from which they were collected, it will be possible to define the extent of the effects they cause.

Additionally, by collecting small microplastics and identifying them at the single particle level, it will be possible to create some indicators regarding the potential impact of microplastics on marine biodiversity and their distribution in the ocean. Currently, there are two methods for collecting microplastics in coral reef waters: chartering a ship and towing the nets, or having people swim and tow the nets. The former requires the ship to be routed from the port to the collection site, and costs such as chartering fees are incurred. While the latter can be accessed directly from the beach, the area in which it can be retrieved is very limited.

With the goal of improving these shortcomings and efficiently collecting more microplastics in coral reef waters, we are developing a disassembly-reassembly type USV, which we will introduce in this paper. The USV under development will tow the net while automatically following a set trajectory by a biologist to collect microplastics. The time and trajectory of microplastic collection will be recorded by GNSS data. Since the work is intended to be carried out even on remote islands, we designed a structure that could be disassembled, transported by aircraft, etc., and reassembled on-site.

Keisuke Watanabe, Koki Amano, Gaku Minato, Yasutaka Taniguchi, Konosuke Watanabe

2. USV System Development

2.1. *System architecture*

An external view of the developing USV is shown in Fig.1. The dimensions of the USV is shown in Fig.2. As shown in Fig.1, this USV consists of four hulls. The size of the hull was determined to fit within the typical packaging size used in typical Japanese delivery services. As shown in Fig.2, the thruster, computer, and GNSS antenna are attached to the aluminum frame. The center of the hull has a large opening so that the USV can collect microplastics while navigating, and a net for collecting microplastics is attached to the rear of the USV. The hardware components of control system is shown in Fig.3. This USV is equipped with two microcomputers. One computer is for motion control, and the other is for image acquisition and processing. Image processing is supposed to be used for collision avoidance when the tide is low and the reef is exposed above the sea surface. As for the algorithm of trajectory tracking, we hire optimal control based 1-type servo [5].

2.2. *Drag force measurement of a hull*

In order to estimate the maximum speed of the USV and construct a nominal model for designing the optimal control system, it is necessary to estimate the drag force coefficient. Therefore, as shown in Fig. 4, the hull was set in a circulating water tank and the drag force was measured by varying the flow velocity and angle of attack. The flow velocity was increased every 0.2 m/s and was measured from 0.4m/s up to 1.0 m/s. The measurement results are shown in Fig.5. As the graph shows, it is almost parabolic, confirming that the drag force is proportional to the square of the flow velocity. The drag force coefficient when moving forward is estimated using the equation, where D is drag force, ρ is water density, A is representative area of the hull, u is velocity.

$$C_d = \frac{D}{\frac{1}{2}\rho A u^2} \qquad (1)$$
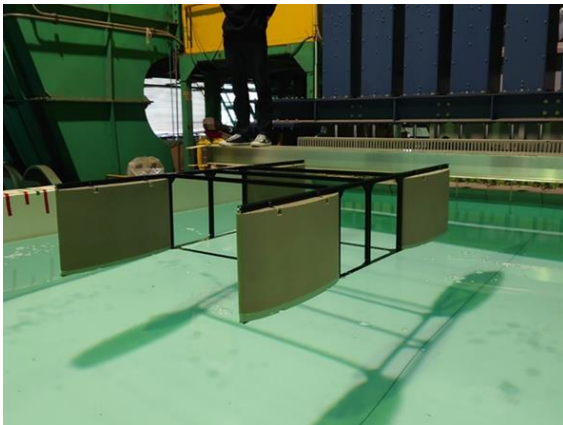
Fig. 1. External view of the USV
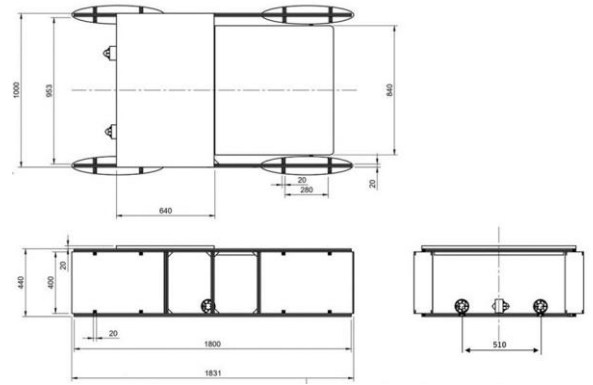
Fig. 2. Dimensions of the USV
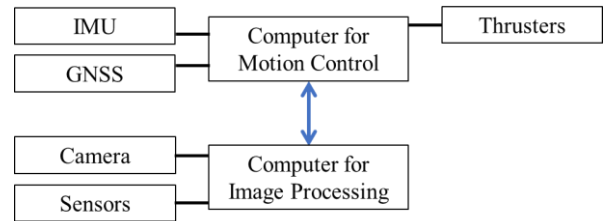
Fig. 3. Control system architecture
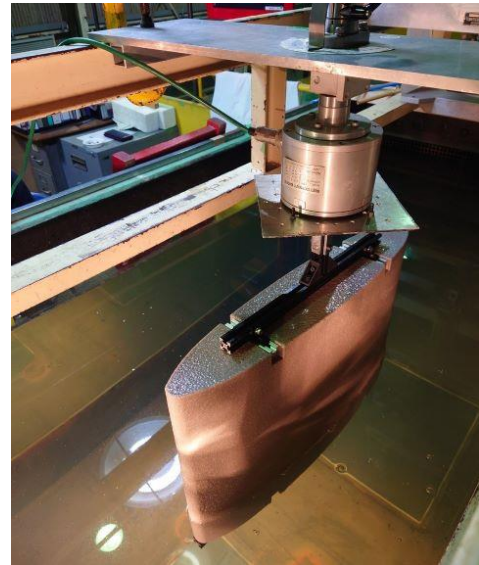
Fig. 4. Hull drag force measurement.
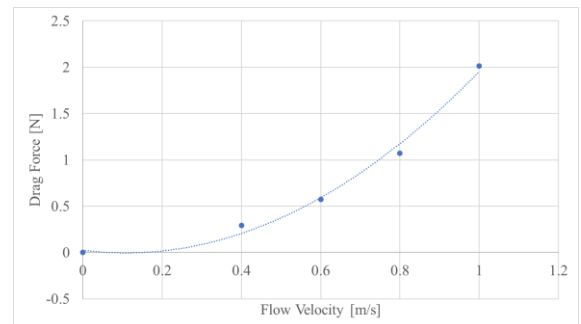
Fig.5 Drag force result

The drag force coefficient became 0.18 in average.

3. Image Recognition Experiment Using AI

Since it is necessary to avoid obstacles during automatic navigation, we are considering an image object recognition function. We conducted an experiment using YOLOv5 using our university's pier column as an example. Fig.6 is an example of an image of the "pillar" labeling that YOLO is trained to learn. About 100 photos were taken to show the different shapes of the pillars on the quay and the changes in the exposed parts of the pillars due to the ebb and flow of the tide, and were used as training data. Fig.7 shows the results of image recognition based on the above learning. It can be seen that the pillars at the bottom of the quay could be recognized, but the probability of recognition was only about 60%, and when sailing from an angle, it was often not possible to recognize them. Fig.8 shows the results of learning by adding additional training data, including images from diagonal directions. The recognition rate was over 80%, and the pillars could now be recognized even in images viewed from an angle. Image recognition using AI is extremely important as a function of USV, and in addition to detecting obstacles, it may also be applied to recognizing coral reefs using underwater cameras in coral reef surveys. YOLO is a very powerful tool, but it is necessary to prepare various images as training data, and how to obtain images for learning at the first site is a future issue. One possible solution would be to fly a UAV (Unmanned Aerial Vehicle) in the research area and use the images captured by the UAV as learning data. Floating debris such as large pieces of wood or plastic bottles can break the microplastic mesh if the USV swims over them, so it must have a function to find them ahead and avoid them.

Fig.6 An image example for pillar labeling

Fig.7 An example of image recognition results

Fig.8 Improved recognition result after more learning

4. Trajectory Tracking Experiment

In order to confirm the resolution of GNSS and confirm the trajectory tracking performance, we conducted a trajectory tracking experiment using a catamaran-type USV that has been developed in our laboratory, rather than the USV currently under development. The results are shown from Fig. 9 to Fig.11. Fig.9 is a map of the trajectory in the actual ocean area. Fig.10 shows the results of thrust forces. Fig.11 shows converting the latitude and longitude to the XY coordinate system with the starting point as the origin. It was set to sail to the (-10,10) point and then return to the origin. On the day of the experiment, the wind was strong and there was some deviation from the straight trajectory, but we were able to confirm that it had sufficient performance to collect microplastics.
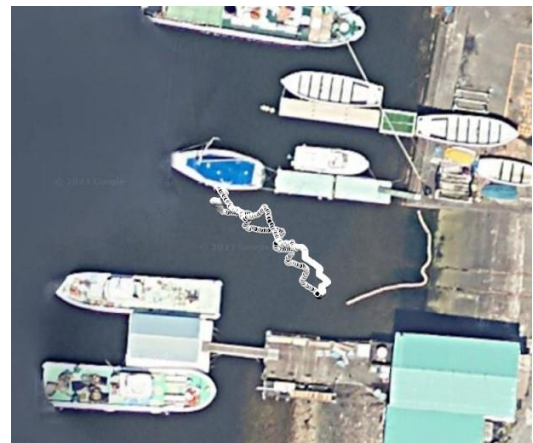
Fig.9 Trajectory result

Fig.10 Thruster forces

Keisuke Watanabe, Koki Amano, Gaku Minato, Yasutaka Taniguchi, Konosuke Watanabe
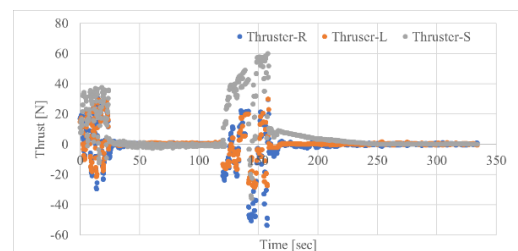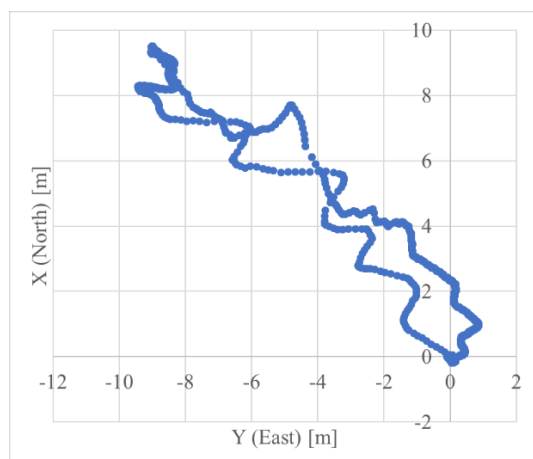
Fig.11 XY coordinate of the trajectory
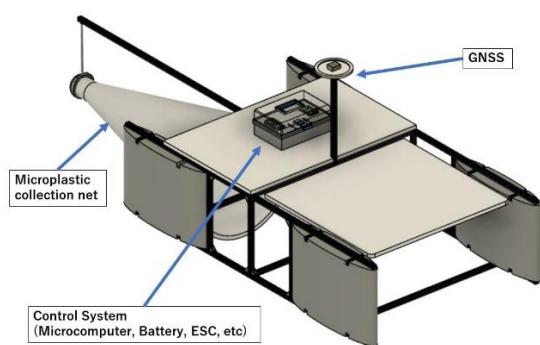
Fig.12 The total image of the USV

5. Summary

In this study, we designed a USV with the goal of automating and reducing the cost of microplastic collection in coral reef waters, which is currently done manually. A split-assembly type USV was designed to facilitate transportation to remote islands. The conceptual design drawing is shown in Fig.12. In order to realize the obstacle avoidance function through image processing, we implemented an object detection system using AI. In order to construct a nominal model for control system design, the hull's fluid force coefficient was identified through experiments using a circulating water tank. Through trajectory tracking experiments using existing USV, we have confirmed that trajectory tracking is possible within an error range of approximately 1m. We plan to conduct microplastic collection experiments.

References

1. Anthony L. Andrady, "Microplastics in the marine environment", Mar. Pollut. Bull., 62 (8) (2011), pp. 1596-1605
2. Julien Gigault, Alexandra ter Halle, Magalie Baudrimont, et.al., "Current opinion: what is a nanoplastic?", Environ. Pollut., 235 (2018), pp. 1030-1034
3. Christina Ripken, Domna G. Kotsifaki, Síle Nic Chormaic, "Analysis of small microplastics in coastal surface water samples of the subtropical island of Okinawa, Japan", Science of The Total Environment, 2021, Vol. 760, 143927
4. Richard E. Engler, "The complex interaction between marine debris and toxic chemicals in the ocean", Environ. Sci. Technol., 46 (22) (2012), pp. 12302-12315
5. K. Watanabe, K. Amano, "Conceptual Design of USV/UUV Combined Autonomous Platform for Offshore Structure Inspection," 2023 IEEE International Conference on Marine Artificial Intelligence and Law (ICMAIL), Taipei, Taiwan, 2023, pp. 1-6.

Authors Introduction

Dr. Keisuke Watanabe

He received his B.Engineering, M.E., ph.D in 1991,1993,1996,respectively from the Department of Naval Architecture and Ocean Engineering, the University of Tokyo. He is a Professor of the School of Marine Science and Technology, Tokai University in Japan. He is a member of JASNAOE, MTS, IEEE etc..

Mr. Koki Amano

He received his B.Oceanography in 2023 from the School of Marine Science and Technology, Tokai University in Japan. He is currently a master student in the Graduate School of Marine Science and Technology, Tokai University.

Mr. Gaku Minato
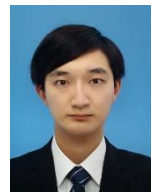
He is currently an undergraduate student in the School of Marine Science and Technology, Tokai University.

Mr. Yasutaka Taniguchi

He is currently an undergraduate student in the School of Marine Science and Technology, Tokai University.

Mr. Konosuke Watanabe

He is currently an undergraduate student in the School of Marine Science and Technology, Tokai University.

Optimization method to improve visual SLAM in dynamic environment

Yufei Liu

Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku, Kitakyushu, 808-0196, Japan

Kazuo Ishii

Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku, Kitakyushu, 808-0196, Japan
Email: liu.yufei124@mail.kyutech.jp, ishii@brain.kyutech.ac.jp

Abstract

Aiming at the problem that the robustness of the classic visual SLAM system is greatly affected by dynamic target feature points in the environment, a method is proposed to use a target detection algorithm to identify and eliminate dynamic target feature points. First, use the target detection algorithm YOLOv5 to identify the collected environmental images, and select the surrounding environment. Objects identified as dynamic targets in the environment, and then the target detection results are integrated into the feature extraction of the visual SLAM front-end, the feature points belonging to the dynamic target part of the extracted image feature points are removed, and the remaining static feature points are used to map Construction and positioning, and finally testing on the TUM data set. The results show that after using the target detection algorithm to eliminate dynamic feature points, the root mean square error of the absolute trajectory error of the visual SLAM system in highly dynamic scenes is reduced by 97.89%, effectively improving the positioning accuracy and robustness of the system.

Keywords: Visual SLAM, Deep learning, Feature detection, Position estimation, YOLOv5

1. Introduction

With the development of autonomous mobile robot platforms, robots are widely used in areas such as search and rescue operations or delivery services required at industrial sites and hotels. In these scenarios, the robot needs to understand the entire area and the precise location of the target object in the map based on the map to complete autonomous navigation. In order to achieve autonomous navigation, mobile robots need to complete two tasks: attitude estimation and map construction. Therefore, Simultaneous Localization and Mapping (SLAM) is one of the research hotspots in the field of mobile robots. It is a robot that only relies on its own sensors to obtain external information of the unknown environment to complete pose estimation and build an environment model. Technology. This technology is currently widely used in agriculture, disaster relief, auxiliary medical and other robotic fields. According to the different sensors used, SLAM is divided into visual SLAM and laser SLAM. Visual SLAM uses cameras mounted on the robot as sensors, including monocular, binocular and depth cameras. Compared with lidar, cameras are cheaper and more convenient to install and debug. Moreover, cameras can collect more semantic information in the environment. Therefore, visual SLAM has gradually become a major research hotspot in the field of autonomous navigation of mobile robots.

2. Problem Description

2.1. *Overview of visual SLAM technology*

In recent years, various visual SLAM (Simultaneous Localization and Mapping) techniques have emerged, such as ORB-SLAM, ORB-SLAM2, VINS Mono, RTABMap, PTAM, LSD-SLAM, DSO, among others, playing a crucial role in the field of autonomous robot navigation [1]. The general algorithmic process of visual SLAM begins with the acquisition of images using a camera to generate a sequence of images. Subsequently, features are extracted from adjacent frames, and the motion of the camera is estimated by minimizing pixel intensity values through feature point matching. Pixel matching methods between images include optical flow and direct methods. Optical flow involves extracting image feature points and estimating camera motion using triangulation or Epipolar Geometry. Direct methods, on the other hand, directly use pixel blocks or extract image corner points to calculate motion estimation based on grayscale values.

Following the motion estimation, noise is filtered from the data to obtain the optimal pose estimation. Maximum a posteriori probability is then used to estimate the global map. Campos et al. [2], building on the research by Artal et al., released ORB-SLAM3 in 2020, which is a feature-based SLAM system and currently represents a notable approach in the field.

2.2. *Impact of Dynamic Points on Visual SLAM*

ORB-SLAM3, as one of the classic algorithms in visual SLAM, consists of two main components in its algorithmic pipeline: the front end and the back end. The front end, also known as visual odometry, processes captured images through feature extraction and matching. It solves the Epipolar Geometry relationship between corresponding feature pixels, thereby estimating the pose parameters of camera translation and rotation. The back end includes nonlinear optimization, map estimation, and loop closure detection. Nonlinear optimization, differing from the previous use of Kalman filtering, employs Bundle Adjustment (BA) to simultaneously optimize the six degrees of freedom of camera pose parameters and the poses of landmarks in space.

The effectiveness of the Bundle Adjustment (BA) method in optimizing camera pose results depends on whether the extracted matching feature points from the captured images exclusively static object features are. However, real-world robot operating environments often include dynamic objects. In scenarios with dynamic targets, both feature-based SLAM algorithms and direct methods SLAM algorithms struggle to differentiate between feature types in regions with moving objects. The matching points of dynamic target feature points can lead to misalignments, resulting in reduced accuracy in the front-end visual odometry's estimation of camera pose. This, in turn, causes a loss of camera pose tracking, leading to significant deviations between the robot's motion trajectory and the constructed environmental map.

3. Elimination of the dynamic points

3.1. *Dynamic SLAM based on geometric methods*

Currently, there are two main methods for mitigating the impact of dynamic target feature points in detection scenes. One approach is the traditional geometric-based method. Sun et al. [3] detect moving objects by comparing differences between adjacent frames, but this method suffers from poor real-time performance. Wang et al. [4] filter matching feature points in adjacent frames using Epipolar Geometry. They also perform clustering on depth images collected by RGB-D cameras to identify independent dynamic objects in the scene. The accuracy of this algorithm is significantly influenced by the pose transformation matrix between adjacent frames. If there are many highly dynamic objects in the scene, it can lead to substantial algorithmic errors. Lin et al. [5] use a fusion of image depth information and visual ranging to detect moving objects. While this approach can determine the position of moving targets in the scene, the algorithm's accuracy is compromised due to the uncertainty of depth information and cumulative errors in calculating the transformation matrix between adjacent frames.

The underlying principle of these methods assumes that the feature points of dynamic objects deviate from the standard constraints obtained in a static scene through processes such as triangulation, fundamental matrix estimation, epipolar lines, and reprojection error. During the pose estimation process, these dynamic feature points are treated as outliers. The correctness of feature matching can be determined by examining whether the extracted feature points violate these constraints, and dynamic points are subsequently excluded. However, the accuracy of this method depends on the proportion of static feature points in the scene. In scenarios with a high density of dynamic objects, it can significantly impact the reliability of pose estimation and the accuracy of map construction.

3.2. *Dynamic SLAM based on deep learning*

With the development of deep learning, many researchers have introduced deep learning algorithms into visual SLAM to mitigate the impact of dynamic targets and improve its accuracy. The DynaSLAM (Dynamic SLAM) algorithm proposed by Berta Bescos et al. [6] optimizes ORB-SLAM2 using Mask R-CNN (Region-based Convolutional Neural Network). By combining geometry with deep learning, DynaSLAM filters out dynamic feature points in the scene. The algorithm has shown excellent results on the TUM dataset, but its use of Mask R-CNN for pixel segmentation leads to low real-time detection efficiency, limiting its practical application in real-world environments. DDL-SLAM (Dynamic Deep Learning SLAM) [7] uses DUNet (Deformable Unity Networking) and semantic masks obtained through multi-view geometry to detect dynamic objects. It then employs an image restoration strategy to reconstruct the background obscured by dynamic objects. Since the calculation of dynamic object masks occurs at the pixel level, this method also falls short of achieving real-time performance. In contrast to Mask R-CNN, another more efficient object detection model is YOLOv5 (You Only Look Once Version 5) [8]. YOLOv5 achieves detection speeds of up to 45-155 frames per second (fps), which is 9-30 times faster than the maximum 5 fps achieved by Mask R-CNN. If the results of YOLOv5's object detection can be integrated into dynamic visual SLAM algorithms, it can partially compensate for the low efficiency of Mask R-CNN.

To address the issue of reduced localization accuracy and robustness of visual SLAM in dynamic environments due to the influence of dynamic targets, this paper introduces a target detection thread based on the ORB-SLAM3 algorithm to detect dynamic targets. Using the YOLOv5 target detection network to identify objects in input images, the semantic information of dynamic target objects in the images is determined. Additionally, while visual SLAM extracts feature points in the front end, a module is added to the tracking thread of visual SLAM to eliminate dynamic feature points based on the results of

target detection. Finally, only the remaining static feature points are used to estimate the pose change matrix between adjacent frames, reducing the impact of dynamic targets. This ultimately achieves high-precision estimation and localization of the environmental trajectory. The improved algorithmic flowchart is shown in Fig.1.

4. Experimental verification and result analysis

4.1. *Experimental environment construction and data set selection*

his experiment involved testing the optimized visual SLAM system proposed in this paper using video sequences from the TUM dataset. The experimental results were analyzed to evaluate the localization accuracy of the SLAM system. The experiments were conducted on the Ubuntu 20.04 operating system, with a 12th Gen Intel(R) Core(TM) i9-12900H 2.50GHz CPU, an NVIDIA GeForce RTX 3060 GPU with 12GB of VRAM, and the PyTorch deep learning framework. The algorithm's performance was tested on the fr3_walking_xyz, fr3_walking_half, and fr3_walking_static dataset sequences. The fr3_walking_xyz dataset depicts two individuals walking and conversing in a fixed scene, with both the camera and people in motion, representing a high-dynamic scene. The fr3_walking_half dataset builds upon this by having the camera move along a semi-circular trajectory in the air. The fr3_walking_static dataset, on the other hand, features relatively stationary objects, representing a low-dynamic scene.

4.2. *Visual SLAM front-end feature extraction effect after integrating YOLOv5*

Capturing a frame for comparison with the original ORB-SLAM3 algorithm's front-end feature extraction, the results are illustrated in Fig. 2. The lower image shows the feature extraction before integrating YOLOv5, while the upper image demonstrates the effect after fusion. It is evident that the visual SLAM front end, after integrating YOLOv5, accurately identifies objects such as people and computers. It successfully removes feature points on the dynamic target "person" while retaining feature points on the static object "computer" within the detected bounding box of the dynamic target. This refinement enhances the precision of the results.
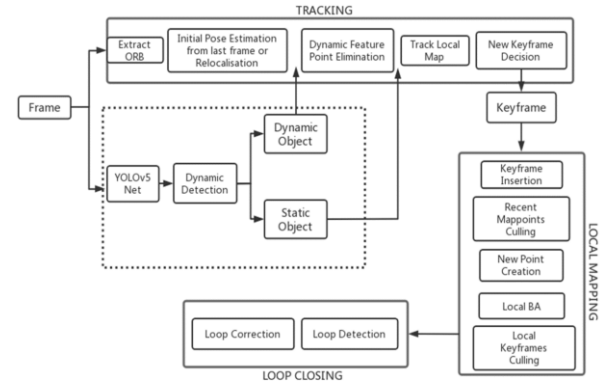
Fig.1 Algorithm framework

4.3. *Experimental data indicator analysis*
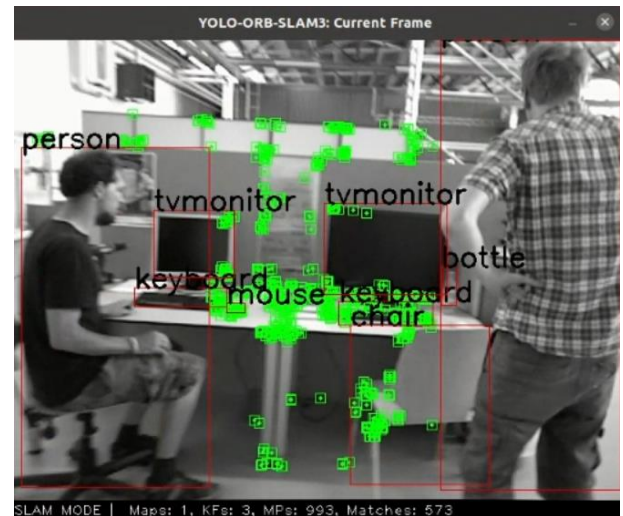
a.   Fusion algorithm

b.   Original algorithm

Fig.2. Comparison of front-end feature extraction results before and after integrating YOLOv5

Absolute Trajectory Error (ATE) describes the difference between estimated poses and ground truth poses, providing an intuitive expression of the algorithm's

Yufei Liu, Kazuo Ishii

global performance and accuracy. Relative Pose Error (RPE) calculates the differences in pose at the same timestamps, typically used for estimating odometry errors. The Root Mean Square Error (RMSE) is then employed to compute the overall value of this error. ATE and RPE are two parameters that reflect the robustness and stability of a visual SLAM system [9]. A lower RMSE value calculated from ATE and RPE indicates a better fitting performance.

Capturing a frame for comparison with the original ORB-SLAM3 algorithm's front-end feature extraction, the results are illustrated in Fig. 2. The lower image shows the feature extraction before integrating YOLOv5, while the upper image demonstrates the effect after fusion. It is evident that the visual SLAM front end, after integrating YOLOv5, accurately identifies objects such as people and computers. It successfully removes feature points on the dynamic target "person" while retaining feature points on the static object "computer" within the detected bounding box of the dynamic target. This refinement enhances the precision of the results. The improvement of the ATE and RPE performance of this algorithm compared with the original ORB-SLAM3 algorithm is shown in Table.1 and Table.2.
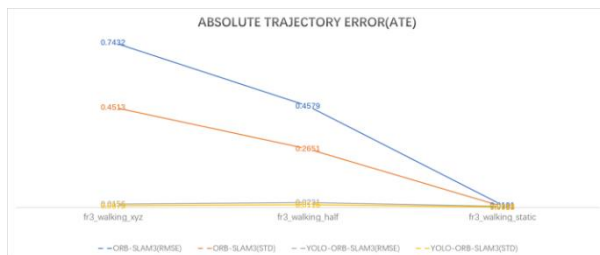
Table.1 ATE performance comparison
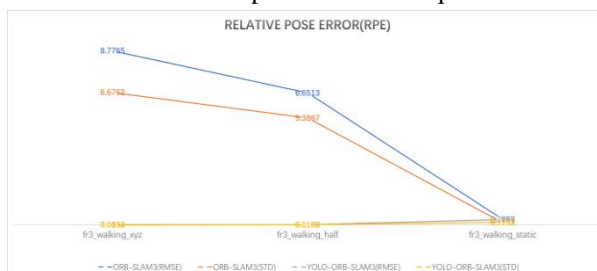
Table.2 RPE performance comparison

This indicates that the proposed algorithm not only enhances the detection capability in dynamic environments but also preserves the accuracy of map construction and localization in static scenes, as achieved by the original ORB-SLAM3 algorithm.

References

1. Zhaopeng G, Liu H, University P, et al. A survey of monocular simultaneous localization and mapping[J]. CAAI Transactions on Intelligent Systems, 2015.

2. Campos C, Elvira R, Rodríguez J J G, et al. Orb-slam3: An accurate open-source library for visual, visual‑inertial, and multimap slam [J]. IEEE Transactions on Robotics, 2021, 37(6): 1874-1890.

3. Sun Y, Liu M, Meng M Q H. Improving RGB-D SLAM in dynamic environments: A motion removal approach[J]. Robotics and Autonomous Systems, 2017, 89: 110-122.

4. Wang R, Wan W, Wang Y, et al. A new RGB-D SLAM method with moving object detection for dynamic indoor scenes[J]. Remote Sensing, 2019, 11(10):1143.

5. Lin S F, Huang S H. Moving object detection from a moving stereo camera via depth information and visual odometry[C]//2018 IEEE International Conference on Applied System Invention (ICASI). IEEE, 2018:437-440.

6. Bescos B, Fácil J M, Civera J, et al. DynaSLAM: Tracking, mapping, and inpainting in dynamic scenes[J]. IEEE Robotics and Automation Letters, 2018, 3(4) : 4076-4083.

7. Ai Y, Rui T, Lu M, et al. DDL-SLAM: A robust RGB-D SLAM in dynamic environments combined with deep learning[J]. IEEE Access, 2020, 8: 162335-162342.

8. Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016:779-788.

9. Chang Z, Wu H, Sun Y, et al. RGB-D Visual SLAM Based on Yolov4-Tiny in Indoor Dynamic Environment[J]. Micromachines, 2022, 13(2):230.

Authors Introduction

Ms. Yufei Liu

She graduated from Zhejiang Normal University in China with a bachelor's degree in 2017. Now she is studying as a master's student at Kyushu Institute of Technology in Japan.

Dr. Kazuo Ishii

He received his PhD from the University of Tokyo, Japan, in 1996. In 2011, he joined Kyushu Institute of Technology and is currently a professor in the Department of Human Intelligent Systems. His research interests include information communications and marine robotics. He is a member of IEEE.

Design of flexible mechanism for flexible manipulator

Huang Jiawei

Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku, Kitakyushu, 808-0196, Japan

Kazuo Ishii

Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku, Kitakyushu, 808-0196, Japan
Email: huang.jiawei311@mail.kyutech.jp, ishiii@brain.kyutech.ac.jp

Abstract

Based on the study of the kinematic limitations of the rigid manipulator structure and the characteristics of the existing flexible manipulator in different categories, this paper proposes a structure assumption of flexible manipulator based on the advantages of high load of rigid joint and high flexibility of flexible manipulator. This paper designed a flexible manipulator structure driven by wires. Using the forward kinematics analysis, the sport model of the structure. The motion range of the end of the manipulator arm and the length changes of the wires were simulated with the movement of the model.

Keywords: Flexible manipulator, Wires driven, Kinematic model

1. Introduction

With the development of science and technology, rigid robot devices have been widely used in social production activities. It is well known that the industrial robot is applied in the manufacturing industry. The rigid manipulator consists of a rigid connecting rod connected by a rigid motion pair, which can quickly and accurately complete the corresponding control tasks [1]. However, due to the rigid structure of the manipulator, its flexibility is poor, and it cannot work properly in the complex and changeable environment. Its structural characteristics limit its application in dangerous and complex environments, such as exploration and disaster relief, nuclear power hazard management, space missions and other environments. If the manipulator is to be applied to the above engineering fields, it needs to meet the requirements of safe interaction between the machine and the environment, safe human-computer interaction, high flexibility and intelligence. The flexible manipulator is made of soft materials. Compared with the rigid robot, the manipulator has the characteristics of multiple redundant degrees of freedom, flexibility and safe human-computer interaction. Its characteristics greatly make up for the shortcomings of rigid robots, so it has been widely concerned and studied by scholars and institutions [2].

2. Classification of prior research

The development status of flexible manipulator is introduced according to the driving mode classification. The flexible manipulator is divided into the following driving modes: SMA driven, pneumatic - hydraulic driven, and wire driven, etc.

2.1. SMA driven

SMA, full of Shape Memory Alloys, can be heated to eliminate the deformation of alloys at a lower temperature, making it revert to the initial state and achieving the effect of "memory". It can be driven by direct heating of current, and its deformation can be controlled by heating and cooling. It has been widely used in aerospace and industrial intelligent manufacturing. As shown in the Fig.1, the manipulator can realize large spatial bending and complete the grasping action of different positions.

Fig.1 SMA flexible mechanical arm [3]

2.2. Pneumatic-hydraulic driven

Pneumatic or hydraulic driven respectively use compressed air and hydraulic oil as the power source of the mechanical arm bending. Pneumatic - hydraulic driven manipulator is usually made of elastic material as the main body, surrounded by a driving cavity. The bending of the manipulator can be realized by changing the volume deformation of the driving cavity by changing the input of the power source as showed in Fig.2. When the inner rubber air bag is filled with compressed air, the internal pressure rises, the inner rubber air bag expands along the radial direction, and then the force transmission

effect of the outer fiber is transformed into the axial contraction force to achieve the driving effect.

Fig.2 Pneumatic driven flexible manipulator [4]

2.3. *Wire driven*

Wire driven is a popular driving method for flexible manipulator. The bending motion of the flexible arm is generally realized by the stretching of the wire driven by the motor. The wire driven flexible manipulator is driven by the wires embedded in the interior of the flexible manipulator to realize the deformation movement of the flexible manipulator as showed in Fig.3. Usually, mechanical equipment such as motor is used to generate traction on the wire. The wire embedded in the flexible manipulator with eccentric eccentricity generates axial force on the manipulator while generating bending moment on the neutral surface, making the flexible bending motion of the manipulator.
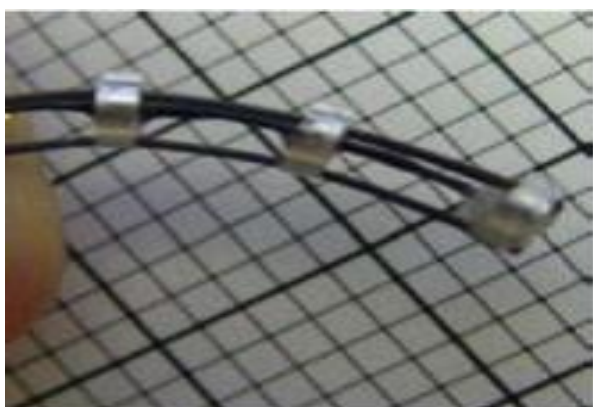
Fig.3 Wire driven flexible manipulator [5]

3. Model scheme

The flexible robot arm is composed of three layers of single-section robot arm, the height of the single-section robot arm is 200mm, and the height of the entire robot arm is 600mm. The central flexible column is a soft plastic material with good flexibility and bearing capacity. The hardness and elastic modulus of soft plastics can be controlled by additives to adapt to different application scenarios. The flexible column is used as the main body

of the flexible manipulator. The flexible column is non-stretchable, incompressible and can be twisted arbitrarily. Bend the flexible column by tensing and relaxing the three driving wires to tilt the support plate. The outer part is covered with a spring to protect the flexible manipulator. The single-section mechanical arm is shown in Fig. 4. Integrated model scheme is showed in Fig.5.
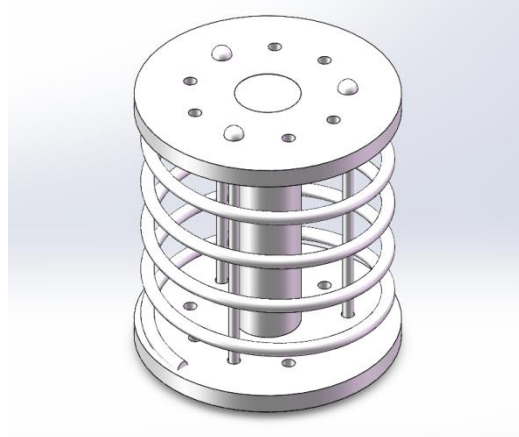
a.   A single segment robotic arm in an upright position
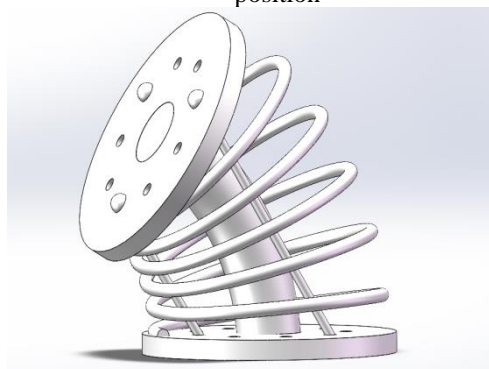
b.   A single segment robotic arm in a bent position
Fig.4 Single segment robotic arm
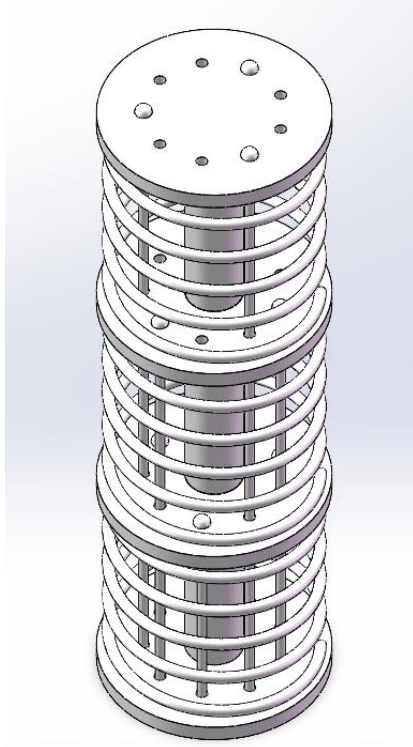
Fig. 5. Integrated model scheme

4. Kinematic analysis of model

The spatial mapping relationship of the robot arm is shown in Fig.6. This paper analyzes the kinematics model of the flexible manipulator through forward kinematics. The forward kinematics analysis of the manipulator arm is to calculate the end pose of the manipulator arm according to the information of the driving space.
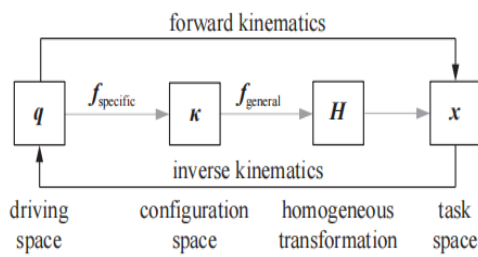
Fig.6 Spatial mapping of kinematic [6]

4.1. *The mapping relationship from driving space to configuration space*

According to the length variation of three wires, the bending angle θ and the rotating angle α and radius of curvature R of the mechanical arm are solved by geometric method. The solution method is to establish a set of equations by using bending angle θ, rotating angle

α and curvature radius R to represent the corresponding arc length and chord length of the projection of the line from the anchor point of the lower support disk to the curvature radius of the single-section mechanical arm, so as to obtain the array of bending angle $θ$ and rotation angle $α$ and curvature radius R represented by the length of the driving wires .The mapping relationship from driving space to configure space is Eq.(1).

$$\begin{cases} \alpha = \arctan\dfrac{\sqrt{3}(H_1 - H_2)}{H_1 + H_2 - H_3} \\ \theta = \arcsin\dfrac{\sqrt[2]{H_1^2 + H_2^2 + H_3^2 - H_1H_2 - H_2H_3 - H_3H_1}}{6r} \\ R = \dfrac{(H_1 + H_2 - H_3)r}{2\sqrt{H_1^2 + H_2^2 + H_3^2 - H_1H_2 - H_2H_3 - H_3H_1}} \end{cases} \quad (1)$$

4.2. *The mapping relationship from configuration space to task space*

Solving the mapping relationship between configuration space and task space that is the homogeneous transformation matrix from local reference frame to global reference frame is solved. Matrix elements are composed of configuration variables α、 θ and R. The solution is to use the D-H method, and the transformation relationship between the coordinate system from the lower support disk to the upper support disk is regarded as a series of sequential rotation and translation results.

D-H method (Denavit-Hartenberg method) is a matrix general method for establishing the relative position and attitude relationship of serial robots. It uses four parameters of connecting link length, joint angle, link offset and link twist to describe the structure and kinematic relationship of the series mechanism. These four parameters are collectively referred to as D-H parameters. The spatial geometric relation of each link relative to the fixed reference coordinate system is described by homogeneous transformation, and the spatial geometric relation between two adjacent links is expressed by homogeneous transformation matrix. The equivalent homogeneous coordinate transformation matrix of the end actuator coordinate system relative to the base coordinate system can be derived by the homogeneous transformation, and the kinematics equation of the robot can be established [7]. Generation steps of homogeneous transformation matrix is showed in Fig.7.
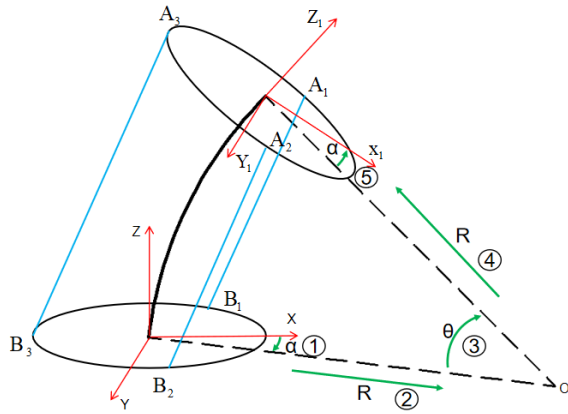
Fig. 7 Generation steps of homogeneous transformation matrix

The final homogeneous transformation matrix from configuration space to task space is Eq. (2), Eq. (3) is the rotation matrix. Eq. (4) is the translation matrix.

$$H(k) = \begin{bmatrix} \theta(k) & \Psi(k) \\ 0 & 1 \end{bmatrix} \quad (2)$$

$$\theta(k) = \begin{bmatrix} cos^2\alpha cos\theta + sin^2\alpha & cos\alpha sin\alpha(cos\theta - 1) & cos\alpha sin\theta \\ cos\alpha sin\alpha(cos\theta - 1) & sin^2\alpha cos\theta + cos^2\alpha & sin\alpha sin\theta \\ -cos\theta sin\alpha & -sin\alpha sin\theta & cos\Theta \end{bmatrix} \quad (3)$$

$$\Psi(k) = [Rcos\alpha(1 - cos\theta) \quad Rsin\alpha(1 - cos\theta) \quad Rsin\theta] \quad (4)$$

5. Flexible manipulator motion space simulation analysis:

The motion space diagram of the single-section flexible manipulator is shown in Fig.8.
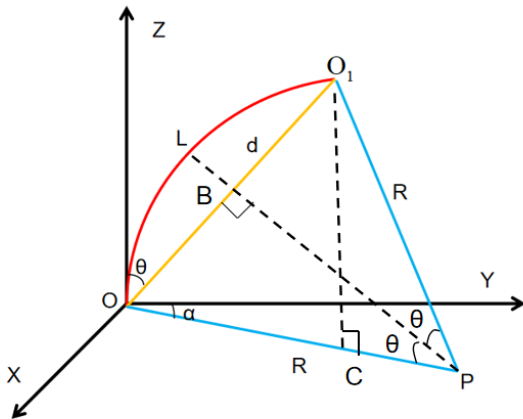
Fig. 8 Space analysis diagram of flexible manipulator motion

The red part OO_1 of the circular arc represents the continuous flexible body. The head end point O is the fixed point, and the end point O_1 is the free moving end. Let the length of the continuous flexible body (spring) be L. The arc OO_1 corresponds to the chord length d. When the continuous flexible body is not subjected to external forces, its initial state is reconnected with the Z-axis. The angle between the projection of arc OO_1 on the *X-O-Y* plane and the positive direction of the *Y*-axis is set as α, and the angle between string d and the positive direction of the *Z*-axis is set as θ.

The free end point O_1 of the center line of the flexible arm is obtained. The expression in the coordinate system is Eq.5. The bending Angle is 0 to 90° and the rotation Angle is 0 to 360°.

$$\begin{cases} x = \dfrac{L}{2\theta}(1 - cos2\theta)sin\alpha \\ y = \dfrac{L}{2\theta}(1 - cos2\theta)cos\alpha \quad (5) \\ z = \dfrac{L}{2\theta}sin\theta \end{cases}$$

The length L of the prototype single section flexible arm in the natural state is set to 200mm, and the end point motion relation is obtained by using MATLAB to get the end point motion range diagram (Fig.9).
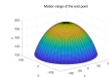
Fig. 9 Flexible manipulator motion space simulation

6. The simulation of length changes of the wires:

The vectors representation of the driving wires is obtained by superposition of the vectors, and then the maximum wire length variation of the single flexible arm is obtained by solving the modulus of the vector. The activity limit of the initial single-section mechanical arm is the rotation angle 0°-360° and the bending angle 0°-90°. Through simulation analysis in Fig.10, when the bending angle is 90° and the rotation angle is from 0° to 360°, the length of the drawstring changes the most. The simulation results are shown in the figure below. The maximum variation value of wires length of single section manipulator is about 95mm-265mm.
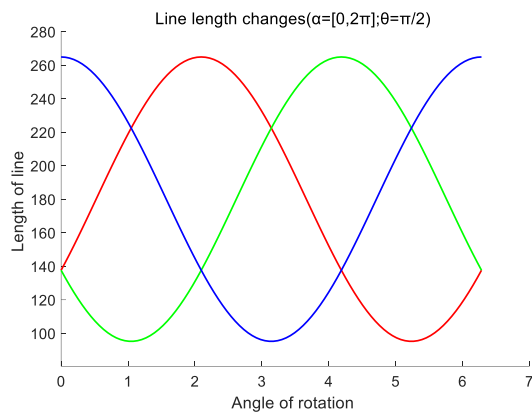
Fig. 10 The simulation of length changes of the wires

Dr. Kazuo Ishii

He received his Ph.D. degrees from Tokyo University, Japan, in 1996. In 2011, he joined Kyushu Institute of Technology, where he is currently a Professor of the Department of Human Intelligence Systems. His research interests include Underwater Robot, Neural Network and Neural Network

References

1. XIAO Xiao. Research Status and Development Trend of Industrial Robot [J]. Wireless Internet Technology,2021,18(23):49-50.

2. Tsinghua University has made new progress in the field of micro flexible robot [J]. Robot Technology and Application, 2023, (02):5.

3. Wang Yurong, Xu Shengxin. SMA driven FangZhang fish more joint research and development of flexible manipulator analysis [J]. Journal of electronic testing, 2020, (3): 18 to 19 + 25. DOI: 10.16520 / j.carol carroll nki. 1000-8519.2020.03.006.

4. Yu Youhe, Li Jian, Zhou Yan et al. Pneumatic driving bionic flexible manipulator design [J]. Mechanical and electrical information, 2022, (17) : 23-26. DOI: 10.19514 / j.carol carroll nki cn32-1628 / tm. 2022.17.006.

5. Zhu Xiaojun, Wang Xueqian, Ma Yunxuan et al. Rope hyper redundant space flexible manipulator displacement teleoperation system design and experimental research [J]. Robot, 2022, 44 (01) : 9 to 18. DOI: 10.13973 / j.carol carroll nki robot. 210228.

6. Engineering - Mechanical Engineering and Machinery; University of Sousse Researchers Publish New Study Findings on Mechanical Engineering and Machinery (Forward Kinematic Model Resolution of a Special Spherical Parallel Manipulator: Comparison and Real-Time Validation) [J]. Journal of Engineering,2020,2955-.

7. TANG Yue, Zheng Jinhui, Men Zhengxing et al. D-h modeling and movement simulation is industrial robot [J]. China's heavy equipment, 2020, (01) : 31-32 + 40. DOI: 10.14145 / j.carol carroll nki. 51-1702 / th. 2020.01.011.

Authors Introduction

Mr. Huang Jiawei

He received his bachelor's degree in engineering in 2020 from the Mechanical design, manufacture and automation; Guangdong ocean university in China. He is currently a master student in Kyushu Institute of Technology, Japan

Research on Driver Drowsiness Detection Method based on Deep Learning

Shi Puwei

Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka 808-0196, Japan

Kazuo Ishii

Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka 808-0196, Japan
Email: shi.puwei537@mail.kyutech.jp, ishii@brain.kyutech.ac.jp

Abstract

Drowsiness driving will pose a serious threat to the lives of drivers and others. Determining the state of the driver through face recognition has the advantages of low cost and convenience. Therefore, this study deploys the face recognition model to the mobile phone, and finally realizes the recognition of the driver's Drowsiness by the front camera of the mobile phone. The whole research is divided into three parts. The first part is to train the face 68-keypoints recognition model based on the YOLO face algorithm. The second part is to deploy the trained model to the mobile phone using ONNX and NCNN. The third part is to calculate EAR and MAR using several facial key points, and finally complete the recognition of the driver's drowsiness state using EAR and MAR.

Keywords: YOLO face, ONNX, NCNN, EAR (Eye Aspect Ratio), MAR (Mouth Aspect Ratio)

1. Introduction

Nowadays, the car has become an indispensable means of transportation, but at the same time, driver drowsiness driving this dangerous driving behavior is common, the driver himself and others have brought indelible damage. Drowsiness driving is due to the driver's high-intensity, long time driving, resulting in excessive energy and physical exertion, and then the reaction ability and control level decreased, affecting the normal driving operation of the situation. The probability of a driver causing a traffic accident in drowsiness increases dramatically. In Canada, more than 20% of traffic accidents involve drowsiness driving. Surveys also show that 20% of traffic accidents in the United States involve drivers who are in drowsiness [1]. Therefore, it is important to study driver behavior and state recognition methods to alert and intervene when dangerous situations occur, to ensure the safety of drivers themselves and other people and vehicles on the road.

The current objective detection methods for drowsiness driving are mainly divided into based on driver physiological parameters; based on driver operating behavior; and based on vehicle driving information. In 2007, Yamaguchi et al. extracted amylase from the driver's saliva and determined whether the driver was in a drowsiness state by analyzing the changes in its composition [2]. In 2013, Mbouna et al. extracted the driver's eye state and pupil movement state as driver fatigue features and used SVM to classify the driver's drowsiness state, achieving a good recognition effect [3].

Mott G E et al. concluded that the angle of the driver's steering wheel and the lateral position of the vehicle had the highest correlation with drowsiness driving by comparing 87 vehicle parameters related to drowsiness driving in waking and moderate fatigue states [4].

As far as current research is concerned, recognition and early warning through physiological parameters is easy to interfere with normal driving and affect the drowsiness change process; identification through vehicle driving information is subject to factors such as driver skills and road conditions, and the recognition accuracy is relatively low. The drowsiness recognition method based on the driver's behavior and state has the characteristics of low equipment cost and non-invasiveness. At present, the recognition of driver characteristics is mainly concentrated in areas such as eyes, mouth and so on. At the same time, the use of cameras for face recognition also needs to consider the impact of changes in light on the recognition ability of the model.

Methods of drowsiness recognition based on face can be divided into two types. One is based on the entire image. In this method, the model will infer and recognize the entire image, and then crop the different facial feature areas. The other is based on facial feature points. The model will annotate the specified number of key points on the entire image, and then infer facial features based on the relationship between key points. Currently, the most common key points annotations include 98 key points, 68 key points, and 29 key points. This study chooses to use the most common 68 key points for model modification and dataset building.

Shi Puwei, Kazuo Ishii

In summary, this study aims to recognize the facial feature points of drivers through computer vision methods, establish a driver fatigue recognition method, and complete the deployment of the model on the smartphone side.

2. YOLO Face-based Face Recognition Model

2.1. *Yolo Face*

YOLO is a classic one-stage target detection algorithm based on CNN (Convolutional Neural Network). The original YOLO algorithm does not contain key points detection. It can only annotate target objects on images with bounding boxes. Later, some authors developed YOLO pose for human pose detection based on YOLO. YOLO pose adds 17 human body key points detection based on YOLO. The YOLO face used in this study is based on YOLO pose, changing the 17 key points to five key points regression. YOLO face treats face detection as a general target detection task. Therefore, its network structure is not much different from the overall YOLO pose, using only some special network modules to replace the modules in the original structure [5].

The original YOLO face includes only 5 key points detection. This study chooses to use 68 key points for extracting face-related features. Therefore, the original YOLO face needs to be modified. The modification is mainly reflected in the data structure. The data structure of each target in the original is a vector of length 16. The length of data vector modified is 138.

2.2. *Dataset for training*

This study combines the 300w-3D dataset and the LaPa dataset to build training and validation sets for the face recognition model [6], [7]. Among them, the 300w-3D dataset has the characteristics of large data volume, accurate annotation, gender balance, age distribution balance, and bright/dark balance. However, the open/closed eye data in the 300w-3D dataset is not balanced, especially the closed eye data is relatively small. This will lead to poor model generalization ability, and eventually fail to detect closed eye data. Therefore, this study screened out the closed eye data from the LaPa dataset and combined it with the 300w-3D dataset to form a new dataset. The model is trained and validated using this new dataset.

2.3. *Training Result*

YOLO-tiny is selected as the network structure. After training for 300 epochs, the model has basically converged (Fig. 1).

a. Loss of training
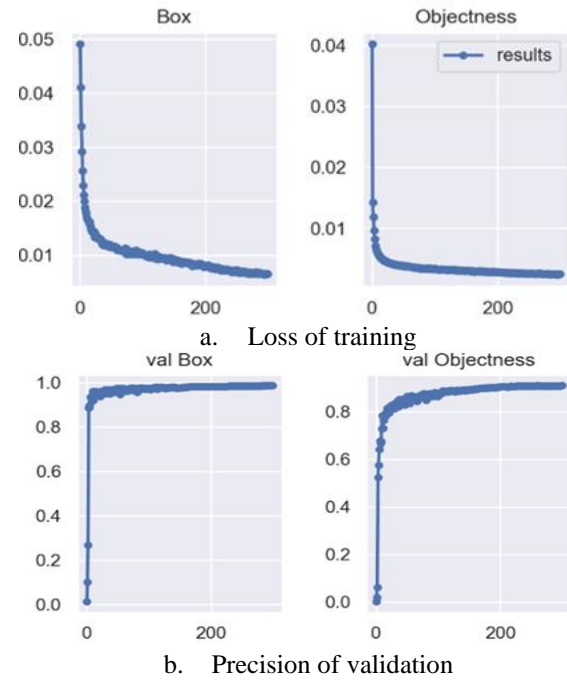
b. Precision of validation
Fig. 1. Result of training

The detection performance of the model is evaluated in terms of mAP (Mean Average Precision). mAP@0.5 represents the change of mAP when the IOU threshold is 0.5, and mAP@0.5:0.95 represents the change of mAP when the IOU threshold gradually increases from 0.5 to 0.95. These two mAPs are both greater than 0.9 after the end of training, and the training effect is good (Fig. 2).
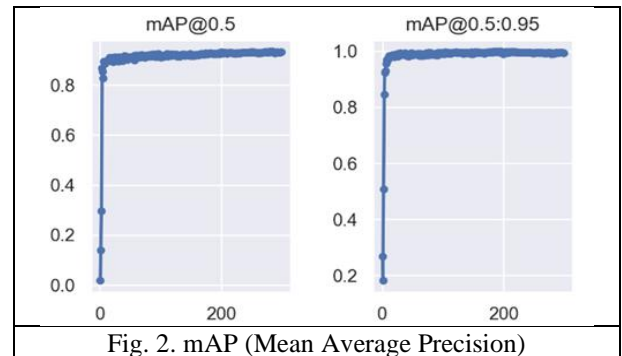
Fig. 2. mAP (Mean Average Precision)

To verify that the model still has good detection capabilities under different lighting conditions, several images are selected, and new images were obtained by decreasing the average brightness of every image by 50 in turn, and they were sent to the model. The results show that the model also has good detection ability in low brightness conditions (Fig. 3).

Fig. 3. Face Detection under Different Brightness

3. Android Smartphone Model Deployment

This study uses ONNX and NCNN as "intermediate representations" to build an inference engine based on the NCNN framework to deploy the trained face recognition model on Android phones.

In general, deep learning models rely on related frameworks and environments, which are not suitable for installation in mobile phones and other environments. In addition, deep learning models require a lot of computing power, which is difficult to run on phones without optimization. ONNX and NCNN as "intermediate representations" can solve this problem. ONNX (Open Neural Network Exchange) was jointly released by Facebook and Microsoft in 2017 for standard description of model structure and parameter calculation graphs. It can optimize network structures and reduce environmental dependencies. NCNN is a lightweight neural network inference framework developed by Tencent for mobile platforms [8], [9]. In short, the trained model needs to be converted into ONNX form and then into NCNN form before it can be deployed to a smart phone. (Fig.4)

Fig. 4. Model transformation process

The effect is shown in Fig. 5, the experimental device is an Android phone equipped with Snapdragon 888Soc (in strong light and low light environments, respectively).
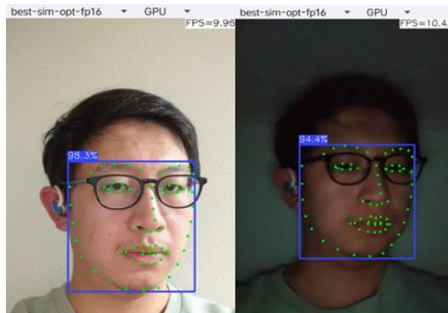
Fig. 5. The effect of the model running on a mobile phone

4. Drowsiness Detection Method by face key points

4.1. *Eye and mouth state recognition based on EAR and MAR*

After completing the extraction of facial feature regions or feature points, it is necessary to study the relationship between the open and closed state of the eyes and mouth and drowsiness based on these characteristics. The relationship was investigated in this study using EAR (Eye Aspect Ratio) and MAR (Mouth Aspect Ratio). They are based on the same principle, the value of them can reflect the opening and closing status of the eyes and

mouth visually. Taking eye as an example, among the 68 facial feature points, each eye occupies 6 feature points, distributed as shown in Fig. 6. [10]
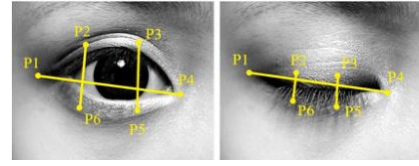
Fig. 6. Feature points of eye

When the eyes are open and closed, the relative position of P1~P6 changes greatly, especially the vertical coordinate changes between the P2/P6 and P3/P5 point pairs are more obvious, and the horizontal coordinate distance between the P1/P4 point pairs basically does not change when the eye state changes. According to the change law of the relative distance between the 6 feature points when the eye state changes, the current state of the eye can be clearly extracted, and the specific calculation is as Eq. (1).

$$EAR = \frac{\|P2 - P6\| + \|P3 - P5\|}{2\|P1 - P4\|} \quad (1)$$

4.2. *Experiment for drowsiness detection based on EAR and MAR*

This study uses the dataset from NTHU Computer Vision Lab for related experiments and analysis. EAR and MAR are computed for each frame of all the videos in the training subset of the dataset. To minimize the effect of noise on the data, exponential smoothing filter and Gaussian filter are applied to the EAR and MAR of each sample respectively. The EAR and MAR graphs can be obtained as shown in Fig.7 and Fig. 8. The downward spike in the graph of EAR is the process of winking, and the upward spike with a certain width in the graph of MAR is the process of yawning.

By analyzing 142 video samples (71 drowsiness samples and 71 non-drowsiness samples), it can be found that the change in the frequency of winking does not reflect the drowsiness state, because there are individual differences in winking habits, and the frequency of winking rises in the drowsiness state of some people and decreases in other people. But there is always a longer period of eye closing when people are in the drowsiness state. Therefore, the presence of significantly longer than normal winking time for eye closing time can be used for drowsiness detection (Fig. 9). In addition, for MAR, the n-sigma rule can be used to distinguish between yawning and normal speaking (Fig.7 and Fig. 10). The existence of yawning can also be used for drowsiness detection. In summary, drowsiness detection methods can be constructed based on yawning and longer eye closing time.

Shi Puwei, Kazuo Ishii
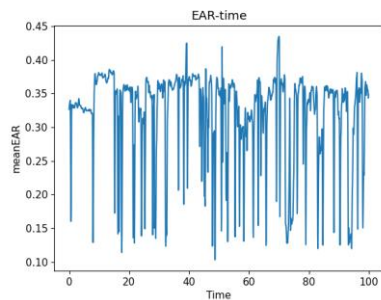
Fig. 7. EAR of a non-drowsiness sample
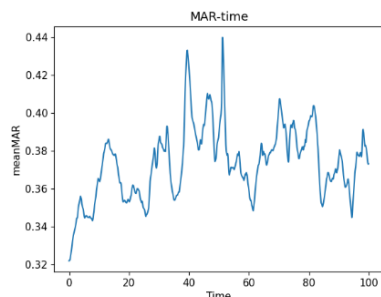
Fig. 8. MAR of a non-drowsiness sample
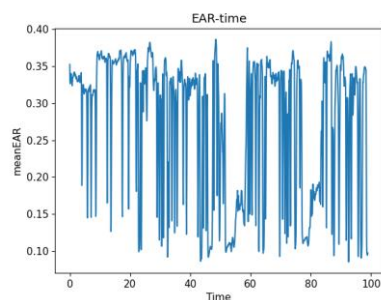
Fig. 9. EAR of a drowsiness sample
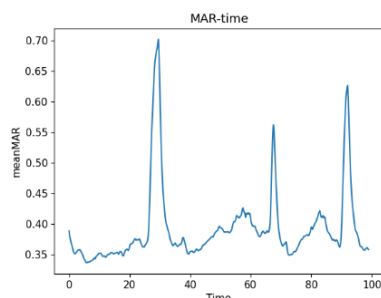
Fig. 10. MAR of a drowsiness sample

5. Conclusion

In this study, a 68-keypoints face recognition model was implemented based on YOLO face, after which the model was deployed to Android phones through ONNX and NCNN frameworks, and finally EAR and MAR were calculated by key points and their relationship with driver drowsiness driving was studied. According to the experiment, it was concluded that based on yawning and longer eye closing time can be used as a standard for drowsiness determination. In the future, it is planned to add this standard into the Android inference engine, and the network structure of the face recognition model is adjusted to improve performance.

References

1. Li Xiaoxing, Deep Learning Based Fatigue Driving Detection Methods, master's thesis, University of Science and Technology of China, 2020.
2. Yamaguchi M, Deguchi M, Wakasugi J, Hand-held monitor of sympathetic nervous system using salivary amylase activity and its validation by driver fatigue assessment, Biosensors& Bioelectronics, 2006, Vol. 21(7), pp. 1007-1014.
3. Mbouna R O, Kong S G, Chun M G, Visual Analysis of Eye State and Head Pose for Driver Alertness Monitoring, IEEE Transactions on Intelligent Transportation Systems, 2013, Vol. 14(3), pp. 1462-1469.
4. Mott G E, Forsman P, Short K R, Efficient driver drowsiness detection at moderate levels of drowsiness, Accid Anal Prev, 2013, Vol. 50(1), pp. 341-350.
5. Delong Qi, Weijun Tan, Qi Yao, YOLO5Face: Why Reinventing a Face Detector, arXiv:2105.12931(CS/CV), 2021, https://arxiv.org/abs/2105.12931.
6. Xiangyu Zhu, Zhen Lei, Xiaoming Liu, Face Alignment Across Large Poses: A 3D Solution, http://www.cbsr.ia.ac.cn/users/xiangyuzhu/projects/3DDFA/main.htm
7. Yinglu Liu, Hailin Shi, Hao She, A New Dataset and Boundary-Attention Semantic Segmentation for Face Parsing, The Thirty-Fourth AAAI Conference on Artificial Intelligence, 2020, pp. 11637-11644.
8. Open standard for machine learning interoperability, https://github.com/onnx/onnx/tree/main
9. A high-performance neural network inference framework optimized for the mobile platform, https://github.com/Tencent/ncnn
10. Tereza Soukupová, Real-Time Eye Blink Detection using Facial Landmarks, 21st Computer Vision Winter workshop, Slovenia, 2016.

Authors Introduction

Mr. Shi Puwei

He received his Bachelor's degree in Automation from Guangxi University, Institute of Electrical Engineering in China.
He is currently a master student in Kyushu Institute of Technology, Japan

Dr. Kazuo Ishii

He received his Ph.D. degrees from Tokyo University, Japan, in 1996. In 2011, he joined Kyushu Institute of Technology, where he is currently a Professor of the Department of Human Intelligence Systems.
His research interests include Underwater Robot, Neural Network and Neural Network.

An Analysis of Translational Motion for a Mobile Robot with Line-Symmetric Rollers Arrangement

Kenji Kimura

Department of Control Engineering, National Institute of Technology, Matsue College,
14-4 Nishi-ikuma-cho, Matsue-shi,Shimane,690-8518, Japan

Kazuo Ishii

Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku, Kitakyushu, 808-0196, Japan
Email: k-kimura@matsue-ct.jp, ishii@brain.kyutech.ac.jp

Abstract

In fields such as logistics, robots are required to have efficient mobility. There are various types of rollers used in mobile robots. Among them, omni-rollers have excellent omnidirectional mobility and are easy to control. In this study, mechanism kinematics has been proposed that assumes arbitrary changes in the roller arrangement position on a circular mechanism and the roller arrangement has been evaluated from the viewpoint of speed efficiency. Furthermore, we aim to evaluate the mobility of mobile robots by focusing on their translational components. Moreover, we have examined the behavior of the area of the area generated by the end point of the robot velocity vector and evaluated the velocity efficiency.

Keywords: Omni-roller, Translational motion, Motion analysis of mobile robot

1. Introduction

In recent years, industries such as logistics require efficient mobile robots for object transport. Development of such mobile robot vehicles has attracted great attention among the research community. The degrees of freedom for omnidirectional movement (if holonomic properties are provided) is three, expressed as the sum of a translational component of 2 degrees of freedom and a rotational component of 1 degree of freedom. The holonomic movement mechanism is easy to control and has excellent maneuverability in all directions due to the independent drive characteristics of the wheels.

In particular, there are many moving mechanisms equipped with three omni rollers, and the kinematics for these mechanisms have been derived [1]. The three-roller arrangement of this mechanism has a basic structure of an equilateral triangle.

This basic arrangement is also used in the soccer robots of RoboCup Meddle-sized-reague, such as RV-infinity [2], Musashi 150 [3], and NuBot [4]. Thus, a regular polygon with the highest degree of symmetry among triangles was used, and no theoretical research had been conducted. However, recently, researchers have started focusing on the kinetic energy of drive rollers for moving mechanisms [5] and spherical conveyance [6].

In a movement mechanism using a sphere as wheels (or a mechanism using a sphere as a conveyance object), the optimal angle of the sphere rotation axis with respect to the movement direction [7] and the placement location of the two drive rollers that drive the sphere should be determined. The optimal location has been identified to be on the equator [8]. In a mechanism using rollers as wheels, a transformation matrix is defined that associates the input roller speed with the output robot speed (translation/rotation), and the "Image volume" and "Orthographic area" are used as an evaluation function for movement efficiency. The roller arrangement was evaluated in a previous study [9].

The present study focuses on sectional area by considering the kinetic energy of a mobile robot in the line-symmetric roller's arrangement which has one-dimensional of freedom. Focusing only on the robot's translational motion, we derive a relational expression between velocity efficiency and roller arrangement, and subsequently analyze the behavior of the end point region area of the robot velocity vector.

The rest of this study is as follows: Chapter 2 discusses the kinematics of mobile robots. Chapter 3 derive sectional area function. Chapter 4 conducted the simulation. Finally, we present the summary and future tasks.

2. Kinematics of Transfer Mechanism in case of isosceles triangle three rollers arrangement

In a previous study, we have defined "Image volume" and "Orthographic projection area" as evaluation functions analysis roller contact location ([9]).

Section 2.1 introduces only on the robot translational motion and analysis robot motion.

Section 2.2 discusses kinematics for a mechanism that adapts three omni-rollers in line symmetry roller's arrangement.

2.1. *Liner Transformation mapping for correspondence of roller speed and robot speed*

As shown in Figure 1(a), the mobile robot that has a common radius of all omni-wheels adapted the *i-th* rollers ($i = 1,2,3$) contact point P_i on a circle. $X-Y$ is the global coordinate system (origin O). The robot translation speed is $V = [V_x, V_y]^T$ and φ denotes robot direction. Robot rational speed $L\dot{\phi}$ ($\dot{\phi}$: robot angular velocity) roller peripheral speed v_i are decomposed as translation and rotational components. Contact point P_i are adapted angle θ_i on the circle that has a radius L. Thus, the correspondence of $[v_1, v_2, v_3]$ and $[V_x, V_y, \dot{\phi}L]$ is represented in Figure 1(a).

As shown in Figure 1(b), linear transformation mapping $f_A : [V_x, V_y, \dot{\phi}L]^T \rightarrow [v_1, v_2, v_3]^T$ and $f_A(W) =$Image$f_A \in \mathbf{R}^3$ is a parallelepiped domain from cubic domain W.

In this study, we represented a sectional area of $f_{A^{-1}}(W)$ for a horizontal plane as follow:

$$f_{A^{-1}}(W) \cap \{V_x V_y - \text{plane}\}$$
$$= \{(V_x, V_y, 0)| \ |v_1|, |v_2|, |v_3| \leq 1\} \quad (1)$$

Where

$$f_{A^{-1}}(W) = \{(V_x, V_y, L\dot{\phi})| \ |v_1|, |v_2|, |v_3| \leq 1\} \quad (2)$$

$$W = \{(v_1, v_2, v_3)| |v_1|, |v_2|, |v_3| \leq 1\} \quad (3)$$

2.2 *Kinematics for isosceles triangle three rollers arrangement*

In our previous work [9], we derived kinematics for arbitrary roller arrangement position as $(\theta_1, \theta_2, \theta_3)$. In this study, we restrain an isosceles triangular roller arrangement (line symmetry roller's arrangement) in Eq. (6) of the previous study [9]

Substituting $\theta_1 = \theta$, $\theta_2 = 360° - \theta$ and $\theta_3 = 0°$ with $(\theta_1, \theta_2, \theta_3)$ in Eq. (6) of [9].

Inverse kinematics ($[v_1, v_2, v_3]^T$ is determined from $[V_x, V_y, L\dot{\phi}]^T$) is represented as follows:

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} -\sin\theta & \cos\theta & 1 \\ \sin\theta & \cos\theta & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} V_x \\ V_y \\ L\dot{\phi} \end{bmatrix} \quad (4)$$
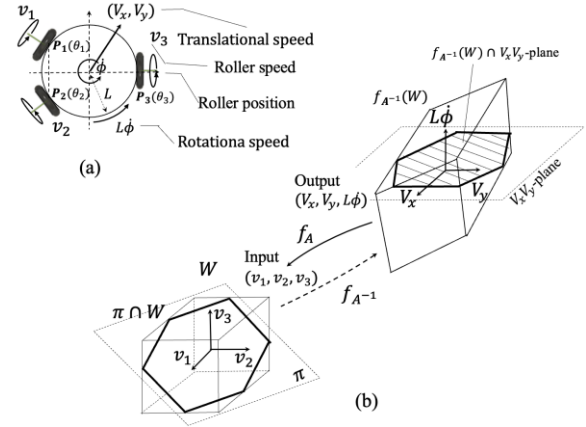
Figure 1 Correspondence between three omni-rollers speed $[v_1, v_2, v_3]^T$ and $[V_x, V_y, \dot{\phi}L]^T$ (robot mobile speed (V_x, V_y) rotational speed $\dot{\phi}$). (a) mobile robot and (b) translation mapping.

Forward kinematics ($[V_x, V_y, L\dot{\phi}]^T$ is determined from $[v_1, v_2, v_3]^T$) is represented as follows:

$$\begin{bmatrix} V_x \\ V_y \\ L\dot{\phi} \end{bmatrix} = \begin{bmatrix} -\dfrac{1}{2\sin\theta} & \dfrac{1}{2\sin\theta} & 0 \\ -\dfrac{1}{2-2\cos\theta} & -\dfrac{1}{2-2\cos\theta} & \dfrac{1}{1-\cos\theta} \\ \dfrac{1}{2-2\cos\theta} & \dfrac{1}{2-2\cos\theta} & -\dfrac{\cos\theta}{1-\cos\theta} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$
$$(5)$$

3. Derive sectional area function

In this section, we calculate area using Eq. (1). Section 3.1 shows analysis of roller's speed space, while Section 3.2 shows analysis of robot speed space.

3.1. *Analysis of roller's speed space*

(A) *Setup for Cubic domain*

As shown in Figure 2(a), the cubic domain is composed of the following eight apexes: $B_1 = (1, 1, 1)$, $B_2 = (-1, 1, 1)$, $B_3 = (-1, -1, 1)$, $B_4 = (1, -1, 1)$, $B_5 = (1, 1, -1)$, $B_6 = (-1, 1, -1)$, $B_7 = (-1, -1, -1)$, $B_8 = (1, -1, -1)$. Additionally, sectional domain denoted by π is symmetrical with respect to the origin O. For the sectional shape to be hexagonal like Figure 2(a), six dots $P(p, -1, 1)$, $Q(-1, q, 1)$, $R(-1, 1, r)$, $S(s, 1, -1)$,

T $(1, t, -1)$, U $(1, -1, u)$ (real parameter : $-1 \leq p, q, r, s, t, u \leq 1$) should satisfy the following condition.

$$p = s, q = t, r = u \tag{6}$$

As the two faces, $B_1 B_2 B_3 B_4$ and $B_5 B_6 B_7 B_8$ (face to face), are parallel, side PQ and TS are parallel and have same length.

(B) Plane representation as roller speed existent set

Focusing on $L\dot{\phi}$ -component of Eq. (5), only transrational motion is equivalent to $\dot{\phi} = 0$.

$$v_1 \sin \theta + v_2 \sin \theta - 2 v_3 \sin 2\theta = 0 \tag{7}$$

Eq. (8) represents a plane equation including origin by three-dimensional combinate $[v_1, v_2, v_3]^T$.

$$\pi = \{(v_1, v_2, v_3) | v_1 \sin \theta + v_2 \sin \theta - 2 v_3 \sin 2\theta = 0\} \tag{8}$$

(C) Normal vector of plane

Eq. (7) is equivalent to following expression.

$$\sin \theta \langle \begin{bmatrix} 1 \\ 1 \\ -2\cos\theta \end{bmatrix}, \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \rangle = 0 \tag{9}$$

Thus. $[1, 1, -2\cos\theta]^T$ is normal vector of plane π. Also, $\sin \theta > 0$ and $-2\cos \theta > 0$ for all $90° \leq \theta < 180°$.

Figure 2(b) shows a section of the cubic domain with respect to rectangle $(B_1 B_3 B_7 B_5)$. It can be observed that normal vectors shift counterclockwise direction from v_3-axis (when $\theta = 90°$) to line $B_3 O$ (when $\theta = 180°$). Thus. $\pi \cap W$ satisfies the following property.

Property
(i) Sectional shape is symmetrical with respect to origin
(a) $\theta_1 = 90°$ $\Leftrightarrow$ Rhomb
(b) $90° < \theta_1 < 180°$ $\Leftrightarrow$ Hexagon

(ii) When $\theta_1 = 120°$ it has minimal area.

3.2. Analysis of robot speed space

Using Eq. (5), image of $\pi \cap W$ by transformation mapping $f_{A^{-1}}$ can be represented as follows:
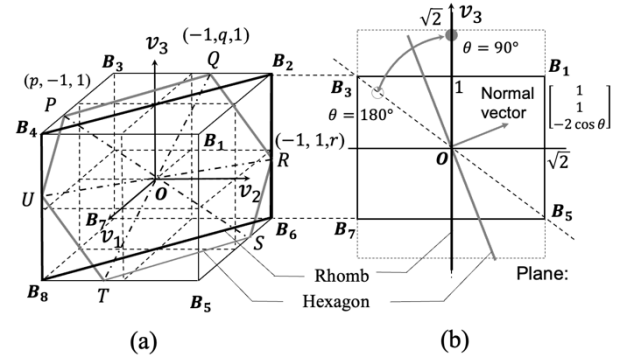
Figure 2 Sectional shape of roller's speed space $\pi \cap W$ in case of an isosceles triangle roller arrangement. (a) Isometric view. (b) Sectional view.

$$f_{A^{-1}}(W) \cap \{V_x V_y - \text{plane}\} = \tag{10}$$

$$\{(V_x, V_y) | v_1 \acute{A}_1 + v_2 \acute{A}_2 + v_3 \acute{A}_3, |v_1|, |v_2|, |v_3| \leq 1\}$$

where

$$\acute{A}_1 = \begin{bmatrix} -\frac{1}{2\sin\theta} \\ -\frac{1}{2-2\cos\theta} \\ \frac{1}{2-2\cos\theta} \end{bmatrix}, \quad \acute{A}_2 = \begin{bmatrix} \frac{1}{2\sin\theta} \\ -\frac{1}{2-2\cos\theta} \\ \frac{1}{2-2\cos\theta} \end{bmatrix}, \quad \acute{A}_3 = \begin{bmatrix} 0 \\ \frac{1}{1-\cos\theta} \\ -\frac{1}{1-\cos\theta} \end{bmatrix} \tag{11}$$

From Eq. (6) and property of transformation mapping $f_{A^{-1}}$, $f_{A^{-1}}(\pi \cap W)$ is decomposed as three parts.

$$f_{A^{-1}}(\pi \cap W) = 2 f_{A^{-1}}(\triangle OPQ) \tag{12}$$

$$+ 2 f_{A^{-1}}(\triangle OQR) + 2 f_{A^{-1}}(\triangle ORP)$$

Where

$$\pi \cap W = 2(\triangle OPQ + \triangle OQR + \triangle ORP) \tag{13}$$

Thus. Sectional area of $f_{A^{-1}}(W) \cap \{V_x V_y - \text{plane}\}$ is represented as following.

$$D_{Sec}(\theta) = \| f_{A^{-1}}(P) \times f_{A^{-1}}(Q) \|$$

$$+ \| f_{A^{-1}}(Q) \times f_{A^{-1}}(R) \| + \| f_{A^{-1}}(R) \times - f_{A^{-1}}(P) \| \tag{14}$$

Where $P(p, -1, 1)$, $Q(-1, q, 1)$, $R(-1, 1, 0)$

Kenji Kimura, Kazuo Ishii

$$p = \frac{\sin\theta - \sin 2\theta}{\sin\theta}, q = \frac{-\sin\theta - \sin 2\theta}{\sin\theta} \qquad (14)$$

4. Simulation

This section presents the simulation findings, including the evaluation values of "horizontal cross-sectional area: D_{Sec}"

Simulations were performed at the various symmetric roller arrangement (Eq. (14) in $90° \leq \theta < 180°$).

As shown in Figure 3, $D_{Sec}(90°) = 4.05[m/s]^2$ and $D_{Sec}(180°) = \infty \ [m/s]^2$. minimum value is $D_{Sec}(120°) = 3.85[m/s]^2$. Thus. The worst efficiency roller arrangement in only translational motion is the equilateral triangular shape.
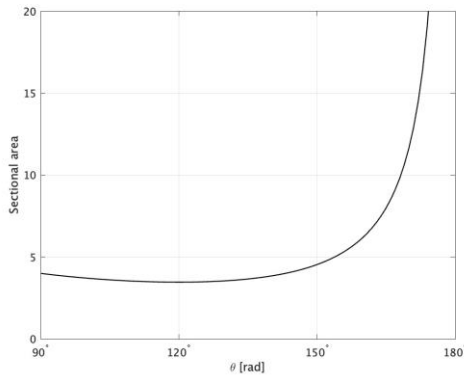
Figure 3   Behavior of Area of domain of end point of robot speed $D_{Sec}(\theta)$ ($90° \leq \theta < 180°$).

5. Conclusion

In this research, we focused on the translational motion of mobile robots. Furthermore, we derived the area of the region where the end point of the robot's velocity vector exists as an evaluation function, investigated its behavior, and determined the minimum value. We get fact that worst efficiency roller arrangement is the equilateral triangular shape.

This is a one-parameter problem because the roller arrangement is given as an isosceles triangle.

Acknowledgements

Reference

[1] J.Tang, K.Watanabe, et al., "Autonomous control for an omnidirectional mobile robot with the orthogonal-wheel assembly," Journal of the Robotics Society of Japan. Vol. 17, No. 1,pp. 51-60, 1999.

[2] Y.Yasohara, K.Shimizu et al., Development of ball handling mechanism for RoboCup MSL, *30th Fussy System Symposium*, pp. 616-617, 2014.

[3] S. Chikushi, M. Kuwada, et al., Development of Next-Generation Soccer Robot"Musashi150"for RoboCup MSL,*30th Fussy System Symposium*, pp. 624-627,2014.

[4] R. Junkai, X. Chenggang, X. Junhao et al., A control system for active ball handling in the RoboCup middle size League, Chinese Control and Decision Conference(CCDC.)2016.

[5] K. Kimura,   K. Ishii, Efficiency Problem of Spherical Robot in Transfer Kinetic Energy, Journal of Robotics, Networking and Artificial Life, Vol.9, issue 1, pp.87-92, 2022.

[6] K. Kimura, Koki Ogata, Hiroyasu Hirai, Takumi Ueda, K. Ishii, Forward Kinematic of a Sphere Considering Slipping and Motion Analysis in Three Rollers, Journal of Robotics, Networking and Artificial Life, Vol. 10, issue1, pp.25-32.2023.

[7] K.Kimura, Y.Abematsu, H.Hirai, K.Ishii, Evaluation of Two Roller Arrangement on a Hemisphere by Kinetic Energy, Journal of Robotics, Networking and Artificial Life, Vol.9, issue 3, pp.233-239, 2022.

[8] K. Kimura, K. Ishii, The Spherical Robot Transfer Problem With Minimal Total Kinetic Energy, Proceedings of International Conference on Artificial ALife and Robotics (ICAROB2021), pp.266-270, 2021.

[9] K. Kimura, Y. Shigyo, K. Ishii, An Analysis of Robot Speed Efficiency for Mobile Robot Adapted Three Omni Rollers Using Linear Transformation, Journal of Advances in Artificial Life Robotics Vol. 3,issue4, pp. 242–249.2023.

Authors Introduction

Dr. Kenji Kimura

He is a Lecturer at the Department of Control Engineering, National Institute of Technology, Matsue College, and a Visiting Associate Professor at Chuo University. He received his ME (mathematics) from Kyushu University in 2002 and his PhD (engineering) from the Kyushu Institute of Technology in 2020. His research interests are spherical mobile robots.

Dr. Kazuo Ishii

He is a Professor in the Kyushu Institute of Technology, where he has been since 1996. He received his Ph.D. degree in engineering from University of Tokyo, in 1996. His research interests span both ship marine engineering and Intelligent Mechanics. He holds five patents derived from his research.

Development of Teaching Materials for Robot Programming for Junior High School Students: Student-Based Educational Activities

Kenji Kimura

Department of Control Engineering, National Institute of Technology, Matsue College,
14-4 Nishi-ikuma-cho, Matsue-shi,Shimane 690-8518, Japan

Youta Takano

Department of Control Engineering, National Institute of Technology, Matsue College,
14-4 Nishi-ikuma-cho, Matsue-shi,Shimane 690-8518, Japan
Email: k-kimura@matsue-ct.jp, y8otaman@gmail.com

Abstract

With the extensive use of robots in recent years, an age of co-existence between humans and robots is expected to arrive in the future. Therefore, providing robot education in the early stages of elementary and junior high school is necessary to stimulate students' interest in robots. Furthermore, educational institutions are becoming more active in robot education as a part of their contribution to the local community. For elementary school students, a Beauto Racer (Vstone) has been adopted as a teaching material for robot education, and teaching materials have been developed for teaching line tracing. In this study, we will develop educational materials for junior high school students on maze exploration using a Beauto Rover (Vstone). The students of Matsue National College of Technology took the initiative in this activity. The purpose is to improve their teaching skills.

Keywords: Robot education, Programing, Maze search, Student-based educational activities

1. Introduction

AI robotic research is being promoted at an accelerated pace, as obvious from RoboCup's goal of "building an autonomous robot team that can beat the world champion soccer team in 2050" [1]. Thus, an era in which humans and robots will live together in harmony is expected in the near future; therefore, providing a strong awareness regarding robots to future children at an early age is critical. As shown in Table 1, many lesson designs use commercially available microcomputers.

As an example of robotic education using Lego Mindstorms NXT, the company offers numerous courses in robotic education for students belonging to first to sixth grades. This includes student-centered robotic contests centered on line tracing [2].
Furthermore, robotic education using Beauto Racer (Vstone) is being developed as a curriculum for elementary school students to learn programing on the theme of line tracing [3].
In a junior robotic contest, participants used Lego Mindstorms EV3 to perform line tracing for simulating tomato harvesting [4].
Moreover, embedded programing using Lego Mindstorms NXT was employed to teach maze exploration programing [5].

Table 1 Commercial microcontrollers for robotic education

Commercially available microcontrollers	Lesson content	Literature
Lego Mindstorms NXT	Line trace Maze Search	[2][5]
Beauto Racer (Vstone)	Line trace	[3]
Lego Mindstorms EV3	Line trace	[4]

While most of these efforts are faculty-led, there are also student-led collaboration activities in the community that are aimed at improving the basic skills of students and other members of society ([6]).

A student-led working group was organized in this study to produce educational materials that will enable students to quickly and easily learn maze search programming.

The rest of this study is as follows: Chapter 2 introduce information on lesson design and microcontrollers to be used for maze exploration programming . Finally, we present the summary and future tasks.

Kenji Kimura, Youta Takano

2. Planning and Preparation

2.1. View trovers and educational equipment

A Beauto Rover (Vstone) was used as the microcontroller for robotic education. To organize multiple wiring, a body is designed for it, as shown in Figure 1. It was designed using Slid Works (Figure 1(a)), prepared for processing by SPACEGEAR-U44 (MAZAK) (Figure 1(b)), and laser was processed using an acrylic plate (Figure 1(c)). This plane has a rectangular shape with dimensions of 90 [mm] × 125 [mm], and an image of the microcontroller after attaching the protective cover is provided in Figure 1(d).

As shown in Figure 2, a single infrared sensor (①) was mounted on the head of the mechanism. First, the power switch of the microcontroller (②) was turned on, and then a program created by an application on a PC was sent to the CPU board (③) by connecting it to a USB connector (④). Finally, the program proceed button (⑤) to execute the program is turned on, and the front tires (⑥) rotate to perform the motion on execution.

(a) (b)

(c) (d)
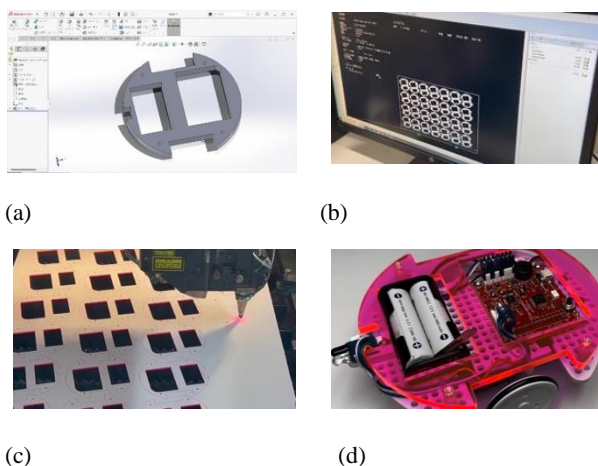
Figure 1 Process of creating a protective cover for the microcontroller: (a). designing using SolidWorks, (b). preparation for laser machining, and (c). laser processing. (d). Microcontroller with a protective cover mounted
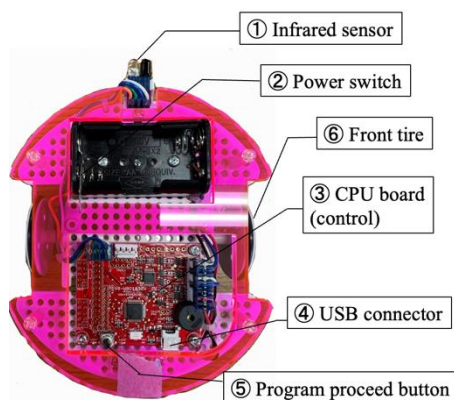
Figure 2 Function of each part of the robot

2.2. Lesson design for maze exploration programing

Assuming that we have lectured on how to transfer a program to the microcontroller and execute it, we will prepare problems 1–5, which are the key problems to be solved in a maze search. Table 2 shows the maze search programing materials (powerpoint-slide) prepared for the lecture.

The first problem is "1 second forward" and "turn right (90-degree rotation)." Participants were asked to think about the number of rotations and how much they have rotated with respect to time (Figure 3(a)).

For problem 2, the exercise is to repeat problem 1 four times, where the useful "loop" function is taught and guidance is provided (Figure 3(b)). For problem 3, students are taught about conditional branching by helping them understand the relationship between senders and the IF function (Figure 3(c)).

A driving program is given as a mission in problem 4 to avoid obstacles when they are placed in predetermined positions (this would use the knowledge from problems 1-3) (Figure 3(d)).

Problem 5 is implementing a program that will work with different placements of obstacles (note the difference from problem 4) (Figure 3(e)).

Table 2 Class design

	(data) item	Notes on the Guidance
Question 1	Actions for going straight and turning right.	Understand the relationship between the number of revolutions and time required to make a 90° turn (relationship between the number of revolutions and angle of the turn)
Question 2	Iterative operation	Ability to use the loop function.
Question 3	IF Function	Ability to use the role of sensors and IF syntax
Question 4	Traveling along an approximate known route	Summary of stationary, straight ahead, right/left turn (basic driving)
Question 5	Unknown Maze Running	Understand the difference between driving on an approximate known pathway

(a)

(b)

(c)

(d)

(e)
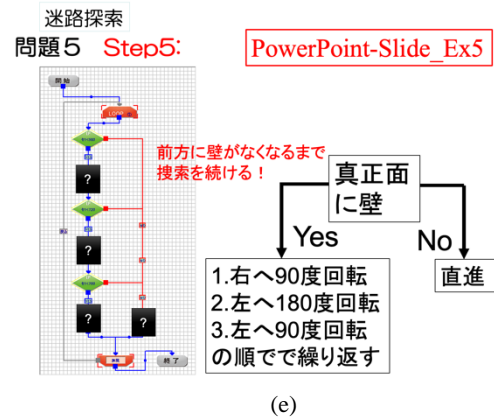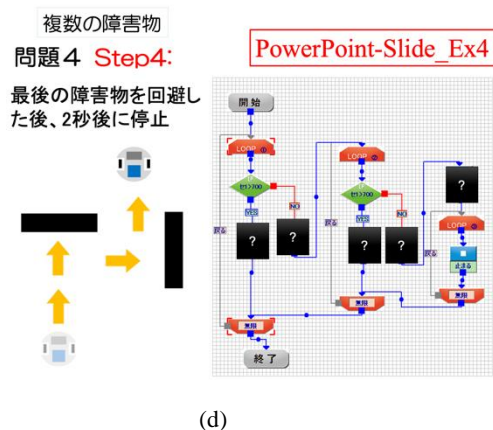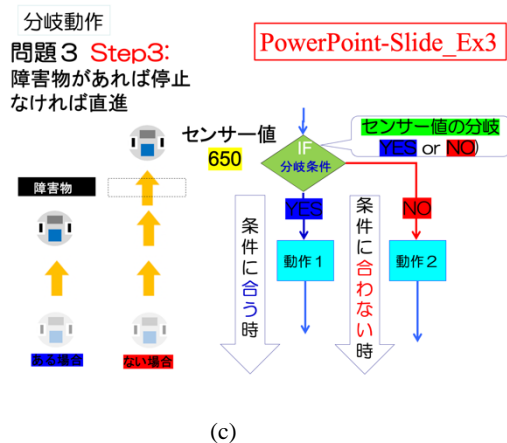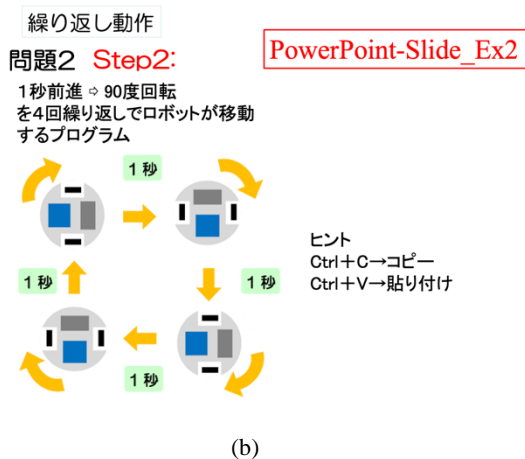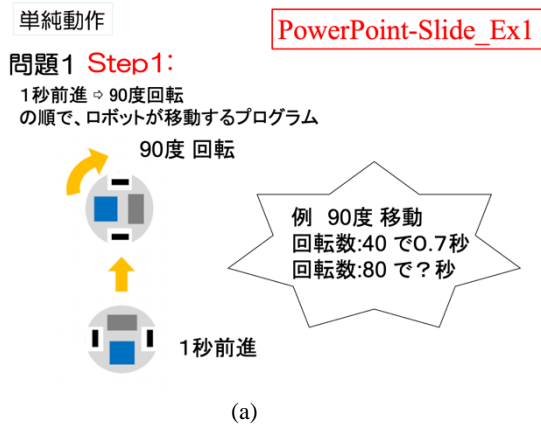
Figure 3 Maze search programming materials (powerpoint-slide for lecture) and five Processes. (a). Basic driving, such as going straight, turning right, and turning left. (b). Iterative operations, (c). IF syntax, (d). multiple obstacle avoidance, and (e). maze search.

2.3. Maze selection

Obstacles were selected to create a maze, and three materials used for creating a maze are listed in Table 3. First one is recycling boxes, which are all around us and have a flat surface. They are relatively responsive to sensors, but their nonuniform size makes it difficult to freely change the maze path (Figure 4(a)). Further, in the case of paper cups, the size is uniform and its unit price is low. However, its shape resembles a part of a cone, and the infrared sensor light reflects irregularly and penetrates the gap between neighboring paper cups (Figure 4(b)). Finally, the sensors responded well in the cubic box case dimensions: 100 mm × 100 mm × 100 mm). This is because they have flat surface, and the infrared sensor responded relatively well because of its uniform size. Furthermore, with cubic box, it is easy to freely change the maze path (Figure 4(c)). Based on the above, cubic boxes were adopted as the obstacle for preparing maze.

The maze corridor's width should be set at 200 mm, which is equivalent to the width of two cubes. This allows the robot to turn right and left in the passageway.

(a)

Kenji Kimura, Youta Takano

(b)
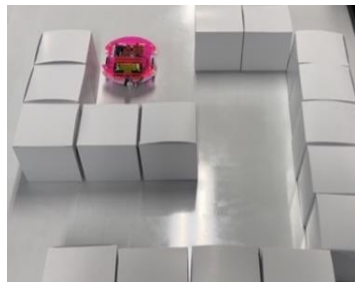
(c)

Figure 4 Maze production choice: (a). recycled box, (b). paper cup, and (c). cube box

Table 3 Three candidates for maze production

Candidate obstacles	Advantage	Disadvantage
Recycling bin	Gather easily	Uneven size
Paper cup	Low unit price Uniform size	Sensor recognition is unstable
Cube box [100 mm per side] (commercially available)	High degree of freedom for placement Uniform size	High cost

3. Conclusion

In this study, a Beauto Rover (Vstone) was used as the microcontroller for robot programing. With only one infrared sensor attached to the back of its head, this microcontroller can be easily operated by junior high school students. To make robot programing learning feasible, we produced programing materials that provided steps for realizing a successful maze exploration.

We plan to conduct several workshops at our university using the developed robot programming material for future research projects. In addition, we plan to conduct pre- and postlecture questionnaires to investigate changes in interest and awareness on maze exploration and to evaluate the educational materials we have developed.

References

[1] I. Noda, H. Minakata, K. Kobayashi, RobotCup Aiming for the Year 2050 (Part 1), Journal of Japan Society for Fuzzy Theory and Intelligent Informatics, Vol. 29, No. 1 (2017) pp. 2-13.

[2] T. Hakata, H. Shibari, Y. Shimada, H. Otsuka, M. Nagata, T. Matsumoto, H. Yamamoto, Education of Embedded System for Senior Students via Line Tracer Robot, Research Reports of Kumamoto-NCT, Vol. 3 (2011) pp. 29-35.

[3] N. Tominaga, Development of Programing Learning Curriculum that Targets Elementary School Students, Ritsumeikan Teacher Educational Studies, Vol. 4 (2014) pp. 81-90.

[4] T. Matsuo, Y. Takemura, T.Sonoda, et al.Toward Smart Tomato Greenhouse:The 6th Tomato-Harvesting-Robot Competition and Regulation Changes Aiming at Practical Application, Journal of Robotics, Networking and Artificial Life, Vol. 9, No. 1 (2022) pp. 13-19.

[5] Y. Takemura, T. Kawasaki, K. Okazaki, Discussion about the Effect of the Manufacturing Education University Student using Robots, The 29th Fuzzy System Symposium, (2013) pp. 310-312.

[6] H. Yonemori, N.Tsuguta, D. Sugiyama, Student-Based Regional Collaboration Activities, Japanese Society for Engineering Education, (2020) pp. 68-69.

Authors Introduction

Dr. Kenji Kimura

He is a Lecturer at the Department of Control Engineering, National Institute of Technology, Matsue College, and a Visiting Associate Professor at Chuo University. He received his ME (mathematics) from Kyushu University in 2002 and his PhD (engineering) from the Kyushu Institute of Technology in 2020. His research interests are spherical mobile robots.

Mr. Youta Takano

He is 5th year student at the Department of Control Engineering, National Institute of Technology, Matsue College. His research interests include educational material for robots and STEM education.

Design and Software Production of Robotics Educational Design
for Elementary and Junior High School Student

Youta Takano

Department of Control Engineering, National Institute of Technology, Matsue College, 14-4 Nishi-ikuma-cho,
Matsue-shi, Shimane,690-8518, Japan

Kenji Kimura

Department of Control Engineering, National Institute of Technology, Matsue College
Email: d1923@matsue-ct.jp, k-kimura@matsue-ct.jp,

Abstract

The Ministry of Education, Culture, Sports, Science and Technology (MEXT) has been promoting cross-curricular learning including STEAM education in its educational policy for 2021. The purpose of this study is to have students experience not only control engineering but also basic programming techniques, and to have them become interested in mathematical subjects in general, which are the basis for control engineering. The educational design and accompanying software were designed using Beauto Balancer2 (Vstone) educational robot and Scratch, a visual programming editor developed by the Scratch Foundation. In addition, workshops were conducted, and the lesson design was evaluated by a questionnaire.

Keywords: Robot education, Control engineering, Visual Programming, Beauto Balancer

1. Introduction.

The Ministry of Education, Culture, Sports, Science and Technology (MEXT) has made programming education mandatory for elementary and junior high schools starting in 2021. The purpose is to develop children's programming abilities, discover their hidden potential, and create opportunities for them to play an active role in society in the future [1].

In addition, the Ministry of Education, Culture, Sports, Science and Technology (MEXT) promotes STEAM education and other cross-curricular learning to apply learning in each subject and other areas to discovering and solving problems in the real world. This is aimed at fostering problem-solving skills that transcend subject boundaries [2].

Hirata developed an experimental teaching material that enables students to identify control objects in both the time and frequency domains and to design control systems while considering the relationship between them [3]. This teaching material is intended for beginning students of control engineering.

Kato conducted an objective verification of educational effectiveness through robot building and robot contests. Here, as an objective method of measuring effectiveness, criteria for behavior evaluation were established and implemented, and the behavior of the subjects was evaluated by TAs [4]. This study cites education in the setting of a robotics class for elementary and junior high school students.

In recent years, it has become necessary to use some form of programming in controlling robots. However, there are few educational materials for elementary and middle school students that can help them understand and become interested in programming and control engineering concepts. There have been no control engineering-related studies targeting elementary and middle school students or aiming to educate them about new concepts through robotics.

In this study, we will develop extensions to existing educational robots and plan lessons that will enable elementary and junior high school students to become interested in a wide range of fields and challenge themselves to learn. The robot used in this study, "Beauto Balancer2 (Vstone)," is a teaching material aimed at learning control. Therefore, it is necessary to develop new software to use it as a teaching material for programming classes.

2. Creation of home-grown software

2.1. *Examination of specifications for home-grown software*

In this study, Beauto Balancer 2 (Vstone) will be used.
Beauto Balancer 2 comes with the factory software (Balancer 2 Programmer).

Youta Takano, Kenji Kimura

The objectives sought for the home-made application used in this study are as follows.

· To have the robot embody the behavior envisioned by the subject.
· The difficulty level of the application should be such that elementary and junior high school students can operate it without difficulty.
· Use of libraries (written in C) provided by Vstone

To achieve these goals, we have defined the functions to include or coexist with the functions of the original software (Balancer 2 programmer) and to introduce the concept of visual programming so that elementary and junior high school students can easily operate the software.

To achieve this, the following three compositions and description methods were considered.

A   Desktop application only (written in Python)
B   Console application (written in C)
C   Scratch script + console application.
       + Genuine software (written in C)

Table 1   Study table for the software specification

	advantage	disadvantage
A	Use of external libraries facilitates UI tuning	The stock library needs to be re-written in advance. Even if the library is ported to Python, there is no guarantee that the serial communication system will work properly.
B	User operation unit (so-called front end) and the language of the internal processing part (i.e., the bag-end) Uniformity is possible, allowing for flexible design and description.	The nature of console applications limits UI tuning. High degree of difficulty in introducing visual programming concepts.
C	Troubleshooting is easier because software is created or diverted by dividing functionality.	Software tends to be somewhat cumbersome to configure, as software is divided by function, which is less consistent than other proposals

Table 1 shows the results of the study and review of the specifications (A, B, and C). Initially, we adopted proposal A, giving priority to the user interface (UI). However, the communication problems could not be solved, and the development process would have been too large, so we gave up on it. Next, they decided to adopt proposal C and develop it in earnest, as it was judged to be more effective in terms of man-hours among the three proposals.

2.2. *Software Development*

As a result of the study in section 2.1, we adopted C and produced two applications. The overall system configuration and process flow are shown in Figure 1. The first is a visual programming editor produced using Scratch 3.6. The second is a communication software written in C.

The flow of the entire system, its usage, and its structure are described. First, a txt file was generated from the visual programming editor, and using Scratch's original block definition function, a block was created to add a string to the list according to the contents of the action (e.g., forward, stop). After the participants created their own actions, we had them generate them using Scratch's txt file output function for the list contents.

Next, we asked participants to start up the communication software. To reduce the burden on the participants, we implemented the software so that it could be automatically loaded and executed by simply clicking on the application to open it.

When the communication software is opened, it automatically establishes USB HID communication and reads the txt file. At this time, we designed the software to read the txt file sequentially and issue communication commands according to the contents of each line.
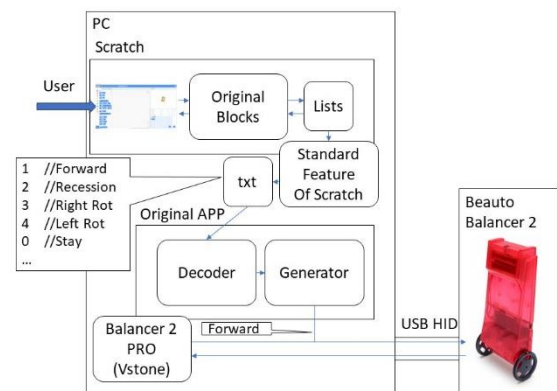
Figure 1   Overall system design

3. Mock Class for Experiments

3.1. *Design of mock class*

A mock class was conducted for elementary and junior high school students (hereinafter referred to as participants). The evaluation was based on the results of the participants questionnaires. The dates of the experiments were as follows Table 2.

Table 2   Planning mock classes

Date	participant
December 16, 2023	4 persons (Breakdown: Fourth grade. 2 students, (one each in sixth grade and eighth grade))
Title. Student-centered lectures and hands-on practice ♪ 〜Let's feel mathematics, physics, and control by using butte balancer! 〜	

The author was the instructor and one teaching assistant (TA), and the maximum number of participants was set at five. Each mock class was held as a part of a craft class hosted by the school, and participants were recruited from the public through posters distributed to elementary and junior high schools in Shimane Prefecture. The aim of the classes was to encourage students to become interested in mathematics, physics, and all other fields by experiencing the mechanisms that make objects move as desired. The mock class consisted of three steps (Step.1-3).

Step.1 "Let's get used to the robot! "
After giving the students the course materials and explaining how to use the robot and its software, we ask them to touch the robot. Since this is the first time for the participants to touch the machine, we ask them to touch it for the time being to familiarize themselves with it. At this point, we ask them the question, "What is gain? and asked them to answer the question while touching the genuine software. Figure 2 is a slide showing how to use Beauto Balancer 2.
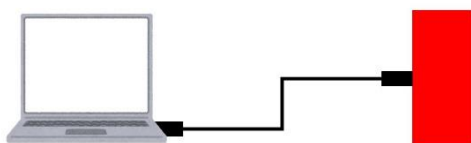
Figure 2   Slide of Class Step 1

Step.2 "Let's feel physics and mathematics! "
Next, mathematics, physics, and control engineering were explained to elementary school students in an easy-to-understand manner using specific examples. At first, basic knowledge and concepts of mathematics and arithmetic (functions, derivatives, and integrals) were explained as necessary knowledge to explain physics concepts. Figure 3 shows a slide explaining integration.
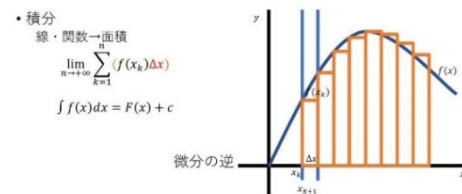
Figure 3   Slide of Class Step 2

Step.3 "Move the robot as you wish! "
We defined the movements we wanted the subjects to perform with the robot and asked them to perform them in sequence. This was done using special software. The explanation was given while displaying the Scratch screen Figure 4.
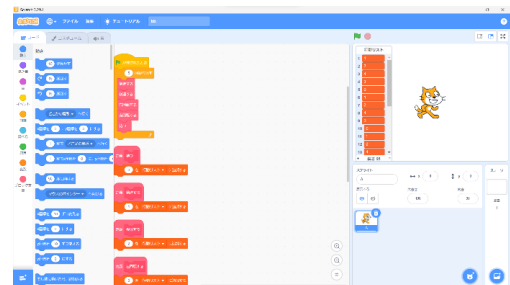
Figure 4   Demonstration of Class Step 3

3.2. *Evaluation of mock class*

To test the effectiveness of the program, a pre-class questionnaire and post-class questionnaire are administered to the subjects. Subjects were asked 5 questions before and after each class. The questionnaire was given on a 5-point scale, with the higher the number, the more positive the evaluation. Table 3 shows the contents of the pre-lesson questionnaire and Table 4 shows the post-lesson questionnaire.

Table 3   Pre-Class Questionnaire

Q1: Do you like mathematics?
Q2: Do you like science?
Q3: Can you visualize the control?
Q4: Are you concerned about what math and arithmetic. learned in school is used for?
Q5: Wha science is used for in school. Are you curious?

Youta Takano, Kenji Kimura

Table 4   Post-Class Questionnaire

Q1: Have you developed an interest in mathematics?
Q2: Were you interested in science?
Q3: Did you have an interest in control engineering?
Q4: Did you want to know more about what math and. arithmetic is used for in school?
Q5: What is the science you learn in school used for? Do you want to know more about it?

Figure 5 shows the results of the questionnaire. The horizontal axis represents each question 1-5 in the questionnaire and the vertical axis represents the average of the response results.
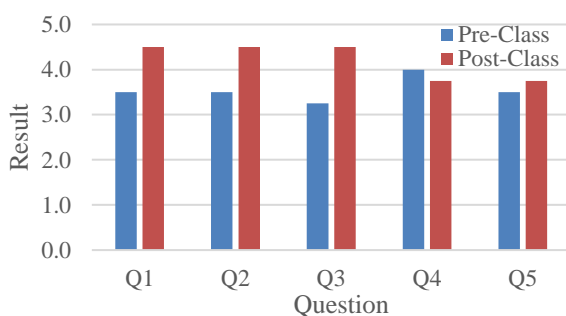
Figure 5   Questionnaire results

3.3. *Discussion of Verification Results*

From Figure 5, the percentage increase in the survey results was calculated and is shown in Table 5. As a result, there is a large increase in some areas. However, Q4 can be taken as showing a decrease.

Table 5   Percentage increase in survey results

	Q1.	Q2.	Q3.	Q4.	Q5.
Rate of increase [%]	29	29	39	- 6	7

Due to the nature of the mock classes held this time, it is difficult to conduct a long-term survey such as a follow-up survey, and the shortcoming is that it is inevitably a short-term survey. In addition, because the recruitment method was open-ended, the results were likely to be influenced by the characteristics of the subject population. For example, if the subjects were junior high school students planning to take entrance exams to difficult-to-enter private high schools, the results of the survey would have changed significantly. Therefore, it is necessary to gather subjects with similar characteristics and ages, and to conduct long-term observation using their parents' and academic records.

4. Conclusion.

In this study, we developed software for robot education for elementary and junior high school students and planned and verified mock classes. The results of the verification showed that the level of interest and concern increased.

In the future, we plan to conduct mock classes with subjects with similar characteristics. In addition, we plan to conduct a long-term post-lesson survey.

Acknowledgements

Reference

[1] Ministry of Education, Culture, Sports, Science and Technology, Guide to Elementary School Programming Education (Third Edition) ,2020.

[2] Ministry of Education, Culture, Sports, Science and Technology, Promotion of STEAM education and other cross-curricular learning, 2021.

[3] Mitsuo Hirata, An experimental apparatus for control education to understand the relation between frequency domain and time domain, Proceedings of the 57th Automatic Control Union Conference, 2014, pp.606-609.

[4] Koichi Kato, et al.. Educational Effects of Creating a Robot and Participating Robot Contest, Annual reports of the Faculty of Education, Gunma University, series issue 46, 2011, pp.125-147.

Authors Introduction

Mr. Youta Takano

He is 5th year student at the Department of Control Engineering, National Institute of Technology, Matsue College. His research interests include educational material for robots and STEM education

Dr. Kenji Kimura

He is a Lecturer at the Department of Control Engineering, National Institute of Technology, Matsue College, and a Visiting Associate Professor at Chuo University. He received his ME (mathematics) from Kyushu University in 2002 and his PhD (engineering) from the Kyushu Institute of Technology in 2020. His research interests are spherical mobile robots.

Development of visual inspection system for low-reflective material utilizing of a string shadow

Keiji Kamei, Tomorou Kawahara
Department of Information Systems, School of Engineering, Nishinippon Institute of Technology
1-11 Aratsu, Kanda, Miyakogun, Fukuoka 800-0394, Japan

Yoshiyuki Daimaru
Nissan Motor Kyushu Co. Ltd., 1-3 Shinhama-cho, Kanda, Miyakogun, Fukuoka 800-0395, Japan
Email: kamei@nishitech.ac.jp

Abstract

Visual inspection of products in manufacturing plant is usually conducted by workers. However, that inspection process by workers occasionally happen oversight due to careless check, so that automatic visual inspection system is needed from manufacturer. Products made of low-reflective material are difficult to inspect by light reflection because contours of defect areas are got blurred. To solve this problem, we propose to inspect products utilizing of a string shadow. Form our research, a string shadow was different between contours of defect area and its of non-defect. On the other hand, former research did not success to detect defect area due to determination algorithm for defect. In this paper, we apply linear approximation of a string shadow to defect detection. From the experimental results, we succeed in detecting defects.

Keywords: Visual inspection system, low-reflective material, string shadow

1. Introduction

In the manufacturing industry, visual inspection in manufacturing plants is used to ensure the quality of products and parts. Most of the current visual inspections are performed by workers. However, visual inspection process by workers occasionally causes human errors such as variation in pass/fail criteria for defective parts or overlooking defective parts due to individual differences. For those reasons, demand for development of visual inspection system by image processing has been increased [1]. A template matching inspection method based on image recognition is an automated method of visual inspection [2], [3]. In inspection by template matching, the shape and location of the inspection target must be fixed. That method is an effective method if inspection items are simple and will not change significantly in the future. However, it is not suitable for the visual inspection system because it requires large cost when there are complex decisions, multiple types of inspection items or flexible responses to market trends.

In recent years, many studies has been conducted on product inspection systems that apply image recognition based on artificial intelligence(AI) [4]. In fact, we have been conducting research on image recognizers using artificial intelligence in our research [5]. However, AI-based product inspection systems cannot explain the reasons for wrong decisions due to the problem of the AI being a black box [6], [7]. So, this research does not apply AI for visual inspection system.

This research aims to construct a visual inspection system that can detect defects whose position is not fixed and whose shape is unknown. In general, such visual inspections detect defects based on the reflection of linear light such as fluorescent lights. On the other hand, this study targets the products made of low-reflective materials which do not reflect lights; so that, our visual inspection system is not able to use reflection of lights. To overcome this difficulty, we propose to utilize a string shadow to detect defective area. The shadow of string has the property of changing shape based on distortion of the material when projected onto low-reflective materials. Our method uses this property to detect distortion of low-reflective materials based on the shadow shape of string. In this research, our method can flexibly detect by characterizing the defective area utilizing a string shadow and determination the defective area not by pattern matching but by the distance to the approximate linear line of a string shadow. From the experimental results, we succeed in detecting defects on material.

2. Method for defective detection

We introduce the method of defective detection utilizing a string shadow in this section.

2.1. *Characterization from a string shadow*

Out method for visual inspection is to light on a string that placed horizontally to the inspection target and to form a string shadow on it, then that judges defective or non-defective from form of a string shadow. Fig. 1(a)

Keiji Kamei, Tomorou Kawahara, Yoshiyuki Daimaru

shows the shadow of a non-defective material and (b) the shadow of a defective material.

For preprocessing, the shadow image from the camera is denoised by Gaussian filter, binarized, and then edge-smoothed again by Gaussian filter by OpenCV [8] as shown in Fig. 2.
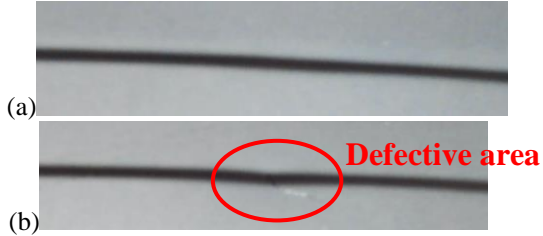
Fig.1 String shadow. (a)non-defective. (b)defective.

Fig.2 Example of after preprocessing.

2.2. *Environment for inspection*

The string shadow is affected by positions for light, camera, string and material as follows.

1. The closer the distance between light and a string, the more the light diffuses and a shadow becomes more fading.
2. The closer the distance between a string and the material, the darker the shadow of a string.
3. When the distance between light and a string is about 45 cm, light is not diffused and shadow is clearly projected.
4. The smaller the angle between the camera and the material, larger the change in the shape of the shadow of the defective area.

From these features for a string shadow, the positional relationship is chosen to have the largest change in the shape of shadow of defective area, as shown in Fig. 3.

2.3. *Linear approximation of a string shadow and judgement based on sum of distance between line and shadow*

Fig. 4 shows that area of non-defective shadow has an almost constant gradient, whereas one part of gradient changes for the defective one. When the contour on the upper side of shadow is approximated by linear line,
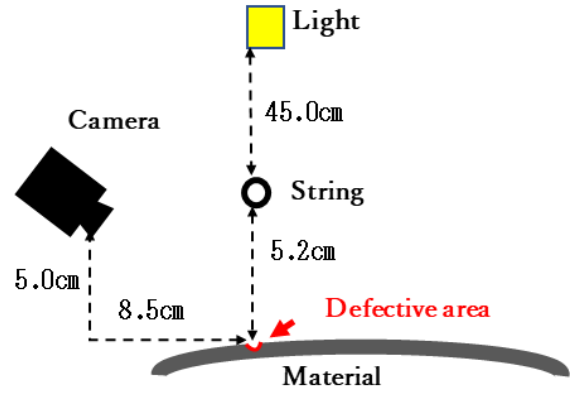
Fig.3 Positional relationship.

defect detection is based on the sum of relative distances between shadow and approximated linear line.

Fig.4 Defective shadow after preprocessing.

When width of linear approximation is Δy, the examples of linear approximated line by start point y and end point $y + \Delta y$ are shown in Fig. 5. A wide black area is created between approximated linear line and shadow contour in defective case as Fig. 5(b).
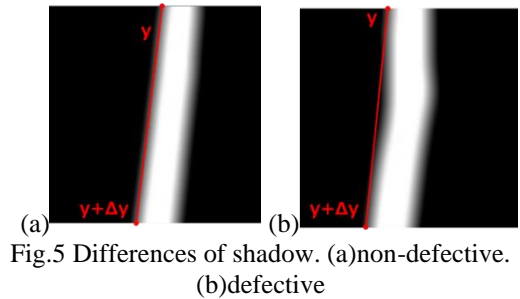
Fig.5 Differences of shadow. (a)non-defective. (b)defective

When the approximated linear line is $f(y)$ and the shadow curve is $g(y)$, gradient a of the approximated linear line is Eq. (1). Thus, linear line, $f(y)$, approximated equation is Eq. (2).

$$a = \frac{g(y_0 + \Delta y) - g(y_0)}{\Delta y} \tag{1}$$

$$f(y) = y_0 + a(y - y_0) \tag{2}$$

where, y_0 is start point of linear line. The sum of relative distances between shadow and approximated linear line is defined as,

$$S = \int_y^{y+\Delta y} |f(y) - g(y)| dy \tag{3}$$

S in Eq. (3) in defective case is larger than non-defective case from Fig. 5. S is obtained by shifting the starting point, then the sum of S is used to determine non-defective and defective. Since this method makes relative

judgments rather than absolute judgments, it is expected to have a high degree of flexibility to suit a wide variety of shapes. On the other hand, a problem with this method is the possibility of successfully approximating the defect area as a linear line depending on the value of Δy. To solve this problem, starting point y for approximated linear line is shifted several times to vary the sum of the relative distances, then determination of non-defects and defects is based on comprehensive results. This will avoid inspection errors by considering approximations which are not coincidental successes of linear line approximations.

3. Experiments

We first examine the variation of shadow contour with and without defects, and then examine the difference in the sum of relative distances between approximated linear line and a shadow contour with and without defects.

3.1. *Sample of visual inspection*

Fig. 6 shows the non-defective and defective samples used to be inspected. The defective sample has two kinds of defectives with different concavities, i.e., dent and depression.
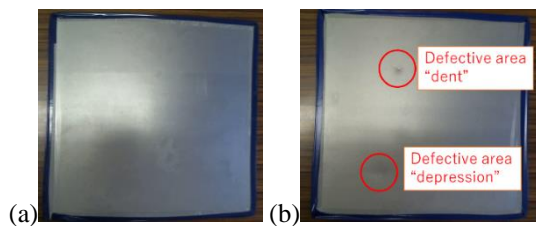
Fig.6 Samples. (a)non-defective. (b)defective

3.2. *Comparison of shadow contour*

Fig. 7 shows the variations in shadow contour coordinates with and without a dent defect. The figure shows that gradient varies linearly without dent, while gradient with dent varies significantly around y at 500 to 700, which is the defective area, in the case of dent defect.
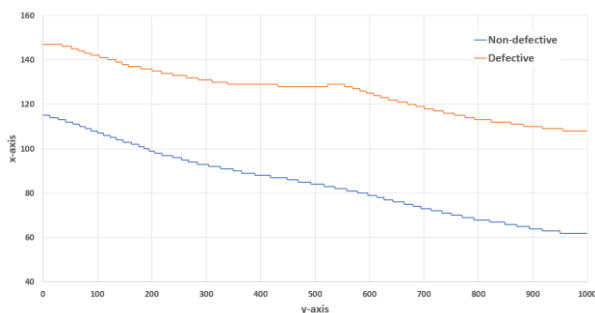
Fig.7 Comparison results of shadow contour(dent)

The same variations are also observed in case of depression. From these results, it can be said that the shadow contours are characterized by a large gradient variation in the defective area.

The product to be inspected in this study does not have a horizontal surface when viewed from the side. In addition, because the height varies depending on the inspection location, the shape of shadow and the location where it is formed are continuously altered. A method of determining by the rate of change of gradient could be considered, but it is difficult to define determination sections because gradient varies depending on the inspection location and the shape of the product and size of defect.

3.3. *Defect detection based on the sum of relative distances between a string shadow and linear line approximation*

In this experiment, two non-defective shadow images are utilized for each defect, shown as Fig. 8(a), to also verify the sum of the relative distances of non-defective shadows, shown as red and green area in Fig. 8(b).
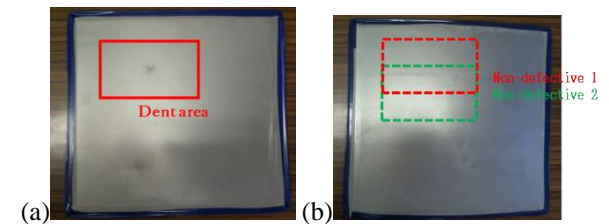
Fig.8 Example of viewpoint. (a)dent defective. (b)non-defective

The results of linear line approximation of dent defect and non-defect are shown in Fig. 9 with the starting point of approximation set to 0 and Δy set to 200. Fig. 10 shows the actual coordinates of a string shadow.
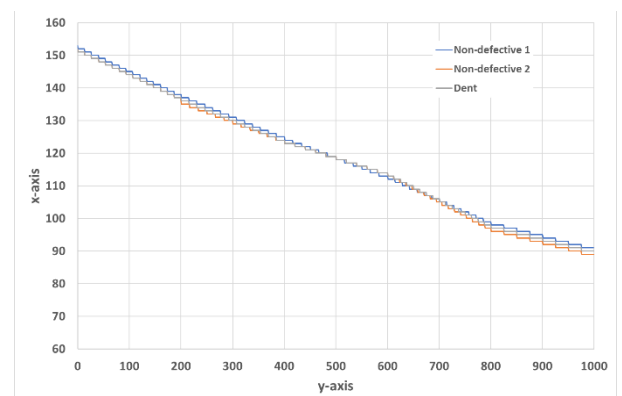
Fig.9 Liner line approx. for "Dent defect"

From figures, there is a large deviation between the approximated linear line and the actual coordinates for defect and non-defects with y-axis between 500 and 600. The same deviation is also happened in case of depression.

Keiji Kamei, Tomorou Kawahara, Yoshiyuki Daimaru

Table 1 shows the results of comparison for the sum of relative distances between a string shadow and linear line approximation in case of dent and depression defects.
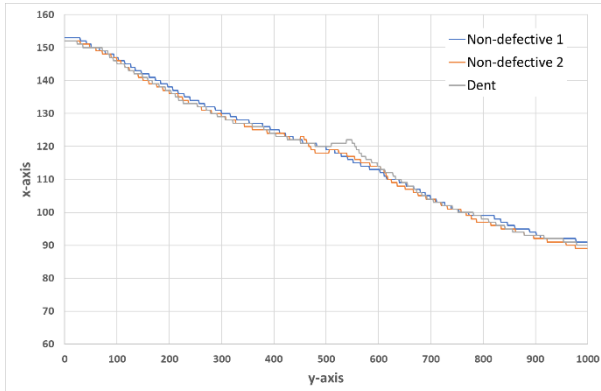
Fig.10 Before liner line approx. for "Dent defect"

Table 1. Comparison of the sum of relative dist.

	Kind of defect	Non-defect1	Non-defect2	defect
Sum of relative dist.	dent	340	344	668
	depression	458	394	576

Table 1 shows that the sum of relative distance with dent is about 300 larger than without, and about 150 larger for depression. It can be said that the sum of relative distances is smaller because of the gradual change in shadow in case of depression defect. For this, it is shown that linear line approximation successfully for slowly varying defects and might fail to detect defects. To solve this, starting point y for approximated linear line is shifted, mentioned in Section 2.3. Let d be the distance of the start point y from the image origin and shift y 19 times in total by +10 in the range of $d = 0$ to 190, then obtain the sum of relative distances for 20 times. Table 2 shows the results of shifting of starting point.

Table 2. Results of shifting. ND1, ND2 and Df correspond to "Non-defective 1", "Non-defective 2" and "defective", respectively.

	dent			depression		
	ND1	ND2	Df	ND1	ND2	Df
Avg.	448.5	459.4	582.3	406.0	399.3	565.4
Max	653	696	881	662	546	828
Min	304	271	435	234	252	341

Table 2 shows that the mean, maximum, and minimum values differ by about 100 or more between the undistorted and the distorted samples. In particular, it can be said that defects in products can be detected by using the maximum value as a criterion for determination because the maximum value differs by about 200 depending on occurrence of defects. From these results, defect detection can be performed flexibly for defect detection problems that cannot be solved by template matching.

4. Discussions

Real-time processing is required in the production environment, and the camera frame rate is 30 fps, so processing must be completed within 33[ms] before the next image is received. The frame rate at 200[kpx] is 20[ms], on the other hand, at 900[kpx] of it is 55[ms]; consequently, a frame rate drop occurred. So, it can be expected that current PC can support up to about 500[kpx], otherwise stronger PC is needed.

5. Conclusions

In this study, defect images are characterized by the shadow of a string to develop a defect detection method for low-reflective materials, and defect features are detected not by pattern matching but by the sum of relative distances to an approximated linear line. From the experiment, it is effective to shift the starting position multiple times because the sum of relative distance varies greatly depending on the starting position of the approximated linear line. It can also be said that defect detection is possible by using the maximum value as a defect determination criterion. We aim to further improve the system by conducting experiments in actual environments and trying out various kinds of samples.

Acknowledgements

References

1. Newman, Timothy S., and Anil K. Jain. "A survey of automated visual inspection," Computer vision and image understanding 61.2 (1995): 231-262.
2. Brunelli, Roberto. "Template matching techniques in computer vision: theory and practice," John Wiley & Sons, 2009.
3. Kamei, Keiji and Moriyama, Hiroyuki, "Recognition Method of Target Objects for Autonomous Tomato Harvesting Robot," Proc. 2019 Int. Conf. on Artificial Life and Robotics, (ICAROB2019), pp.557-560.
4. Huang, Szu-Hao, and Ying-Cheng Pan. "Automated visual inspection in the semiconductor industry: A survey," Computers in industry 66 (2015): 1-10.
5. Kamei, Keiji, et al. "Development of the Image-based Flight and Tree Measurement System in a Forest using a Drone," J. Robotics Netw. Artif. Life 7.2 (2020): 86-90.
6. Hachiya, Tatsuma and Kamei, Keiji. "Factor identification and prediction of power demand using sparse modeling AI," Proc. the 12th Int. Conf. on Innovative Computing, Information and Control, (ICICIC, 2023), pp. 71.
7. DARPA. "Special Issue:DARPA's Explainable Artificial Intelligence (XAI) Program," APPLIED AI LETTERS, Vol.2, Issure 4, 2021.
8. Adrian Kaehler and Gary Bradski. "Learning OpenCV 3," O'Reilly Media, Inc., 2016.

Authors Introduction

Dr. Keiji Kamei

In 2007 he obtained his PhD at the Department of Brain Science and Engineering, Kyushu Institute of Technology (Engineering). From April, 2007 to March, 2014, he was a lecturer, and from April, 2014 to March 2020, he was an associate professor. April, 2020, he has been a professor, Nishinippon Institute of Technology. His research interests are Artificial Neural Networks, Artificial Intelligence, Artificial Life and Computer Science. A member of JNNS, APNNS, IPSJ, IEICE, SOFT.

Mr. Tomorou Kawahara

He is a bachelor research member in Kamei laboratory at Nishinippon Institute of Technology. His research interest is image processing.

Mr. Yoshiyuki Daimaru

He is an employee of Nissan Motor Kyushu Co. Ltd. and a co-researcher on the corporate side of this research.

Development of IoT-Based Remote Monitoring Module for Greenhouse Environment to Facilitate Crop Growth Data Analysis

Moeko Tominga, Yasunori Takemura, Junya Era, Wataru Kaishita
Nishinippon Institute of Technology
1-11, Aratsu, Kanda-machi, Miyako-gun, Fukuoka, 800-0394, Japan

Eiji Mizoe, Tomoyasu Furukawa
Kumamoto Fruits and Vegetables Shipping Association Co. Ltd
3523, Okimachi, Yatsushiro-city, Kumamoto , 866-0013, Japan
Email: tominaga@nishitech.ac.jp

Abstract

As the shortage of agricultural population is becoming an issue, services that remotely monitor the agricultural environment using IoT are being offered in the market. However, small- and medium-scale private farmers are unable to purchase or rent IoT equipment and monitoring tools because they have to spend their budgets on maintenance and utilities in greenhouses to support environmental management, and the construction of data infrastructure has not been widely adopted. Therefore, in this experiment, we developed an agricultural IoT that can be provided to medium and small farmers, and a module that enables collection, accumulation, and visualization of time-series data acquired from sensors. The module provides an architecture that includes multiple sensing, networking, processing, and visualization applications to keep deployment and management costs low. To demonstrate its feasibility, the proposed IoT system was built, experimentally tested, and verified in a greenhouse in Yatsushiro, Kumamoto, Japan, by intermittently monitoring temperature, humidity, carbon dioxide concentration, solar radiation, and soil moisture content in the greenhouse for two years. The results obtained confirm that the proposed IoT remote monitoring module facilitates farmers to monitor the growing environment of their crops and enables them to improve their productivity by progressively increasing their technology level.

Keywords: IoT，Agricultural，Monitoring System, visualization

1. Introduction

Smart Agri and Agri Tech are attracting attention as new forms of agriculture that utilize advanced technologies such as robotics, ICT to enable labor-saving, high-quality production [1]. One of the technologies that support this Smart Agri is the spread of the Internet of Things (IoT), whereby various "things" are now connected to the internet IoT refers to a system whereby sensors are built into various "things", which are then connected to the internet, and a large amount of data acquired via the internet It is a mechanism whereby large amounts of data acquired via the internet are stored in the cloud. This accumulated large amount of data is known as 'big data' (BD), and although BD cannot create value simply by accumulating it, new value can be created by analyzing the accumulated data [2].

In agriculture, where declining productivity due to an ageing population and a lack of bearers is a problem, there are high hopes for this IoT as a way to improve productivity with fewer people. It is desired to acquire environmental information on farms, which is a factor that changes the quality and quantity of agricultural products, and to visualize this information. In addition, one of the challenges for future development is to analyze the acquired data and weather data and utilize them as time-series data for forecasting, thereby improving productivity and ensuring the production of high-quality and stable agricultural products [3]. Here, an example of a farmer growing banana and coffee in Yatsushiro, Kumamoto Prefecture, Japan, is presented. Banana and coffee are mainly grown in tropical regions. Japan belongs to the temperate zone, and farmers across the country are searching for new cultivation methods to enable the cultivation of tropical plants such as banana and coffee in temperate Japan. Farmers in Yatsushiro City, the site of this experiment, also grow bananas and coffee and maintain a tropical environment in vast plastic greenhouses. However, in Japan, which has four seasons, temperatures vary throughout the year and it is difficult to maintain a constant environment inside the greenhouses.

Moeko Tominaga, Yasunori Takemura, Junya Era, Wataru Kaishita, Eiji, Mizoe, Tomoyasu Furukawa

However, in Japan, which has four seasons, the temperature varies throughout the year and it is difficult to maintain a constant environment in the vast greenhouses, which affects the growth and production of the plants. For this reason, many companies are introducing agricultural IoT, collecting environmental data in greenhouses and providing visualization services. However, small- and medium-scale individual farmers have to spend their budgets on the maintenance of their greenhouses, utility costs and other costs to support environmental management, and are unable to purchase or rent IoT equipment and monitoring tools, which means that the construction of data infrastructures has not become widespread. Therefore, in this experiment, we developed an agricultural IoT device that can be provided to medium and small farmers, and built a system for collecting, storing and visualizing time-series data acquired from sensors.

2. Development concept

The concept was set up as follows.

- To collect on temperature, humidity, carbon, dioxide concentration, moisture content and solar radiation in order to investigate environmental information in greenhouse from all perspectives.
- Designed to continue to operate even under severe environmental conditions and to ensure that data acquisition is not interrupted.
- Programmed to be able to be used universally at other sites, and designed so that sensors can be connected using connectors.
- Collect agricultural IoT devices that meet the above conditions using inexpensive and highly accurate sensors.

A system of IoT modules configured according to the concept is shown in Fig. 1. All sensor values are regularly collected by microcomputer and stored in storage in the cloud. The user can view the values of these sensors numerically and graphically at any time and from anywhere.
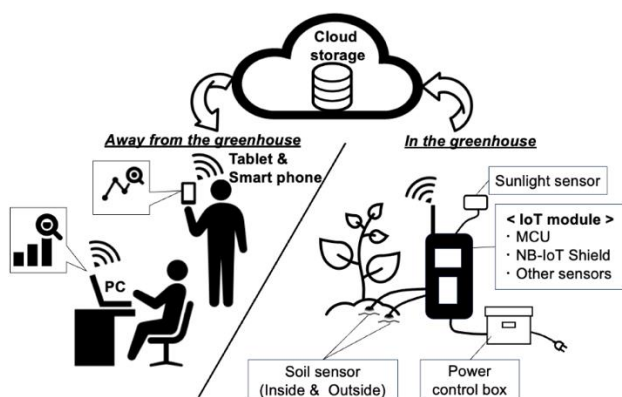
Fig. 1 System image of the IoT module

The developed IoT device is shown in Fig. 2 and the specifications are given in Table 1. As an Micro Controller Unit (MCU), the Arduino, which is also used in other smart farming systems [4], [5], was used. The developed IoT device was equipped with several types of sensors to collect data in order to investigate the environmental information inside the greenhouse from all perspectives.

In order to develop an IoT device that can continue to operate even in a harsh environment, a temperature and humidity sensor was also attached to the inside of the greenhouse, and a durability test was conducted to investigate at what temperature range the device would break down. The results showed that failures occurred at around 50 °C, so a fan was installed inside the chamber to prevent failures due to overheating of the IoT device. In addition, a Watch Dog Timer (WDT) was incorporated into the software to monitor whether the MCU program stopped or communication was interrupted, and if so, measures were taken to restart the program. By taking countermeasures in terms of hardware and software, we were able to develop an IoT device that could continue to operate even in a harsh environment.
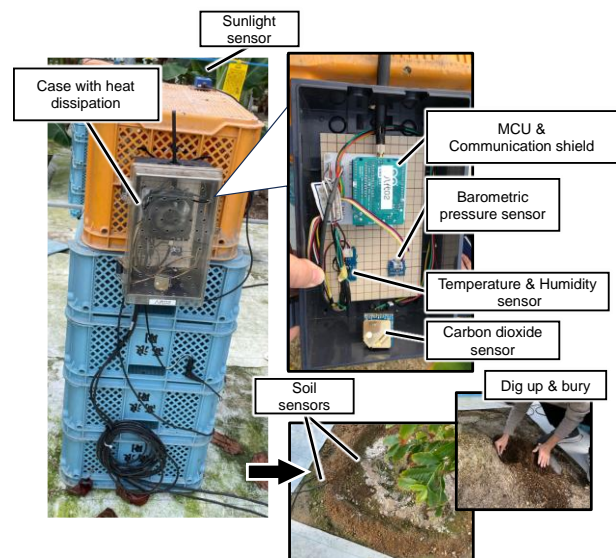
Fig. 2 IoT devices developed (being implemented in a green house in Yashiro).

Table 1 Specification of agriculture IoT module

Size (Length x width x depth)	$280.5 \times 170 \times 130$
MCU	Arduino Uno
Supply voltage	$12[V]$
Line	Long Term Evolution (LTE)
Temperature sensor	BME280
Humidity Sensor	
Barometric pressure sensor	
Carbon dioxide sensor	S-300-3V
Soil moisture censer	WD-3-W-5Y
Sunlight sensor	PVSS-03

Fig. 4 Visualization using SORACOM Lagoon
(Only some data on coffee was extracted)

3. System configuration diagram

The system configuration is shown in Fig. 3. The data acquired from each sensor connected to the MCU was A/D converted and the time-series data was stored in the data server via the LET module. The data server used SORACOM Harvest provided by SORACOM, and SORACOM Lagoon (Fig. 4) was used as the visualization tool.
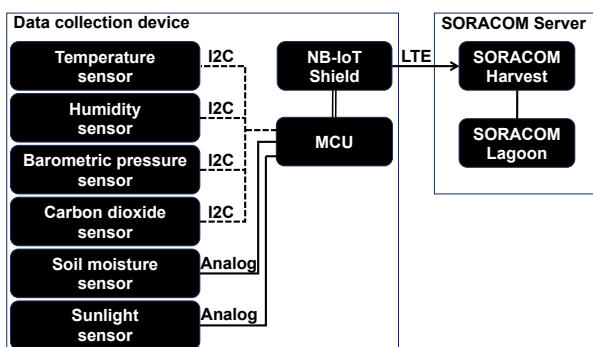
Fig. 3 System configuration diagram

4. Field experiments and time series data obtained

The developed agricultural IoT equipment was actually installed in a plastic greenhouse in Yatsushiro to monitor the growing environment of banana and coffee as shown in Fig. 5. The monitoring period started intermittently in 2021 and lasted for two years until 2023. The data interval was 15 minutes. An example of the data collected is shown in Fig. 6. This is temperature data, where temperatures inside and outside the developed agricultural IoT module were collected for one month. As can be seen from Fig. 6, sensor values were successfully retrieved and stored in cloud storage at 15-minute intervals. The sensor values also show that the internal temperature tends to be 1-5 ℃ higher than the temperature outside the IoT module, but is kept almost the same due to heat dissipation.

However, as a problem, the phenomenon of system stoppages could appear. When the system is restarted, it returns to a state where data can be continuously retrieved again. The data in which this phenomenon appears is shown in Fig. 7. The period during which the sensor values are not updated is the period during which the system is stopped. This has not yet been resolved. As a countermeasure, the system was monitored and improved so that it can be restarted remotely in the event of a stoppage. After the improvements, the system stoppages have not been reproduced, but the program is being logged so that the cause of the reproduction can be clarified.

Fig. 5 Installation of the developed IoT modules
(Right: coffee field, Left: banana field,
Top: before growth, Bottom: after growth)

Moeko Tominaga, Yasunori Takemura, Junya Era, Wataru Kaishita, Eiji, Mizoe, Tomoyasu Furukawa
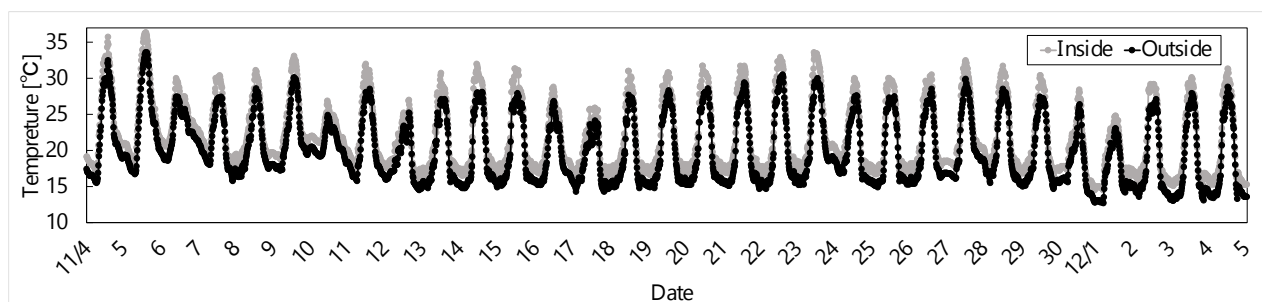
Fig. 6   Temperatures observed by IoT devices(4 November 2023 - 4 December 2023)

Periods when data is not updated to the cloud.
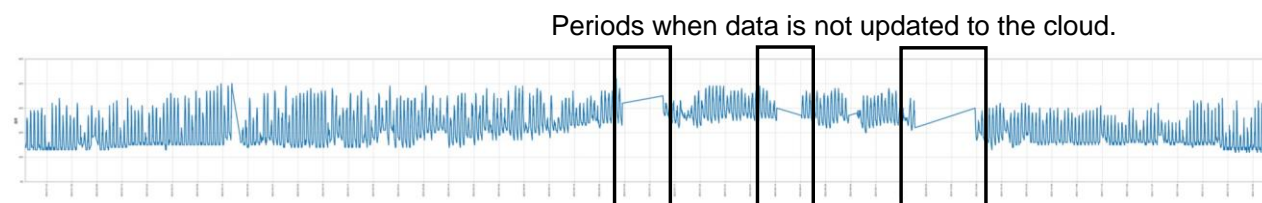
Fig. 7   Temperatures observed by IoT devices(17 January 2022 – 28 December 2023)

5. Conclusion

In this study, an IoT device for agriculture was developed and actual field observation experiments were conducted in greenhouses. As a result, a device with the following functions was developed.

- To continue to operate normally in greenhouses where the humidity is above 90% and the temperature exceeds 30°C.
- Data is uploaded to the cloud every 15 minutes and can be immediately visualized.
- The device can observe multiple parameters (temperature, humidity, barometric pressure, Sunlight, soil moisture and carbon dioxide concentration in the air).
- Robust against communication errors.

In the future, the durability of the system will be increased by improving the modules so that the system can run without stoppages for a long period of time.

Acknowledgements

References

1. J. Pretty, "Agricultural sustainability: concepts, principles and evidence", Philosophical Transactions of the Royal Society B: Biological Sciences, Vol. 363(1491), 2007, pp. 447-465.
2. O. Elijah, T. A. Rahman, I. Orikumhi, C. Y. Leow and M. N. Hindia, "An Overview of Internet of Things (IoT) and Data Analytics in Agriculture: Benefits and Challenges," in IEEE Internet of Things Journal, vol. 5, no. 5, 2018, pp. 3758-3773.
3. M. R. Bendre, R. C. Thool and V. R. Thool, "Big data in precision agriculture: Weather forecasting for future farming," 2015 1st International Conference on Next Generation Computing Technologies (NGCT), Dehradun, India, 2015, pp. 744-750.
4. V. K. Shukla, R. S. Nair, F. Khan,"Chapter 13 - Smart irrigation-based behavioral study of Moringa plant for growth monitoring in subtropical desert climatic condition", In Intelligent Data-Centric Systems, AI, Edge and IoT-based Smart Agriculture, Academic Press, 2022, pp. 227-240.
5. E. Ogunti, "IoT Based Crop Field Monitoring and Irrigation Automation System", International Journal of Innovative Science, Engineering & Technology, Vol. 6 Issue 3, 2019, pp. 124-129.

Authors Introduction

Dr. Moeko Tominaga
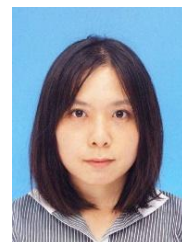
She is an Assistant Professor at the Department of Integrated System Engineering, Nishinippon Institute of Technology, Japan. Her research interests are in multi-agent systems and field robotics.

Dr. Yasunori Takemura

He received his M.S., and Ph.D. degrees from Kyushu Institute of Technology, Japan, in 2006 and 2009, respectively. In 2010, he joined Nippon Bunri University. In 2015 he joined Nishinippon Institute of Technology, where he is currently an Professor of the Department of Engineering. His research interests include robotics, IoT and AI Technology.

Mr. Junya Era

He worked as a steel plant controls engineer from 2012-2019 and has been at Nishinippon Institute of Technology since 2020, learning about the IoT. He is an undergraduate engineering student. His research interests are in IoT Technology.

Mr. Wataru Kaishita

He received his Bachelor's degree from the Nishinnippon Institute of Technology in 2022 and started his postgraduate studies at the Nishinippon Institute of Technology in 2022. He is currently in his first year of master's studies. His research interests are in the field robotics.

Mr. Eiji Mizoe

He is currenntry working in Kumamoto Fruits and Vegitables shipping Association Co. Ltd.. (Kumamoto seikabutsu syukka kumiai).

Mr. Hiroaki Wagatsuma

He is currenntry CEO in Kumamoto Fruits and Vegitables shipping Association Co. Ltd.. (Kumamoto seikabutsu syukka kumiai).

Development greenhouse environment prediction system using IoT data

Yasunori Takemura, Naoya Nishida, Moeko Tominga

Nishinippon Institute of Technology, 1-11 Aratsu, Kanda-town, Miyako-gun, Fukuoka, 800-0394, Japan

Eiji Mizoe, Tomoyasu Furukawa

Kumamoto Fruits and Vegetables Shipping Association Co. Ltd, 3523, Okimachi, Yatsushiro-city, Kumamoto , 866-0013, Japan
Email: takemura@nishitech.ac.jp

Abstract

Currently, the decrease in the working population is a significant problem in agriculture in Japan. In addition, the prices of imported vegetables and fruits have been rising due to international trade problems such as the weak yen. This study has developed an environmental data storage system using IoT to assist the labor force and to understand the cultivation data of crops that have not been cultivated in Japan. Using this big data, we have constructed and evaluated a forecasting system. In this study, we focus on the domestic cultivation of bananas grown in the tropics as the subject to obtain data and make predictions. For learning, a system was constructed to infer the next day's temperature by time-series analysis, using as input data obtained from the IoT. In this paper, we report a comparison of two time series analysis algorithms.

Keywords: Time-series data analysis, ARIMA model, prophet model, IoT, agriculture

1. Introduction

For In recent years, the Internet of Things (IoT) has become increasingly popular, and various " devices" are now connected to the Internet. The IoT is a system in which sensors are embedded in various " products," connected to the Internet, and a large amount of data acquired via the Internet is stored in the cloud. This accumulated large amount of data is called "Big Data (BD)," and today there are tens of billions of IoT devices in the world, all of which continue to generate new data. BD cannot create value simply by accumulating it, but it can create new value by analyzing the accumulated data. In addition, with the development of computer technology, the computation time of AI (Artificial Intelligence) has been shortened, and many companies have developed home appliances and automobiles that incorporate AI, making the IoT and AI a familiar presence. As evidence of this, the number of IoT devices is on an increasing trend with respect to transition data, and is expected to exceed 32.33 billion units by 2022 [1] . In Japan, the government and companies are actively working toward the realization of Society 5.0, and it is expected that a society that uses information in a wide variety of ways will expand in the future [2].

For example, the utilization of IoT databases is expected in various fields. One example is in the field of agriculture, where the number of farmers is decreasing year by year due to the problem of declining productivity caused by the aging of the farming population and the lack of farmers to take on the workforce. Therefore, as a solution, it is expected to improve productivity with a small number of people, and a necessary method to acquire environmental information on farms, which is a factor that changes the quality and quantity of agricultural products [3]. As a solution to this problem, visualization and prediction functions of environmental information using IoT are expected. One of the challenges in future development is to analyze the acquired IoT data and weather data and utilize them as time-series data in forecasting algorithms in order to improve productivity and realize high-quality and stable crop production [4].

The purpose of this research is to build an AI using time-series data with IoT technology, and to generate and validate prediction data using time-series data. The experiment involves time-series prediction of environmental data in a plastic greenhouse in Yatsushiro City, Kumamoto Prefecture, where tropical crops are harvested. The time-series forecasting algorithms were compared using the ARIMA model [5] and Prophet [6], because of their ability to consider the effects of periodicity and external factors.

2. Methodology
2.1. ARIMA (Auto Regressive Integrated Moving Average) model

Yasunori Takemura, Naoya Nishida, Moeko Tominaga, Eiji Mizoe, Tomoyasu Furukawa

The ARIMA model [5], derived by Box & Jenkins, is a model consisting of three components: AR (Auto Regressive) model, MA (Moving Average) model, and summed components (integrated). Time series data may have non-stationary time series processes when a trend analysis is performed, which cannot be handled by the ARMA model. Therefore, non-stationary time series must be converted to stationary time series and then fit into the ARMA model. The ARIMA model, which adds a summation component to the ARMA model, is then applied to the non-stationary time series and aims to make them stationary by taking the "d" th-order difference. The general equation is shown in Equation (1) below.

$$y_t - y_{t-d} = c + e_t + \sum_{i=1}^{p} \emptyset_i y_{t-i+} + \sum_{j=1}^{q} \theta_j e_{t-j} \qquad (1)$$

In the equation, t denotes time, d, i and j denote the time at the regression point in time. e indicates the error height, y is the output, and c, ϕ and θ are constants. In recent years, it has been used as a forecasting system for COVID-19 [7] , and has also been used in research such as stock price forecasting [8].

2.2. Prophet model

Prophet is a library for time series analysis developed by Meta in 2017. The advantages of this model are that it can be built without statistical knowledge, it can be trained even with missing values, and it is easy to interpret the prediction results. Prophet is based on the Generalized Additive Model (GAM) [9], which is a model that uses a It predicts future values by modeling and combining four terms: the trend term (growth: $g_{(t)}$), the seasonality term (seasonality: $s_{(t)}$), the holidays effect term (holidays: $h_{(t)}$), and the error term $\varepsilon_{(t)}$). The model equation is shown in Equation (2) below.

$$y = g_{(t)} + s_{(t)} + h_{(t)} + \varepsilon_{(t)} \qquad (2)$$

3. Evaluation Methodology
3.1. Correlation coefficient

The correlation coefficient expresses the strength of the relationship and correlation between two items as a positive or negative value, and can be objectively analyzed in the range of 1 to -1. 1 is closer to a positive correlation and -1 to a negative correlation. Equation (3) below shows the formula for obtaining the correlation coefficient. In this research, we aim to use data obtained from IoT data and weather data as explanatory variables in the algorithm in order to increase the accuracy of data prediction, and to select data by examining the correlation coefficient with the output values.

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \qquad (3)$$

3.2. Akaike's Information Criterion (AIC) [10]

It is a statistic that evaluates the predictability of a statistical model using the residuals between observed and theoretical values, with a smaller value indicating a better fit. If the log likelihood of the model is L and the number of parameters in the model is p, the AIC is calculated by the following Equation (4). In this paper, the AIC is used as an evaluation criterion when selecting parameters for the ARIMA model. L is the maximum likelihood and p is the number of free parameters. In the present model, we adopted a low value of AIC to avoid a complex model in determining the AR and MA coefficients in the ARIMA model.

$$AIC = -2\ln(L) + 2p \qquad (4)$$

3.3. Coefficient of determination R^2

The coefficient of determination, also known as the contribution ratio, is a measure of the explanatory power of the predicted value of the target variable relative to the observed value of the target variable in a regression analysis. The coefficient of determination is generally expressed as R^2. R^2 takes values between 0 and 1, and the closer to 1, the more valid the regression equation is. When the actual data are (x_i, y_i), the data estimated from the regression equation are $(y_i, \hat{y_i})$, and the mean value obtained from the entire data is $(\bar{x}, \bar{y})$, the following Equation (5) is shown.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y_i})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \qquad (5)$$

4. Data set and Experiment Environment
4.1. Experiment Environment

The field for this experiment was a farmer in Yatsushiro City, Kumamoto Prefecture, who grows bananas and coffee. Since Japan belongs to the temperate zone, farmers across the country are searching for new cultivation methods to make the cultivation of tropical plants such as bananas and coffee possible in temperate Japan. The farmers in Yatsushiro City, which is the focus of this experiment, also grow bananas and coffee in a vast plastic greenhouse, maintaining a tropical environment. However, in Japan, which has four distinct seasons, temperatures vary throughout the year, making it difficult to maintain a constant environment in the vast greenhouses, which affects the growth and production of the plants. For this reason, many companies are introducing agricultural IoT, collecting environmental

data in greenhouses and providing visualization services. However, small- and medium-scale private farmers have to spend their budgets on environmental management, such as maintenance and utility costs in greenhouses, and are unable to purchase or rent IoT equipment and monitoring tools, which has hindered the diffusion of data infrastructure construction. Therefore, in this experiment, we developed an agricultural IoT system that can be provided to small and medium farmers, and collected, stored, visualized, and analyzed time-series data acquired from sensors, with the goal of predicting environmental information for the next day based on time-series data.

4.2. *IoT Device and Data set*

Figure 1 shows the IoT device developed. The developed IoT devices were equipped with multiple types of sensors to collect data in order to investigate environmental information in greenhouses from various perspectives. In addition, we conducted a durability test to determine the temperature range at which the device malfunctions, aiming to develop an IoT device that can continue to operate even under harsh environmental conditions. The results showed that the IoT devices failed at temperatures around 50 degrees Celsius, so a fan was installed inside the chamber to prevent failures due to overheating of the IoT devices. In addition, a Watch Dog Timer (WDT) was incorporated into the software to monitor the microcontroller for program stoppages and communication interruptions, and to restart the program in the event of a stoppage. We developed IoT devices that can continue to operate even in harsh environments by implementing countermeasures on both the hardware and software sides.

From this device, temperature, humidity, air pressure, carbon dioxide, solar radiation, and soil moisture in the greenhouse can be obtained. In addition, we also obtained meteorological data supplied by the Japan Meteorological Agency as necessary data for forecasting, such as altar instantaneous wind speed, precipitation, sunshine duration, maximum wind speed, and average wind speed, and examined whether these data could be utilized as explanatory variables for the model.

a.   Device   b.   Setting Environment
Fig. 1. IoT Device

Table 1. The planning and control components.

Item	Correlation Coefficent
Temperature in plastic greenhouses 1 day ago	0.886252
Average temperature 1 day before [JMA].	0.578104
Lowest temperature 1 day ago [JMA].	0.567764
Highest temperature 1 day ago [JMA].	0.546740
Solar radiation 1 day ago	0.438165
Moisture content 1 day ago (outside)	0.365323
Maximum instantaneous wind speed 1 day ago [JMA].	0.109610
Precipitation 1 day ago [JMA].	0.086130
Sunshine hours 1 day ago [JMA].	0.077532
Moisture content 1 day ago (inside)	0.015999
Maximum wind speed 1 day ago [JMA].	-0.195937
Average wind speed 1 day ago [JMA].	-0.198114
Air pressure 1 day ago	-0.284419
Humidity 1 day before	-0.467905
Carbon dioxide concentration 1 day ago	-0.680186

The correlation coefficients between the output target, temperature in the plastic greenhouse, and the other data sets are shown in Table 1, and the explanatory variables needed to predict the temperature in the greenhouse were extracted from the threshold values. As a result, five items were extracted as positively correlated: the temperature in the greenhouse one day before, the average temperature one day before, the minimum temperature one day before, the maximum temperature one day before, and the amount of solar radiation one day before. Two items, carbon dioxide concentration one day before and humidity one day before, were selected as negatively correlated items, and the temperature in the greenhouse one day before had a high correlation coefficient with the objective variable to be predicted. Therefore, the high correlations were adopted as explanatory variables in the model. As a pre-processing step for time series forecasting, the data set is divided into two types of data sets: training data and validation data. In this experiment, in order to predict the temperature in the greenhouse on the next day, the validation data was set as the last three days, and the other data was used as the training data. The other data were used as training data. Data for training The training data are For the training data, we used 14464 data sets from January 18 to December 25, 2022. We used the temperature inside the greenhouses as input and the temperature, carbon dioxide level, average temperature, humidity, and solar radiation of the previous day with high correlation coefficient as

Yasunori Takemura, Naoya Nishida, Moeko Tominaga, Eiji Mizoe, Tomoyasu Furukawa

explanatory variables. In addition, 145 pieces of data from December 26 to 28 were utilized as test data for accuracy verification.

5. Experiment Result

Analysis of the ARIMA model chorreogram shows that the data has a period of one year, and the candidate parameters for AR(p) are in the range of 0 to 4, and the candidate parameters for MA(q) should be designed in the range of 0 to 2. Figure 2 and 3 show the prediction results when good parameters are set based on these results. Figure 2 shows the forecasting results without explanatory variables. The black dots indicate the values that are the correct answers, the orange line shows the predictions of the ARIMA model, and the blue line shows the Prophet predictions. The value of R^2 is 0.98 for the ARIMA model, and 0.83 for Prophet's model. The values of R2 for the test data are 0.32 for the ARIMA model and 0.74 for the Prophet model. This indicates that the Prophet model gives better predictions when there are no explanatory variables. Next, Figure 3 shows the results of the model with additional explanatory variables. The black dots indicate the values that are the correct answers, the orange line shows the predictions of the ARIMA model, and the blue line shows the Prophet predictions. The ARIMA model, the value of R^2 in the training data is 0.80, and for the Prophet model, the value is 0.85. The predicted values for the test data were 0.88 for the ARIMA model and 0.85 for the Prophet model. These results indicate that the accuracy improves with the addition of explanatory variables, and that the ARIMA model is more accurate when explanatory variables are added.
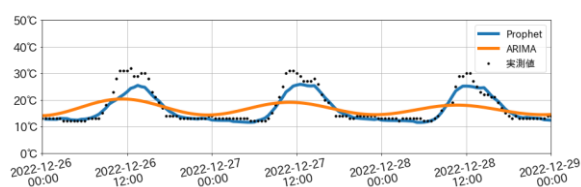
Fig. 2 The experiment data using ARIMA and Prophet model (No explanatory variables)
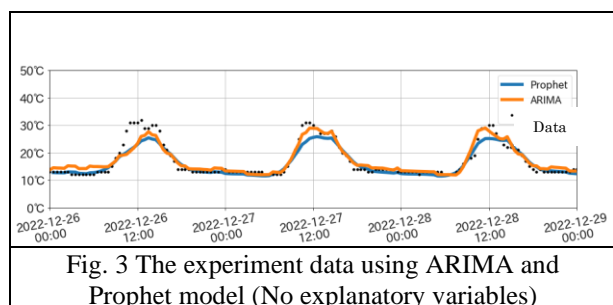
Fig. 3 The experiment data using ARIMA and Prophet model (No explanatory variables)

6. Conclusion

Then paper, we developed a model that acquires data in agriculture using IoT devices and utilizes the acquired data to make predictions. The ARIMA model and the Prophet model were used for development and comparison. In the experiment, we examined the correlation between the output temperature and the data acquired to use the explanatory variables, and compared the models with and without the variables with high correlation values added as explanatory variables. The results showed that the models with explanatory variables were more accurate and the ARIMA model had the best performance. Future issues include repeated experiments to improve accuracy and evaluation through long-term operation of the forecasting system.

Acknowledgements

References

1. Ministry of Internal Affairs and Communications, 2022 White Paper on Information and Communications, Part 2, Chapter., ICT Market Trends, Section 9, Trends in the Postal and Letter Service Businesses., p.111, 2021, in Japanese
2. Ministry of Economy, Trade and Industry of Japan, "Future Investment Strategy 2017 -Reforms towards the realization of Society 5.0-", pp. 39-43, 2017, in Japanese
3. Prime Minister's Office: The Fifth Science and Technology Basic Plan, Chapter 3: Measures to Address Economic and Social Challenges, pp. 16-18, 2016, in Japanese
4. Shingo Otsuka, "Acquisition of Environment Data at Farmers" Vol.30, No.1, p.35-41, 2018, in Japanese
5. George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel, Greta M. Ljung : Time Series Analysis: Forecasting and Control, 5th Edition, 2015.
6. Sean J. Taylor, Benjamin Letham : Forecasting at Scale, 2017.
7. Domenico, Benvenuto, at el. al, "Application of the ARIMA model on the COVID-2019 epidemic dataset, Data in brief, Vol29, pp.1-4, 2020
8. A. A. Ariyo, A. O. Adewumi and C. K. Ayo, "Stock Price Prediction Using the ARIMA Model," 2014 UKSim-AMSS 16th International Conference on Computer Modelling and Simulation, Cambridge, UK, 2014, pp. 106-112, doi: 10.1109/UKSim.2014.67.
9. Trevor Hastie and Robert Tibshirani : Generalized Additive Models, Statistical Science, vol.1, No.3, pp.297-318, 1986
10. Akaike, H., "Information theory and an extension of the maximum likelihood principle", Proceedings of the 2nd International Symposium on Information Theory, Petrov, B. N., and Caski, F. (eds.), Akadimiai Kiado, Budapest: 267-281 (1973)

Authors Introduction

Dr. Yasunori Takemura

He received his M.S., and Ph.D. degrees from Kyushu Institute of Technology, Japan, in 2006 and 2009, respectively. In 2010, he joined Nippon Bunri University. In 2015 he joined Nishinippon Institute of Technology, where he is currently an Professor of the Department of Engineering. His research interests include robotics, IoT and AI Technology.

Mr. Naoya Nishida

He received his Bachelor's degree and M.S. in Engineering in 2021 and 2023 from the Faculty of Engineering, Nishinippon Institute of technology in Japan. He is currently working in Information Technology company.

Dr. Moeko Tominaga

She is an Assistant Professor at the Department of Integrated System Engineering, Nishinippon Institute of Technology, Japan. Her research interests are in multi-agent systems and field robotics.

Mr. Eiji Mizoe

He is currenntry working in Kumamoto Fruits and Vegitables shipping Association Co. Ltd.. (Kumamoto seikabutsu syukka kumiai).

Mr. Hiroaki Wagatsuma

He is currenntry CEO in Kumamoto Fruits and Vegitables shipping Association Co. Ltd.. (Kumamoto seikabutsu syukka kumiai).

The Development of SaaS for Quantifying the Amount of Drifted Debris on the Coast

Ayumu Tominaga

National Institute of Technology Kitakyushu College, 5-20-1 Shii, Kokuraminami-ku, Kitakyushu, Fukuoka, 802-0985, Japan

Ryohei Komori

National Institute of Technology Kitakyushu College, 5-20-1 Shii, Kokuraminami-ku, Kitakyushu, Fukuoka, 802-0985, Japan

Eiji Hayashi

Kyushu Institute of Technology, 680-4, Kawazu, Iizuka, 820-8502, Fukuoka, Japan

Email: tominaga@kct.ac.jp, k18082rk@apps.kct.ac.jp, haya@ics.kyutech.ac.jp

Abstract

In the process of collecting drifted debris along a coastal area, it has been required to quantify the collection results. This study introduced a Software as a Service (SaaS) equipped with image processing capabilities for the detection and identification of debris in photographs. The SaaS was developed to facilitate the automated weighing of the collected debris. In the beach cleanup event in June 2023, 317 images capturing debris found during the coastal cleanup were gathered on a cloud server from 90 participants via this SaaS. These images were subsequently analyzed by object recognition AI executed on a cloud server, leading to the detection and identification of 954 instances of debris. This result represents the actual amount of debris collected during this cleanup event, and indicates that the achievements of the event were quantified via the SaaS.

Keywords: Coastal environment, Cloud edge, SaaS, AWS

1. Introduction

The impact of marine debris is a global issue [1]. The familiar scenery of the beach can be destroyed by a large amount of debris drifting to the coast after being carried from urban areas through rivers. In Japan, an island nation surrounded by the sea, this is a significant problem that needs to be addressed as soon as possible. In recent years, much research has been conducted on quantifying the amount of beach-drifted debris. Especially, methods based on high-resolution images acquired from UAVs or satellites have been developed [2], [3], [4]. In both proposed methods, deep learning has been actively utilized, and certain results have been achieved for a local region of the beach. These methods can quantify the physical situation of the debris spread on the beach compared to human-powered methods. However, these image processing-based methods have not been implemented in an actual field. It is estimated that people engaged in cleanup tasks may not consider utilizing such technologies to improve the effectiveness of cleanup.

Additionally, these latest technologies have not been widely disseminated to the public.

In this study, considering this situation, it is estimated that a system that can be used with already familiar devices is needed for non-engineers or general consumers participating in the quantifying work of debris. In other words, a method that can be completed with only a smartphone and its operation method is important to promote a quantitative understanding of waste. The performance of smartphones equipped with high-resolution cameras has been increasing year by year. Furthermore, the advancement of AI through cloud computing, which is a core technology in Digital Transformation (DX), has created a social situation where high-performance PCs are not needed for individuals. In this study, while utilizing image processing-based litter detection and identification methods similar to previous studies, an attempt has been made to widely establish the work of quantifying debris by developing a Software as a Service (SaaS) to provide users with functions related to image analysis and information collection on the cloud. Users are required to take a photo containing the debris found on the beach during the beach cleanup, and the identification process of the debris category and the quantification process for

Ayumu Tominaga, Ryohei Komori, Eiji Hayashi

each category are completed on the cloud. The SaaS implemented with such functions has been developed. In this paper, to socially implement the proposed method, the verification results are reported.

2. Proposed method

2.1. *Beach debris quantifying though cloud service*

In this study, an application integrating AI and cloud services has been developed to facilitate the quantitative assessment of beach-drifted debris. Fig. 1 represents a use case of the application. Users of this application are participants in a beach cleanup event. Typically, they explore the beach, find debris, and pick it up. In the proposed method, users are expected to take a photo capturing the found debris. After finishing the cleanup work (or during the cleanup), these photos will be uploaded to cloud storage via the internet. Subsequently, the photos will be analyzed by AI upon upload. Finally, the information needed to quantify the debris, such as the type of debris and the quantity of each type, will be gathered as a database on the cloud. As for the expectation of this proposed method, the event that originally gathered labor to collect debris would transform into an initiative constructing a large-scale database focusing on the physical situation of the drifted debris. This database will be provided to the participants or organizers of the event, and it can be used for evaluating the results of the executed event or planning the schedule for the next event.
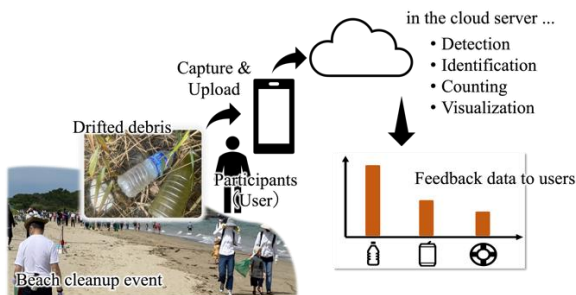
Fig. 1 Usecase of the application in the proposed method

2.2. *Design of the application*

To achieve the usecase as shown in Fig. 1, the application has been designed as follows:

- Users can upload images without depending on a specific location.
- The images will be uploaded to cloud storage and automatically analyzed by AI upon upload.
- AI detects all debris in the image, classifies them, and quantifies the number of each debris type. These

results are then gathered in cloud storage for easy access.
- Users can review the data analyzed by AI through the application.

In this study, a web application with a graphical user interface (GUI) has been selected and developed to meet the previously specified requirements. Cloud storage and CPU resources on the cloud have been utilized to automate image processing. By leveraging cloud computing, a large amount of information about the debris can be collected through smartphones owned by individuals.
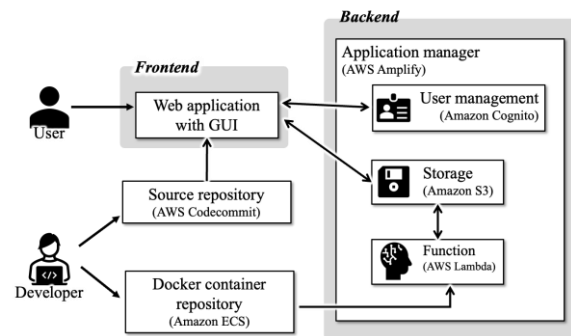
Fig. 2 The system structure and network resources

3. Developed web application

3.1. *Amazon Web Service (AWS)*

In developing the web application, Amazon Web Services (AWS) was chosen as the platform for constructing its network infrastructure. AWS, the cloud computing service provided by Amazon Web Services Inc., has been offering services, including network infrastructure. By integrating various services within AWS, the web application developer can rapidly create an application with extensibility and robustness.

3.2. *The system structure*

Fig. 2 shows the structure of the system (as the web application). As mentioned above, resources such as storage for data or image processing have been implemented by integrating services provided by AWS (e.g., AWS Amplify, S3, AWS Lambda, etc.). The system can be divided into two subsystems: a frontend and a backend. The frontend is software that can be executed on a web browser, essentially a web page including a GUI. The resources required to implement the functions of the application are referred to as the backend, and they exist on the cloud.

In this study, the frontend was designed as a web page with a file browser UI, allowing users to choose images and buttons to initiate the upload of the selected image. On the backend side, the implementation included (1) User identification and management functions, (2) The

storage region on the cloud, and (3) Virtual CPU executing inferencing process. Particularly, the details of the inference processing are explained in the next section in this paper.

Fig. 3 Captured debris by user

Fig. 4 Example of detection and identification debris by the SaaS: Right side was identified "Styrofoam". Left side was identified "Other plastics".

3.3. *Inference process based on DNN*

In this study, a DNN-based inference is used to detect captured debris in an uploaded image and classify its category. In this process, *mmdetection* [5] was used as the DNN framework, and *Mask-RCNN* [6] was employed as the neural network architecture for detection of debris. And, *TACO dataset* [7], it is large dataset for garbage was utilized as the annotated data. In this study, the dataset was classified by eleven categories as below:

- Plastic bottle
- Can
- Glass
- Metal
- Tabaco
- Rubber
- Rope & String
- Styrofoam
- Paper
- Other plastics
- Others

Fig. 3 shows an example of input image including the actual beach-drifted debris, and Fig. 4 shows example of inference result of Fig. 3 image.

4. Experiment in actual environment

The developed system has been implemented in an actual beach cleanup event and verified. The event, held in June 2023 at Hokuto-mizukumi Park (Munakata, Fukuoka, Japan), aimed to have participants use the system. In this cleanup activity, approximately 70 individuals from 30 families participated, spending one hour cleaning up the beach. It is estimated that 30 to 40 smartphones were used to utilize the system. The participants were instructed to upload images capturing the found debris after finished the cleanup activity. The total number of gathered images was 317, and AI's identification processing detected 954 beach-drifted debris. Fig. 5 represents the numbers for each category of debris, with the majority falling under "Other plastics," including plastic fragments and buoys used in fisheries. As shown in Fig. 5, plastics were detected significantly more than other categories. Fig. 6 shows examples of results where debris identified as "Other plastics.". From these results, it was confirmed that:

- There were instances of misidentifying shells or colored natural objects as plastics.
- Single plastic objects were occasionally detected as multiple plastic objects.

So it will suggest (1) participants could easily spot plastics on the beach, leading to intensive collection, and
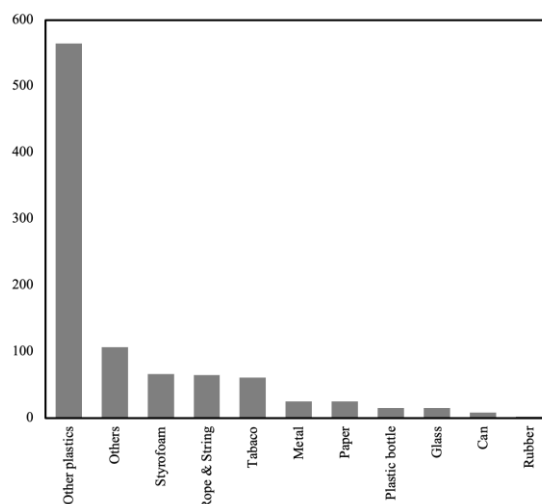
Fig. 5 The result of number of each category of detected debris

Ayumu Tominaga, Ryohei Komori, Eiji Hayashi

(2) to enhance object detection accuracy, reconstruction of the learning dataset for debris is necessary.
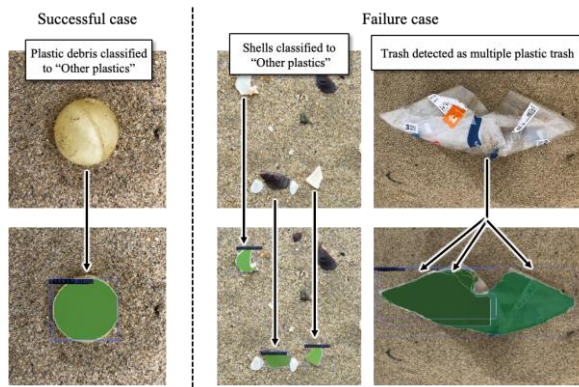
Fig. 6 Example of the input and detected "Other plastics"

5. Conclusion

In this study, a cloud service-based web application was developed to promote the quantification of beach-drifted debris. This web application, which can be used from a smartphone, aims to gather a large-scale beach debris dataset through beach cleanup events. One advantage of this proposal is that many people can easily participate in debris quantification work. In the case of experimental verification, the system successfully collected over 300 images capturing debris and detected over 900 debris with the participation of about 70 cleanup event participants. In the future, efforts will be made to enhance the usability of the application, making it less stressful for users. Additionally, long-term verification, spanning months or even years, will be conducted to address and improve the physical situation of beach-drifted debris.

References

1. S. C. Gall and R. C. Thompson, "The impact of debris on marine life," Marine Pollution Bulletin, Vol. 92, Issues 1–2, pp. 170-179, 2015.
2. K. Sasaki and et. al., "Coastal Marine Debris Detection and Density Mapping With Very High Resolution Satellite Imagery," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 15, pp. 6391-6401, 2022.
3. K. Moy and et. al., "Mapping coastal marine debris using aerial imagery and spatial analysis," Marine Pollution Bulletin, Vol. 132, pp. 52-59, 2018.
4. Y. Taddia, and et. al., "UAV Approach for Detecting Plastic Marine Debris on the Beach: A Case Study in the Po River Delta (Italy)," Drones, Vo. 5, No. 140, 2021.
5. K. Chen and et. al., "MMDetection: Open MMLab Detection Toolbox and Benchmark," arXiv:1906.07155, 2019.
6. K. He and et. al., "Mask R-CNN," 2017 IEEE International Conference on Computer Vision (ICCV), 2017.
7. P. F Proença and P. Simões, "TACO: Trash Annotations in Context for Litter Detection," arXiv:2003.06975, 2020.

Authors Introduction

Dr. Ayumu Tominaga
Dr. Ayumu Tominaga is assistant professor in Department of creative engineering, National Institute of Technology, Kitakyushu College. He received the Ph. D. from Kyushu Institute of Technology. His research interests a field robot such as for beach cleanup robot, and social implementation of a cloud service. He is a member of The Japan Society of Mechanical Engineers (JSME).

Mr. Ryohei Komori
Mr. Ryohei Komori is student in Department of creative engineering, National Institute of Technology, Kitakyushu College. His research interests include development of a SaaS or web service, application for democratizing of AI.

Prof. Eiji Hayashi
Prof. Eiji Hayashi is a professor in the Department of Intelligent and Control Systems at Kyushu Institute of Technology. He received the Ph.D. (Dr. Eng.) degree from Waseda University in 1996. His research interests include Intelligent mechanics, Mechanical systems and Perceptual information processing. He is a member of The Institute of Electrical and Electronics Engineers (IEEE) and The Japan Society of Mechanical Engineers (JSME).

Development of antagonistic wire-driven joint mechanism capable of rapid motion and variable stiffness

Katsuaki Suzuki
Kumamoto Industrial Research Institute
3-11-38 Higashi machi,Higashi-ku, Kumamoto 862-0901, Kumamoto, Japan
E-mail: k-suzuki@kumamoto-iri.jp

Yuya Nishida
Life Science and Systems Engineering, Kyushu Institute of Technology
2-4 Hibikino,Wakamatsu-ku, Kitakyushu 808-0196, Fukuoka, Japan
E-mail: y-nishida@lsse.kyutech.ac.jp

Kazuo Ishii
Life Science and Systems Engineering, Kyushu Institute of Technology
2-4 Hibikino,Wakamatsu-ku, Kitakyushu 808-0196, Fukuoka, Japan
E-mail: ishii@brain.kyutech.ac.jp

Abstract

The advancement of digital transformation in manufacturing is expected to demand novel tasks for industrial machinery unlike before. When adding mechanisms or actuators to enhance machine functionality, concerns arise about potential system enlargement and increased complexity. The design concept of integrating multiple functions into mechanisms using a few actuators and components serves as one means to address these issues. We aim to elucidate the structural arrangement capable of achieving three functions, including normal motion, rapid motion, and variable stiffness, using two electric actuators. In this paper, we introduce a newly devised antagonistic wire-driven joint mechanism aimed at improving the variable stiffness function within the articulated mechanism we have developed, enabling three functionalities.

Keywords: Rapid motion, Variable stiffness, Link mechanism

1. Introduction

Digital transformation is being promoted through the use of robots and digital technologies, such as the Internet of Things and artificial intelligence, with the aim of increasing productivity and manpower saving. In response, robots will be required to perform tasks that have not been done before, and it will be necessary to promote research and development of robots that can perform movements and functions that have not yet been generalized. One area is soft robotics, which can perform rapid motion and variable stiffness. For example, in the food industry, in order to realize the automation of the serving operation of prepared food, it is necessary for the robot to grasp the hardness of the food and to pick and place it according to the hardness of the food so that the soft food does not fall apart. Mochizuki et al. are conducting research on non-contact acoustic impedance estimation using ultrasonic waves before picking to optimize the gripping rigidity in order to reduce damage caused by food picking operations using robot hands [1]. To achieve automation, this system needs to be supplemented with a hand that can change the stiffness of the fingertips. In addition, percussion testing is being researched to predict the firmness of tomatoes to determine their expiration date, to estimate the sugar content of fruit, and to determine whether canned goods are defective [2], [3]. To automate percussion testing, an appropriate hammering mechanism is required in addition to a sensing system.

Adding additional mechanisms and actuators to integrate multiple functions into a robot can increase system complexity and size. Therefore, a design concept that allows mechanisms to have multiple functions with fewer actuators and parts is one way to address these issues.

Katsuaki Suzuki, Yuya Nishida , Kazuo Ishii

Aiming to realize three functions with a small number of actuators, we developed a joint mechanism that realizes normal motion, rapid motion, and variable stiffness with two motors [4], [5]. Our research results have shown that these functions can be realized. However, in the variable stiffness function, an end-effector does not rotate unless the load torque applied to the end-effector exceeds the torque applied to the end-effector by the pre-compressed spring to increase stiffness. In this paper, we propose a new antagonistic wire-driven joint mechanism that can rotate the end-effector at the moment load torque is applied to the end-effector, even in a stiff state, in order to improve the functionality related to variable stiffness of joint mechanisms that can realize the three functions.

2. Antagonistic wire-driven joint mechanism

Fig. 1 shows the proposed mechanism. The proposed mechanism consists of the end-effector, a main joint, two wires, two springs, two linear guides, and two cam units. Assume that the positions of the linear guide and main joint are fixed in space. Additionally, the linear guide and slider are a prismatic pair. The internal structure of the cam unit consists of a slider, a passive pulley, an input pulley, a cam follower, a cam, a worm wheel, a worm, and a motor.The function of the cam unit is to extend and retract the wire and release the energy stored in the spring. Rotation of the motor causes the cam to rotate through the worm gear mechanism. Since the cam follower and the input pulley are connected, when the input pulley rotates, the cam follower performs an orbital motion. Therefore, the input pulley rotates driven by the rotation of the cam. For ease of understanding, we will now discuss the structure using the simplified proposed mechanism shown in Fig. 2.

The main difference between Fig. 1 and Fig. 2 is that the cam unit is represented by an electric actuator. As shown in Fig. 3, during normal motion, the end-effector rotates counterclockwise with actuator 1 in the active state and actuator 2 in the passive state. For rapid motion, as shown in Fig. 4, when the two actuators are active, the two sliders move toward the main joint and energy is stored in the two compression springs. Now, when actuator 1 is in the braking state and actuator 2 is in the passive state, the forces on the wires are out of balance and the two sliders are moving away from the main joint. Since actuator 1 is in the braking state, the compression spring force acting on the sliders acts on the end-effector

through the wires, and the end-effector performs a rapid motion. Next, consider the variable stiffness function. As shown in Fig. 5, first assume a state in which the two springs are compressed by ε_{sp} and the two actuators are braked. Then assume a state in which a load torque is
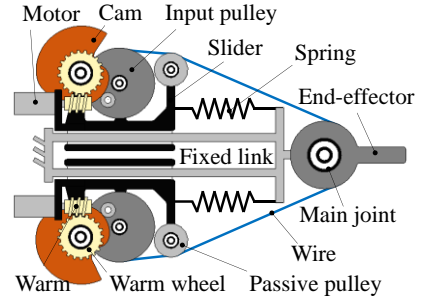
Fig. 1 Conceptual model
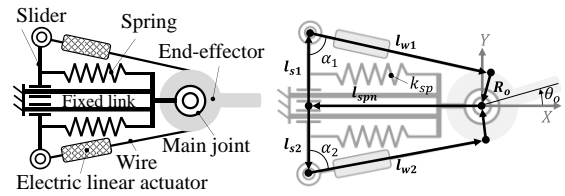
Fig. 2 Simple model
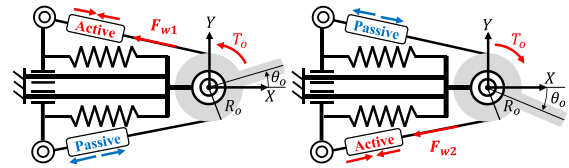
Fig. 3 Normal motion
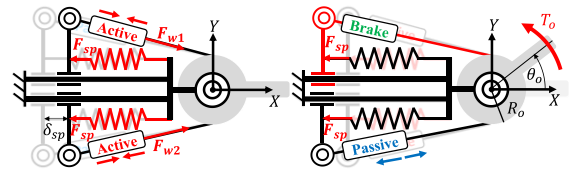
Fig. 4 Rapid motion
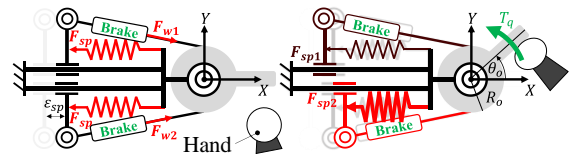
Fig. 5 Variable-stiffness function

applied to the end effector and the end effector is rotated counterclockwise to θ_o. At this time, wire 2 is wrapped around the end effector and slider 2 makes a translational movement toward the main joint, increasing the

compressive force acting on spring 2. Therefore, the torque due to the compression force of spring 2 acts on the end-effector in the direction of resisting the load torque. On the other hand, since the wire 1 originally wound around the end-effector is unwound, the slider 1 makes a translational movement in the direction away from the main joint, and the compression force originally acting on the spring 1 decreases. In other words, the torque due to the compression force of spring 1 acts on the end effector in the direction to assist against the load torque. By using this load torque support mechanism, the end-effector can be rotated at the moment when a load torque is applied to the end-effector. In addition, by adjusting the value of ε_{sp}, the compression force of the spring can be changed, and thus the rotational rigidity of the joint can be varied. The relation between the load torque T_q and the internal force acting inside the mechanism is expressed as follows based on their geometric relation:

$$T_q = k_{sp}R_o\left(\frac{\delta_{sp1}}{sin\,\alpha_1} - \frac{\delta_{sp2}}{sin\,\alpha_2}\right) \qquad (1)$$

In Eq. (1), α_* is expressed as follows:

$$\alpha_* = tan^{-1}\frac{l_{sp*}l_{w*} + l_sR_o}{l_sl_{w*} - l_{sp*}R_o} \qquad (2)$$

The rotational stiffness K of the end-effector is expressed using the following equation by differentiating the load torque T_q exerted on the end-effector by the angle θ_o of the output link:

$$K = \frac{dT_q}{d\theta_o} = K_{b2} - K_{b1} \qquad (3)$$

In Eq. (3), K_{b*} is expressed as follows:

$$K_{b*} = \frac{R_o^2k_{sp}l_{w*}}{l_{sp*}\,sin\,\alpha_*} + \frac{\left[\frac{R_o^2k_{sp}\delta_{sp*}l_{w*}l_s}{l_{sp*}(l_{sp*}^2 + l_s^2)} - \frac{R_o^3k_{sp}\delta_{sp*}}{(R_o^2 + l_{w*}^2)}\right]}{sin\,\alpha_*\,tan\,\alpha_*} \qquad (4)$$

In Eq. (4), with $\varepsilon_{sp} = 0$, $K_{b1} = 0$ for counterclockwise rotation of the output link and $K_{b2} = 0$ for clockwise rotation.

3. Analysis of output characteristics

Using the equations derived in Chapter 2, we analyzed the output characteristics with respect to variable stiffness. Fig. 6 shows the profile of the load torque T_q and angle θ_o of the end-effector at initial spring displacements ε_{sp} of 0, 5, 10, and 15 mm (Eq. (1)). Further, Table 1 shows the values of each parameter used in the analysis. The analysis was performed assuming that a load torque was applied to the end-effector and the

angle θ_o was rotated from 0[deg] (initial posture) to +40[deg] and then from 0[deg] to -40[deg]. The analysis results confirm that the load torque increases as the angle θ_o increases. Furthermore, as the value of ε_{sp} increases, the load torque required to rotate the joint increases. Even when ε_{sp} was set, the load torque applied to the end-effector did not have to exceed the torque applied to the end-effector by the pre-compressed spring. On the other hand, it can be observed that the maximum displacement angle decreases as the value of ε_{sp} increases. This is because as ε_{sp} increases, the distance between the slider and the main joint becomes shorter and the range of motion of the slider becomes smaller. Fig. 7 shows the profile of the rotational stiffness K and angle θ_o of the end-effector when the initial spring displacement ε_{sp} is 0, 5, 10, and 15 mm (Eq. (3)). It can be observed that as ε_{sp} increases, the rotational stiffness of the main joint increases. It can also be observed that the value of each stiffness increases rapidly near the end point of the motion. This is because as the angle θ_o increases, the distance between the slider and the main joint becomes shorter, and the tension generated in the wire increases as α_* becomes smaller.
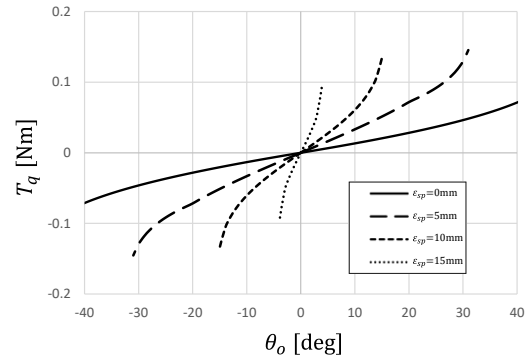
Fig. 6 Relationship between the rotational angle of end effector and load torque

Table 1. Design parameters

k_{sp}	Spring constant	500 N/m
R	Radius of end-effector	0.01m
l_s	Length of slider	0.02 m
l_{spn}	Natural length of spring	0.02 m
$max(\theta_o)$	Maximum movement range	$+40°$
$min(\theta_o)$	Minimum movement range	$-40°$
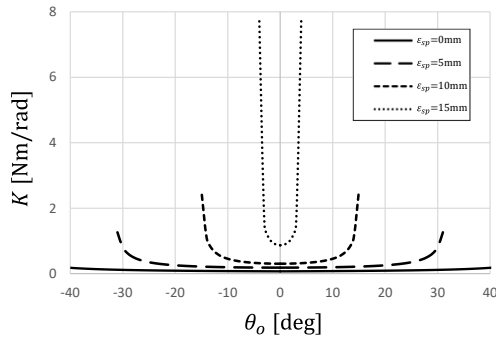
Fig. 7 Relationship between the rotational angle of
end effector and rotational stiffness

4. Conclusion

In this study, we describes the structure of the new mechanism and demonstrated a method for realizing normal motion, rapid motion, and variable stiffness. Additionally, we derived expressions for variable stiffness and analyzed. In the previously developed joint mechanism, the end-effector does not rotate unless the load torque exceeds the torque applied to the end-effector by the spring preloaded to increase the rigidity. As a result of the analysis, it was theoretically clarified that the newly proposed joint mechanism can solve this problem and rotate the end-effector at the moment the load torque is applied. Future issues include the search for a more optimal link ratio and the need to verify the effectiveness of the system by fabricating an actual machine and evaluating its performance.

References

1. Mochizuki, Ryuugo, et al. A Method to Estimate Reflection Coefficient of Ultrasonic Wave at a Boundary of Two Media. Journal of Robotics, Networking and Artificial Life, 2023, 9.4: 343-348.
2. ALIUA, I., et al. THE NONDESTRUCTIVE METHOD OF ACOUSTIC IMPULSE RESPONSE IN CHARACTERIZING TOMATOES FIRMNESS COEFFICIENT DURING STORAGE.
3. YOKOI Masayuki. A Study on Diagnosis Method using Hammering Test. JOURNAL OF OSAKA SANGYO UNIVERSITY Natural Sciences, 2021, 131: 59-69.
4. Suzuki, Katsuaki, et al. A new rotary actuator capable of rapid motion using an antagonistic cam mechanism. Journal of Advances in Artificial Life Robotics, 2020, 1.3: 143-151.
5. Suzuki, Katsuaki. Research on antagonistic wire-driven joints with rapid motion and variable stiffness functions, Ph.D. thesis, Kyushu Institute of Technology, 2021.

Authors Introduction

Dr. Katsuaki Suzuki

He is a Researcher at the Kumamoto Industrial Research Institute, Japan. He received his Ph.D. degree from the Kyushu Institute of Technology in 2021. His research interests include joint mechanisms and their applications.

Dr. Yuya Nishida
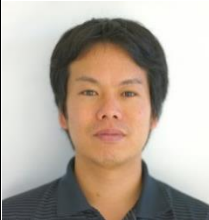
He is an Associate Professor at the Graduate School of Life Science and System Engineering, Kyushu Institute of Technology, Japan. His research interests include filed robotics, its application, and data processing.

Prof. Kazuo Ishii

He is a Professor at the Graduate School of Life Science and System Engineering, Kyushu Institute of Technology, Japan. His research interests include field robots and intelligent robot systems.

Gakken Hills Interdisciplinary Ekiden Competing with Humans, Animals, and Robots

Takuya Fujinaga
Fukuoka University
8-19-1 Nanakuma, Johnan, Fukuoka 814-0180, Japan

Moeko Tominaga
Nishinippon Institute of Technology
1-11, Aratsu, Kanda-machi, Miyako-gun, Fukuoka 800-0394, Japan

Daigo Katayama, Kazuo Ishii
Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka 808-0196, Japan
E-mail: t.fujinaga.wz@fukuoka-u.ac.jp
www.lsse.kyutech.ac.jp/~sociorobo/gakkenekiden/

Abstract

The number of robots exiting in human living spaces has increased, and interaction between humans and robots has become an important issue. The purpose of this study is to propose a future society where humans, animals, and robots coexist. We organize relay race, or ekiden in Japanese, where they compete together as runners. The competition rules have been established to ensure that each runner can compete fairly. This paper explains the competition rules and discuss the fairness of the rules

Keywords: Robotics competition, Ekiden, Human-Robot Interaction, Robot coexistence society

1. Introduction

Various robotics competitions [1], [2] have been held for social implementation with the development of robotics and artificial intelligence. In these competitions, robots mainly compete about functional performance. On the other hand, the content of robotics competitions changes to more challenging tasks aimed at coexistence between humans and robots [3].

Traditionally, robots have been introduced into industrial environments where safety has been ensured by cages. However, in recent years, the environment in which robots work has been changed to match the daily lives of humans. Research on human-agent interaction [4] and human-robot interaction [5] have aimed to the future society in which humans and robots coexist in harmony. On the other hand, in the universal situation of human society, there are many types of mobile objects: humans, animals, and vehicles. There are few examples of demonstration experiments or discussions about the role of robots in such environments.

We have held "Gakken Hills Interdisciplinary Ekiden" as the opportunity to think about the future society in which humans and robots coexist in harmony [6]. This ekiden is the relay race which humans, animals, and robots compete together on the same field. We held the 8th race on May 13, 2023. This paper describes the competition rules, introduces the characteristics and transition of participants from the 1st to the 8th races, and discusses the fairness of the rules.

2. Gakken Hills Interdisciplinary Ekiden

Takuya Fujinaga, Moeko Tominaga, Daigo Katayama, Kazuo Ishii

2.1. *Overview*

We organize this ekiden with the aim of making it a starting point for thinking about the future society in which humans and robots coexist and discussing the directions and challenges of such a society. The 1st race was held in 2014 and the 8th in 2023. Various runners have participated so far: humans regardless of age or gender, animals (mainly dogs) and robots with various movement and operation methods. To ensure fairness for the runners in this ekiden, where a wide variety of runners coexist, we have set our own competition rules. The competition rules have been adjusted every year based on law revision, developments in science and technology, and past competition results. In the following, the current competition rules are described.

2.2. *Rules for all runners*

Runners are divided into categories [P] (Person), [A] (Animal) and [R] (Robot). Category [P] is humans aged 18 years or older, [A] is dogs, cats, and other animals (excluding horses, cows, and other light vehicles) and [R] is robots and electric wheelchairs (excluding flying robots). One of the rules for the [R] sets the use of an electric motor as a prime mover, and electric wheelchairs are classified as the [R].

One team consists of four runners, and, in principle, the four runners are two males and two females. However, the [A] and [R] can be classified as either male or female. Note that within one team, the [A] and [R] can each have up to one runner participating. The race distance differs depending on the category: about 1.6km for the [P] and [A], and about 0.8km for the [R]. For details of the course, see the ekiden website [7].

This ekiden sets advantages so that humans, animals, and robots can compete fairly. The rank in the competition are determined based on the results of applying advantages to the race results. Details of the advantages are stated in 2.4.

2.3. *Rules for robots*

The rules for robot participation are as follows. We set these rules based on Japanese road traffic act.
- The max size is 120cm, 70cm, and 120cm in length, width, and height, respectively.
- An electric motor is used as a prime mover.
- Traveling speed is 6 km/h or less.

- No sharp protruding parts or devices that may cause harm to others.
- The robot has a function that enables it to be stopped quickly in an emergency.

The advantages are applied to different categories of runners in this ekiden, and robots have different advantages depending on their movement and operating methods. Runners in the [R] are classified into the following three classes.
- Direct operation class:
  requires human intervention (e.g. electric wheelchairs, electric kickboards)
- Remote operation class:
  no requires human intervention
- Automatic operation class:
  applicable to three or more of the following five items
  - Do not require a remote controller or other operating device (e.g. human following).
  - Do not require real-time operation by humans (e.g. driving using satellite positioning systems).
  - Make action selections based on environmental recognition (e.g. obstacle avoidance).
  - Have originality (e.g. does not use existing products as a base, or not wheel-based type).
  - Submit a technical poster

Note that in this ekiden, the human intervention is defined as pushing, carrying, or directly touching the robot to adjust its traveling direction by humans.

2.4. *Advantages*

The advantages are determined by the age of the [P], the participation of animals, the robot specifications. The results which are applied the advantages are used to determine the final competition results and the rank. Three types of advantages are defined: age advantage, animal advantage and robot advantage. For the age advantage, let N_p and Age_{total} be the number of participants in category [P] and their total age, respectively. If Eq. (1) is satisfied, Adv_p seconds (Eq. (2)) is subtracted from the race result.

$$Age_{total} > 25 \times N_p \qquad (1)$$

$$Adv_p = Age_{total} \times 2 \qquad (2)$$

For the animal advantage, let T_A be the race time in the [A]. The unit is seconds. The result R_A seconds which applying the animal advantage is calculated from Eq. (3). Note that the method of calculating R_A depends on the value of T_A.

$$R_A = \begin{cases} \dfrac{1}{2}T_A + 120 \ (T_A \le 600) \\ \dfrac{1}{10}T_A + 360 \ (600 < T_A \le 2400) \end{cases} \quad (3)$$

For the robot advantage, let R_A be the race time in the [R]. The unit is seconds. The result R_R seconds which applying the robot advantage is calculated using one of Eqs. (4) to (6). Eq. (4) is for the direct operation class ($R_{R(Direct)}$), Eq. (5) for the remote operation class ($R_{R(Remote)}$) and Eq. (6) for the automatic operation class $R_{R(Auto)}$). Note that T_R is between 480 seconds (8 minutes) and 2400 seconds (40 minutes).

$$R_{R(Direct)} = \frac{1}{10}T_R + 390 \quad (4)$$

$$R_{R(Remote)} = \frac{1}{10}T_R + 270 \quad (5)$$

$$R_{R(Auto)} = \frac{1}{10}T_R + 150 \quad (6)$$

The constant number of these equations for the advantages was determined based on the results of previous competitions results to ensure that all runners can compete fairly.

3. Results and discussion

Table 1 shows the transitions in the number of participants from the 1st to the 8th races. The number of runners per team was six from the 1st to the 5th races, and from the 6th race, it was four due to a change in the rules. Regarding the proportion in each race, the category [P] accounts for 70-90%, the [A] was just under 10%, and the [R] was just over 10%. The number of participants in the [A] so far has been 15, all of them dogs. All runners in the [P] and [A] completed the race. The number of participants in the [R] so far has been 45. For the traveling method of the [R], two of the 45 robots were multi-legged or bipedal walking types rather than wheel-based type. For the operating method, six out 45 robots were semi- or full-automatic. Examples of its

Table 1 Transitions in the number of participants

Times	Team	[P]	[A]	[R]
1	13	70	1	7
2	10	52	1	7
3	7	31	3	8
4	6	30	2	4
5	8	43	1	4
6	10	34	3	3
7	12	40	2	6
8	12	40	2	6

Table 2 Average and standard deviation (S.D.) of 8th race results in each category (m: minutes, s: seconds)

Category	Average	S.D.
[P]: male	8m21s	1m33s
[P]: female	10m19s	1m36s
[A]	7m04s	16s
[R]	6m50s	54s
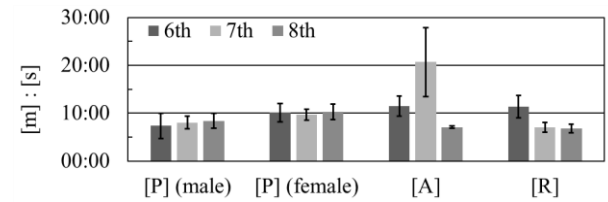
Fig. 1: Average and standard deviation of each race result in each category

operation methods were as follows: human tracking using image processing, driving using waypoints, route tracking using the global navigation satellite system, simultaneous localization and mapping (SLAM) using a laser sensor and an inertial measurement unit. Runners who completed the race within the time limit was 35 robots. The multi-legged and bipedal walking types mentioned above were unable to complete the race. Of the six autonomous robots, two were able to complete the race within the time, one using waypoints and the other using SLAM.

Table 2 shows the average and standard deviation of the results for each category in the 8th race. The results of the [A] and [R] were equal to or higher than those of the [P]. Therefore, rules regarding advantages works, and runners could compete regardless of their characteristics. The details are on the ekiden website [7].

This paper discusses the validity of rules regarding advantages. Fig. 1 shows the average and standard deviation of the competition results for each category since the 6th race. The results for the [A] and [R] are the results of applying the advantages. Note that the race distance for the [R] has been changed since the 6th race, so only the results from then on are shown. The results of the [P] have little variation each race. On the other hand, the [A] was at a slow compared to the [P] in the 6th race, and furthermore, the standard deviation was large in the 7th race. The [R] is also at a slow compared to the [P] in the 6th race. However, in the 7th race, the

Takuya Fujinaga, Moeko Tominaga, Daigo Katayama, Kazuo Ishii

robot advantage was revised based on the previous race results of the [P], so the results of the 7th and 8th races of [R] and [P] are fair. Based on this result, the animal advantage was similarly revised for the 8th race. In the 8th race, the results of the [A] and [R] were equal to or better than those in the [P]. Therefore, these results show that for each runner to compete fairly, it is appropriate to set an advantage based on a certain standard.

4. Conclusions

We have held the ekiden for thinking about a future society in which humans and robots coexist in harmony. The 1st race was held in 2014, and the 8th in 2023. The rules have been revised for each race based on previous race results and law revision. With the spread of robots, laws regarding robots were established, and the rules for the 8th race were determined based on the laws. This paper showed validity of the rules from race results. We have continued to discuss how robots coexist with humans in future society through this ekiden.

References

1. K. Ishii et al., "Report on the 8th Tomato-Harvesting Competition toward Smart Agriculture", Int. Conf. on Artificial Life and Robotics, pp. 484-488, 2023.
2. H. A. Yanco et al., "Beyond Usability Evaluation: Analysis of Human-Robot Interaction at a Major Robotics Competition," Human-Computer Interaction, 19(1-2), pp.117-149, 2004.
3. H. Okada et al. "What is the Purpose of the World Robot Summit Service Category," J. of the Robotics Society of Japan, 37(3), pp. 218-223, 2019. (In Japanese)
4. S. Yamada et al., "Mutual Adaptation to Mind Mapping in Human-Agent Interaction," IEEE Int. Workshop on Robot and Human Interactive Communication, pp. 105-110, 2002.
5. R. Jahanmahin et al., "Human-robot interactions in manufacturing: A survey of human behavior modeling," Robotics and Computer-Integrated Manufacturing, 78, 102404, 2022.
6. T. Fujinaga et al., "Report on the 8th Gakken Hills Interdisciplinary Ekiden," The 31st Symposium on Fuzzy, Artificial Intelligence, Neural Networks and Computational Intelligence, 2023. (In Japanese)
7. Center for Socio-Robotics Synthesis, Kyushu Institute of Technology, "Gakken Hills Interdisciplinary Ekiden," https://www.lsse.kyutech.ac.jp/~sociorobo/gakkenekiden/ (In Japanese, Access 2023-12-12)

Authors Introduction

Dr. Takuya Fujinaga

He is an Assistant Professor at the Department of Electronics Engineering and Computer Science, Fukuoka University, Japan. His research interests are in agricultural robots, underwater robots and field robotics.

Dr. Moeko Tominaga
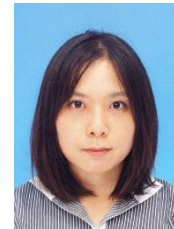
She is an Assistant Professor at the Department of Integrated System Engineering, Nishinippon Institute of Technology, Japan. Her research interests are in multi-agent systems and field robotics.

Dr. Daigo Katayama
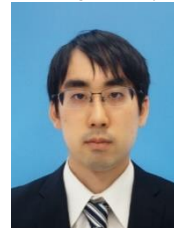
He is an Assistant Professor at Kyushu Institute of Technology, Japan. His research interests are in welfare devices, mobile systems, and field robotics.

Dr. Kazuo Ishii

He received his PhD from the University of Tokyo, Japan, in 1996. In 2011, he joined Kyushu Institute of Technology and is currently a professor in the Department of Human Intelligent Systems. His research interests include information communications and marine robotics. He is a member of IEEE.

Development of a Low-Cost Underwater Robot for Research and Education

Takuya Fujinaga
Fukuoka University
8-19-1 Nanakuma, Johnan, Fukuoka 814-0180, Japan
E-mail: t.fujinaga.wz@fukuoka-u.ac.jp

Abstract

Underwater robots are utilized for underwater observation. Generally, the development of underwater robots demands a significant amount of funds, posing a high barrier for research, development, and education. The aim of this study is to reduce these barriers, facilitating broader implementation of research, development, and education in the field of underwater robotics. This paper describes the development policy and cost of a low-cost underwater robot and discusses considerations based on the results in the robotics competition.

Keywords: Underwater robot, Underwater drone, Remotely operated vehicle, Low-cost, Robotics competition

1. Introduction

Underwater robots are used as a means of underwater observation. Research and development of underwater robots has been studied since the 1960s [1]. Recently, underwater robots are called as underwater drones and become spread to society. When using underwater robots as tools for underwater geological and biological surveys etc., it is beneficial to use commercially available underwater robots. However, when underwater robots are the subject of research and development or education, the specifications required for the robot vary depending on the purpose. Therefore, it is necessary to develop the robot independently. Underwater robots come in a variety of sizes, from large to small, and research and development has been carried out depending on the purpose [2], [3]. There are also underwater robots developed for education [4].

Generally, the development of underwater robots demands a significant amount of funds, posing a high barrier for research, development, and education. Although some underwater robots have been developed for low-cost, the development cost still ranges from $4,000 to $15,000 [5], [6], [7]. Underwater robots differ from robots that move on land or in the air, and the following are the minimum requirements for their development: watertightness, pressure resistance, and adjusting buoyancy etc. There are also various types of tacit knowledge.

The aim of this study is to reduce barrier of the development cost, facilitating broader implementation of research, development, and education in the field of underwater robotics. This paper describes the development policy and cost of a low-cost underwater robot and discusses considerations based on the results of the robotics competition.

2. Method

2.1. *Development policy*

Underwater robots are mainly classified into five types: glider, cruise, hovering, biomimetic, and vehicle. The selection of each type is determined by the target work and research policy. The glider-type generally does not have propulsion devices and can adjust their own buoyancy and center of gravity, enabling long-term observation. The cruise-type has a long and slender structure, so it has high propulsion efficiency and can conduct wide-ranging surveys. The hovering-type can perform more complex movements than the above-

mentioned types by attaching multiple propulsors, making detailed observations possible. The biomimetic-type imitates the characteristics of fish and other animals, can be propelled by fish fins rather than propellers, and are superior to the above types in terms of energy efficiency and consideration for marine life. The vehicle-type can move on the seabed or on the hull of a ship and perform work on those surfaces, such as digging and cleaning. This study adopts the hovering-type for the purpose of detailed underwater observation. The underwater robot developed in this study is designed to operate underwater for about 30 minutes to 1 hour, targeting shallow part of the ocean with an observation depth of 10m or less.

The main components of the hovering-type underwater robot that performs underwater observation are as follows: computers, sensors, thrusters, energy sources. The computers control the underwater robot by controlling the propeller, acquiring sensor information, and selecting the actions. There are various types of sensors depending on the purpose of observation, such as cameras and acoustic devices. Underwater robots also require positioning and navigation equipment and sonar sensor to measure their own posture, position, speed, and surrounding environment. In some cases, sensors account for a large portion of the development cost. Electric thrusters are generally used as propulsion devices, and a high degree of freedom in movement performance is achieved by attaching multiple thrusters. The energy sources need to be selected appropriately to operate the above-mentioned components at the desired operating time. In addition to the above components, communication equipment, work equipment, ballast, etc. are added as necessary.

The low-cost underwater robot in this study also consists of computers, sensors, thrusters, and energy sources. The sensors are equipped with a camera, depth sensor, and inertial measurement unit (IMU). As an option, add underwater lights to conduct underwater observations. In order to control this robot from the ground and check the data acquired by the robot, the external computer on the ground and the internal computer on the underwater robot are connected with an optical fiber cable, and this robot is moored from the ground using this cable. Therefore, this robot becomes a remotely operated vehicle (ROV). However, the development policy is to make it easy to expand the

sensor so that it can be changed to an autonomous underwater vehicle (AUV).

The functional requirements for the low-cost underwater robot are described above. As non-functional requirements, this robot is considered portability and parts availability. For the portability, the design is based on the weight and ease of carrying by one person. To ensure the ease of obtaining the parts, use parts that can be easily purchased at online shopping or home centers. Note that some parts are manufactured using 3D printer and laser beam machining. Regarding the development cost, previous studies has proposed a cost of $4,000 to $15,000 for a low-cost underwater robot. In this study, the development cost is less than $2,000.

2.2. *Low-cost underwater robot*

The design drawing of the low-cos underwater robot is shown in Fig. 1, and the components are shown in Table 1. Although, there is a certain distance offset between each component to show the components in Fig. 1, each component is connected as shown in the actual appearance in Fig. 2. This robot mainly consists of the following.
- Frame (No. 1 and 2 in Fig. 1)
- Main box (No. 8) that stores the computers, depth sensor, IMU, etc.
- Battery box (No. 9) that stores the batteries and circuits
- Thrusters (No. 21 and 22)
- Camera unit (No. 4)
- Underwater lights (No. 5)
- Acrylic plates and joints to connect the above components (No. 16-19)

The energy from the battery box is supplied to the main box via cable gland. A total of six thrusters are installed, total of four ones in the front and rear for movement in the xy-plane, and two in the center for movement in the z-axis direction. The camera unit is independent from the main box, and the data is acquired by connecting its power and signal lines to the main box. Each box has holes for cable glands, and the number of holes can be increased or decreased as the number of sensors or thrusters changes, ensuring expandability. Additionally, when this robot adjusts its buoyancy, it fills the poly vinyl chloride (PVC) pipe of the frame with fresh water to maintain neutral buoyancy. A cap (No. 3) is provided to make it easy to put in and take out the water.

The specifications of this robot are shown in Table 2, and the system architecture is shown in Fig. 3. This robot is a size and weight that can be carried by one person. In terms of watertightness and pressure resistance, it can withstand water depths of 1.5m for 30 minutes. Testing at deeper water depths is a future task. To access the internal computer from the external computer, the ethernet connection is converted to an optic fiber connection inside the main box, and the signal is transmitted to the external computer through the optical fiber cable. The external computer is paired with the joystick device via Bluetooth, and the operator can remotely control the underwater robot by operating the joystick device. The internal computer is connected to each sensor and a micro controller unit (MCU). The MCU sends signals to the electronic speed controllers (ESCs), which controls the thruster speed, and to the underwater lights and indicator LEDs. The LEDs are installed to check the on/off status of the internal computer and the operating status of this robot. Fig. 4 shows the circuit architecture inside the battery and main boxes. The battery box has two power supplies: 11.1V for the thrusters and underwater lights, and 5V for the computer and optical media converter. Note that 5V is converted from 7.4V batteries. For the software, this robot distributes processing between the internal computer and MCU and the external computer (Fig. 5). The former acquires data from each sensor and controls the thrusters, underwater lights, and indicator LEDs, while the latter sends commands to the underwater robot and analyzes the acquired data. This robot realizes distributed processing using robot operating system (ROS).

The development cost for this robot is approximately $1,800 (250,000 yen). The proportions of parts by categories are as follows: 6.0% for the frame, 9.6% for cable-related parts, 30.4% for the main box including internal computer, MCU, and ESCs, etc., and 6.8% for the battery box including batteries, converters, and

(a) Front view   (b) Rear view

(c) Side view   (d) Top view
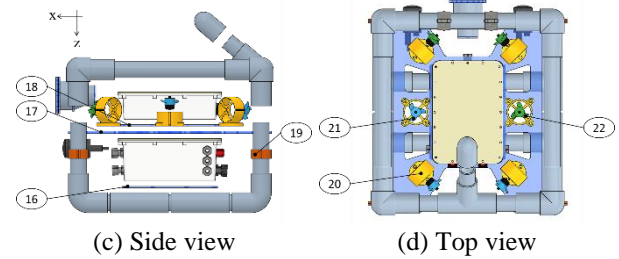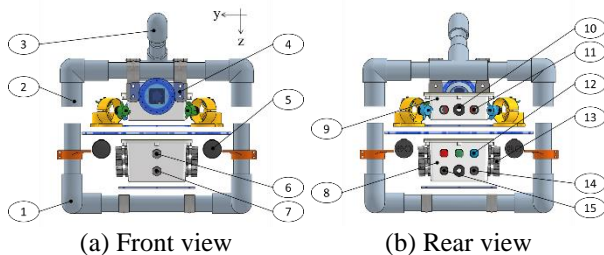
Fig. 1 Design of a low-cost underwater robot (see Table 1 for legend)

Table 1 Components of a low-cost underwater robot

No.	Component	No.	Component
1	Lower frame	12	Indicator LDEs
2	Upper frame	13	Cable gland for thrusters
3	Cap	14	Cable gland for the underwater lights
4	Camera unit	15	Depth sensor
5	Underwater Lights	16	Plate for fixing the main box
6	Cable gland for the camera unit	17	Plate for connecting the two boxes
7	Cable gland for a signal line	18	Plate for fixing the battery box
8	Main box	19	Joint for fixing the two boxes
9	Battery box	20	Thruster guards
10	Cable gland for a power line	21	Thrusters for clockwise propellers
11	Switches	22	Thrusters for counterclockwise propellers

relays, etc., 8.0% for thrusters, 13.6% for sensors, 17.6% for underwater lights, and 8.0% for other including mechanical and electronic parts and 3D printer materials. The frame is used PVC pipe. Products from *TAKACHI Electronics Enclosure* Co., Ltd. are used for the cable gland, main and battery boxes. The internal computer is Jetson Nano and the MCU is Arduino Micro. The type of battery is LiPo. The thruster is a brushless motor for drones, and the motor has a rotation speed of 3100rpm per unit voltage when no

load is applied. The ESC, camera, depth sensor, etc. are products from *BlueRobotics*. The IMU is used BNO055 produced by *BOSCH*.

3. Results and discussion

To evaluate the performance of this robot, it participated in an underwater robotics competition held on October 7, 2023 [8]. This robot entered in the AUV category, but its operation method was ROV. The field of this competition consists of the following areas: for slaloming between buoys, for passing through gates, for observing, for searching an acoustic source, for floating. This robot challenged areas other than the searching an acoustic source area. Fig. 6 shows the depth data and yaw angle data during this competition. This robot started to dive into the field (Fig. 6), moved on the bottom ((ii)), traveled the slaloming area ((iii)), passed the gate, and observed the target ((iv)). Then, this robot turned its direction of travel toward the starting point, that is, rotated its heading 180 degrees, and returned to the starting point. Finally, this robot attempted to float, but the fuse for the 11.1V blown, making it impossible to continue and the trial ended.

The results of this competition confirmed that this robot was able to move on the xy-plane at a speed of approximately 0.15m/s. Fig. 6 showed that the depth and posture of this robot could be measured using each sensor. In the observation area, the USB camera inside this robot was able to read the target panel with QR code and numbers printed on it. On the other hand, regarding the movement in the z-axis direction, the xy-plane of this robot has a larger square measure than other planes, so the underwater drag during floating or diving is also large. For this reason, the load on the thrusters was greater in the z-axis movement than in the xy-plane one, and the fuse worked as a safety measure.

Fig. 2 View of developed low-cost underwater robot

Table 2 Specifications

Dimensions	410 x 400 x 240mm (L x W x H)
Weight	7.0kg (8.0kg when the pipe is filled with water)
Batteries	11.1V, 2.2Ah x 2, 7.4V, 2.2Ah x 2
Computer	CPU: Quad-core ARM A57 @ 1.43GHz, Memory: 4GB
Sensors	USB camera, Depth sensor, IMU
Communications	Ethernet, Optical LAN
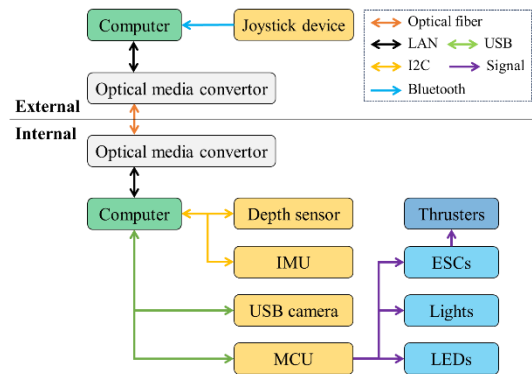
Fig. 3 System architecture for in/external of this robot
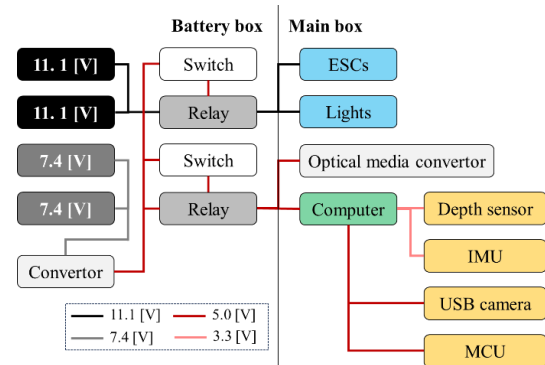
Fig. 4 Circuit architecture
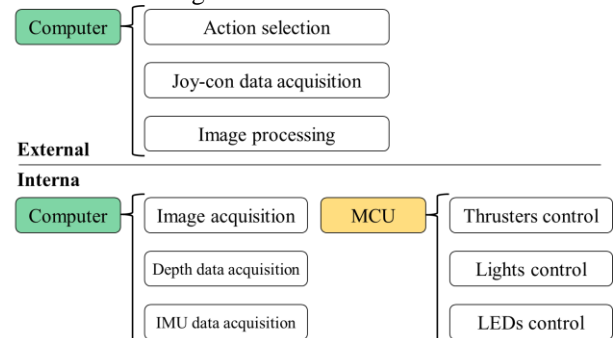
Fig. 5 Software architecture of in/external computer
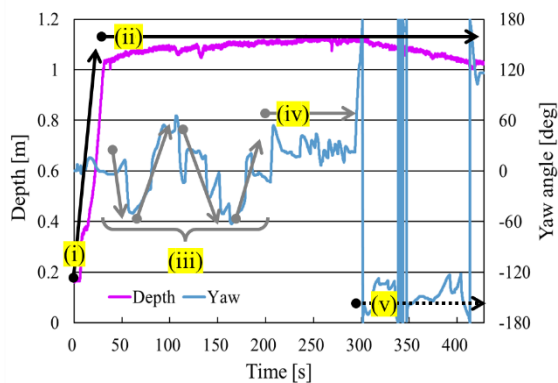
Fig 6 Results of depth data and yaw angle

This robot used a glass tube type fuse, and the holder for it is made of resin. During test before this competition, the plastic holder melted due to the heat generated by the load caused by the rotation of the thrusters. Therefore, as a safety measure, a 5A fuse was used. However, when designing this robot, cables were selected with a permissible current of 15A. Based on these findings, future tasks for this robot include reconsidering the shape of this robot and redesigning the circuit inside the battery box so that it can move in the z-axis direction. Additionally, the purpose of this study is to develop a low-cost underwater robot platform and make it open access. Towards this end, future tasks also include conducting pressure tests, implementing depth and heading control, dead reckoning, and underwater object detection using each sensor.

4. Conclusion

This paper described the development policy and cost of a low-cost underwater robot. This robot was a hovering-type and ROV. The development cost was $1,800. This paper analyzed the sensor data and discussed the motion performance based on the results in the robotics competition. The next tasks are to improve motion performance in the vertical direction and make it an AUV using sensor data. The goal is to open source the requirements and tacit knowledge for the development of the low-cost underwater robot.

References

1. U. Tamaki, "Robots for Underwater World," J. of the Robotics Society of Japan, 22(6), pp. 692-696, 2004. (In Japanese)
2. Y. Nishida, et al., "Underwater Platform for Intelligent Robotics and its Application in Two Visual Tracking Systems," J. of Robotics and Mechatronics, 30(2), pp. 238-247, 2018.
3. T. Matsumura, et al. "Development of a Handy Autonomous Underwater Vehicle "Kyubic," Int. Conf. on Artificial Life and Robotics, pp. 405-408, 2021.
4. H. Yamagata, et al., "Development of Underwater Robot for Teaching Material of Basic Mechanics," J. of the Robotics Society of Japan, 33(3), pp. 181-188, 2015. (In Japanese)
5. B. Allotta et al., "A low cost autonomous underwater vehicle for patrolling and monitoring," J. Engineering for the Maritime Environment, 231(3), pp. 740-749, 2017.
6. B. R. Page, et al. "Highly Maneuverable Low-Cost Underwater Glider: Design and Development," IEEE Robotics and Automation Letters, 2(1), pp. 344-349, 2017.
7. C. Edge at al., "Design and Experiments with LoCO AUV: A Low Cost Open-Source Autonomous Underwater Vehicle," IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, pp. 1761-1768, 2020.
8. Techno-Ocean2023, http://ton23.underwaterrobonet.org/ (In Japanese, Access 2023-12-12)

Author Introduction

Dr. Takuya Fujinaga

He is an Assistant Professor at the Department of Electronics Engineering and Computer Science, Fukuoka University, Japan. His research interests are in agricultural robots, underwater robots and field robotics.

Enhancing Reconnaissance Missions Through Multiple Unmanned Systems in ROS

Anees ul Husnain
Department of Electrical Engineering, Faculty of Engineering, University of Malaya, Malaysia
The Islamia University of Bahawalpur, Pakistan

Norrima Mokthar
Department of Electrical Engineering, Faculty of Engineering, University of Malaya, Malaysia

Takao Ito
Hiroshima University, Japan

Siti Sendari
Universitas Negeri Malang, Indonesia

Muhammad Farris Kyasudeen
University Technology MARA (UiTM), Malaysia

Muhammad Badri M Noor
Universiti Malaya, Malaysia, Ifcon Technology Sdn Bhd, Malaysia

Heshalini Rajagopal
Department of Electrical and Electronics Engineering, MILA University, 71800 Nilai, Negeri Sembilan, Malaysia

Email: norrimamokhtar@um.edu.my

Abstract

The synergistic collaboration between UAVs and UGVs addresses the limitations of individual platforms, offering a versatile solution for reconnaissance tasks in diverse environments. The proposed system employs ROS as the underlying architecture to facilitate seamless communication and coordination among multiple UAVs and UGVs. We delve into the intricacies of developing a robust communication framework that enables real-time data exchange and decision-making, fostering a synchronized and adaptive operation. Furthermore, the article explores strategies for path planning and navigation, considering the unique mobility constraints of UAVs and UGVs. Optimal coverage is ensured through efficient exploration and coverage of the reconnaissance area, under by comparing raster-scan, expanding spiral and zig-zag area exploration approaches. The article concludes by discussing potential extensions, such as the integration of machine learning techniques for enhanced autonomy and the scalability of the system for larger-scale missions by presenting a ROS-based framework that maximizes the synergy between UAVs and UGVs.

Keywords: Heterogeneous unmanned systems, Cooperative Path Planning, Reconnaissance, UAV, UGV, ROS

1. Introduction

With the surge in utilization of unmanned systems for research and their applications, the need for cooperation among heterogenous systems has made to the spotlight. The cooperation among heterogeneous systems exploits the diversity of features in unmanned system to achieve such mission objectives which are near impossible for homogenous systems. Forming a cooperative reconnaissance strategy for multiple unmanned vehicles in unknown areas has proven to be reliable and improves the heading errors of the vehicles [1]. Effective reconnaissance in modern military applications possesses the key for a strategic advantage, and also helps to neutralize surprise attacks, ambushes and can prevent the damages through improvised explosive devices (IEDs) with proactive sensory modules [2]. This is why combining Unmanned Ground Vehicles

Anees ul Husnain, Norrima Mokthar, Takao Ito, Siti Sendari, Muhammad Farris Kyasudeen, Muhammad Badri M Noor, Heshalini Rajagopal

(UGVs) alongside UAVs to coordinate area reconnaissance combines the technical and tactical capabilities of both to achieve the same. However, there are certain challenges linked with combing these unmanned systems.

This work addresses one of such challenges that naturally appears when the UAVs and UGVs are required to perform a coordinated reconnaissance in an unknown area. The nature of terrain and obstacles can affect the conditions required to maintain the reconnaissance formation between the aerial and ground vehicles [3]. A set of conditions is established which comprises two key components: a). The unmanned vehicles, UAV and UGV, have to explore the territory in close proximity. b). Minimize the impact of terrain-based slowness of UGV on the UAV.

This work presents a novel technique and coordination architecture that ensures the conditions above and explores the area effectively.

2. Proposed Method

2.1. UAV / UGV Coordination Architecture

The requirements for reconnaissance formation are primarily derived from the differences in maneuverability between both of the vehicles and impact of terrain on movability of UGV. The relative positions of the heterogenous vehicles, reconnaissance objectives and the environment data (grid map) are combined to generate the planning data, as shown in Fig. 1. The planning data further combines with the relative positions of UAV and UGV, and the trajectory data from UAV is fed to the path planning unit. The path planning unit, considering the conditions mentioned in Section 1, generates instructions for both UAV & UGV for coordinated maneuvers alongside requirements like collision avoidance.
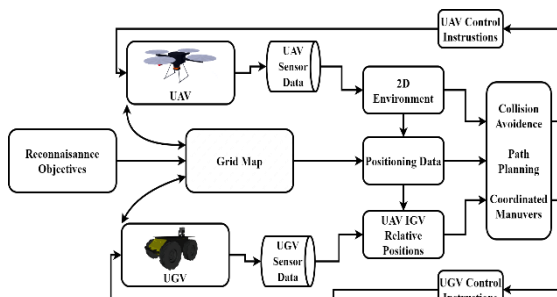
Fig. 1. UAV/UGV Coordination Architecture

2.2. Measure of Objectiveness in coordination

Self-assembly is a technique to attain cohesion in heterogenous robots which is perceived as robotic intelligence. The robots are required to assemble themselves that may not be possible through maneuverability [4]. There are three primary objectives under self-assembling the UAV and UGV to generate maneuver instructions and capitalize on interaction between them to work on complex tasks more efficiently.

(i)   Robustness: To reperform the task while maintaining the cohesion in proximity in case of any missing area to be explored.

(ii)   Versatility: To adjust and readjust the maneuvers because of uneven terrain and ground obstacles for the UGV.

(iii)   Cost reduction: To minimize the cost for the above. Here the cost is computed in terms of the time lapsed and number of turns a robot has to take. through design of large numbers of robots.

2.3. Attraction / Repulsion Mechanism in UAV / UGV

There are two primary tasks which are being performed at all the time: collision avoidance for the ground vehicle and flocking of UAV with respect to the ground vehicle. Collision avoidance in this work is basically a necessary condition for the UGV to maintain a minimum safe distance from its surroundings objects.

This work is inspired from a recent published work which engaged heterogenous robots in an 'invisible binding' by implementing an "Attraction / Repulsion" mechanism using Voronoi partitioning, based on Voronoi Diagrams for collision avoidance [5]. However, our approach exploits the idea of attraction-repulsion for heterogenous robots and generates path instructions in a zone-based coordinate system. Aggregation and dispersion are the two phenomena which are derived from the inter-robot's distance and their distribution in the area needed to be explored. This work mitigates this problem through establishing zones based on the distances and focuses on reconnaissance of the territory.

The proposed model consists of four constraints i.e. placement, repulsion, attraction, and orientation. The placement may take place either randomly or with some prior information from the area of reconnaissance. The model is based on a set of universal equations proposed by Ian Couzin cited by [6]. Referring to Eq. (1), Eq. (2)

and Eq. (3), where dr, do and da represent the radii of repulsion, operation and attraction zones with i and j representing coordinates. Fig. 2 presents an illustration for the three scenarios.

$$d_r(t + \tilde{L}) = -\sum_{j \neq i}^{M_r} \frac{rij(t)}{|rij(t)|} \quad (1)$$

$$d_o(t + \tilde{L}) = \sum_{j \neq i}^{M_o} \frac{Vj(t)}{|Vj(t)|} \quad (2)$$

$$d_a(t + \tilde{L}) = -\sum_{j \neq i}^{M_a} \frac{rij(t)}{|rij(t)|} \quad (3)$$
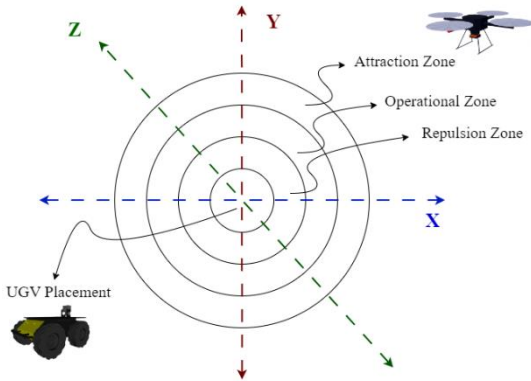
Fig. 2. Attraction Repulsion Mechanism with Placement and operational zone

2.4. Parallel Maneuver Generations for UAV / UGV

Here it is important to mention that either of the heterogeneous robots has to be set as the leader or follower. Moreover, the radii of the zones are adaptable according to the terrain and area under consideration. The maneuver generation takes sensor and positioning data from both unmanned vehicles and create relative positions of UAV and UGV, refer to the Fig. 1. This data generates maneuvering instructions in a 2D workspace, assuming a constant altitude of the UAV. Fig. 3 presents the maneuver generations of for both UAV and UGV. To better understand the flow of an instruction generation, consider the UAV and UGV in the placement zone.
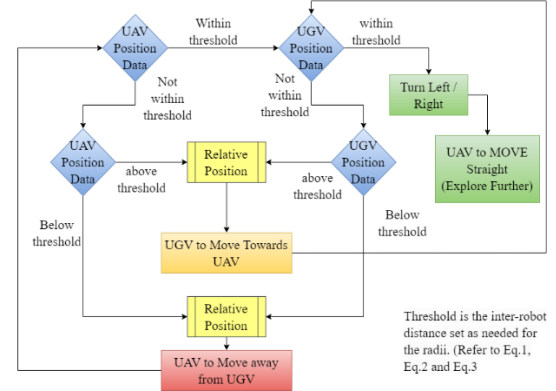
Fig. 3. Parallel Maneuver Generation for UAV /UGV

3. Implementation

3.1. Simulation Setup

Robot Operating System (ROS) is an opensource platform that provides libraries for and tools for a diverse range of robotic applications. These libraries include the hardware abstraction, driver packages, data visualizers, inter-module communication packages and protocols [7]. Gazebo on the other hand simulates the robotic application in a simulated environment [8]. Both these projects are lead by a community called open robotics [9].

For this work, we consider HUSKY and HECTOR which are emulated versions of opensource UGV and UAV, respectively. The integration of Husky and Hector within ROS provides the foundation of collaborative autonomous systems, showcasing the potential coordination among heterogeneous robots. There are four key components for setting up this collaboration:
  (i)   Sensor data Integration
  (ii) Communication of sensor data, control instructions and status of UAV / UGV
  (iii) Motion Control and Mission Planning.
  (iv) Mapping and Navigation.

3.2. Experimentation

The experimental scenarios are established for a comprehensive evaluation of the collaboration and exploration capabilities of Hector and Husky. There are two primary considerations in mind before setting up experimental scenarios:
  (i)   To validate the ability of this collaboration system for exploration and mapping
  (ii) To evaluate the adaptiveness of the system against the challenges in the environment. (Under this work, the challenges refer to static obstacles and uneven terrain.)

Anees ul Husnain, Norrima Mokthar, Takao Ito, Siti Sendari, Muhammad Farris Kyasudeen, Muhammad Badri M Noor, Heshalini Rajagopal

3.3. Flow of Activities

(i) Initialize the ROS-Gazebo environment through ROS launch file. For uneven terrain, there is a stock environment available in Gazebo which has been deployed.

(ii) Deploy Hector UAV and Husky UGV equipped with LiDAR sensors.

(iii) Command Husky for Autonomous Exploration and Hector for aerial mapping.

(iv) Initiate the program for exchanging Lidar and positioning data between Husky and Hector and compute relative positions. The relative data and position histories are stored and explored regions are marked with the coordinates of the environment.

(v) Constraints initialization: This is the main program which generates the thresholds from the relative positions. This module also utilizes LiDAR data from Husky for its collision avoidance.

(vi) Creation of unified map that combines the explored area from both of the platforms.

(vii) Collection of performance parameters which are percentage of area explored, number of maneuvers required, and overall time taken.

(viii) Analysis of exploration data and suggestions for optimizations.

4. Results and Discussion

Since there isn't any notion of geo-referenced coordinates in Gazebo, therefore, the positioning data in the simulation, for both the platforms, is collected in Cartesian Coordinate space. Fig. 4 presents the screenshot of deployment of both platforms while coordinating repulsive motion.

Fig. 4. Hector and Husky Deployment in a stock environment in Gazebo

The synergic collaboration between the two autonomous platforms performed simultaneous mapping based on the combining the aerial and ground data. The UGV is instructed to follow a programmed exploration strategy which can be modified as needed, while the Hector UAV performs the mapping of the aerial mapping of the environment.

Fig. 5 presents the initialization of exploration through Husky in the pattern given to it. While Fig. 6 presents the aerial mapping of Hector of the environment shown in Fig. 4.
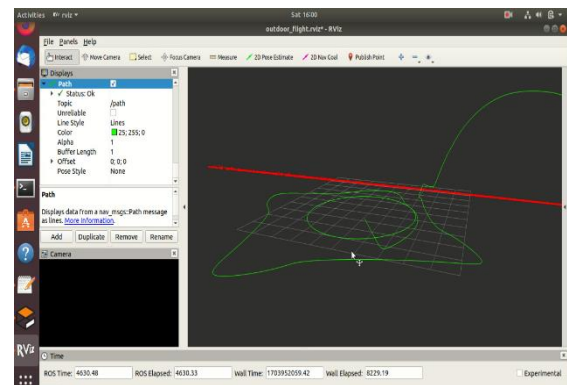
Fig. 5. Expanding Circular Exploration lead by Husky UGV, path trail in RVIZ visualizer
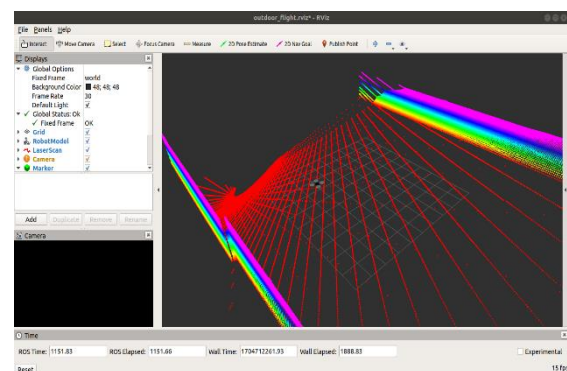
Fig. 6. Aerial Mapping being conducted through Hector UAV, LiDAR data being shown in RVIZ

5. Conclusion

The work demonstrated a successful integration of two heterogeneous autonomous unmanned systems in ROS-Gazebo environment, proving their potential to use them for reconnaissance applications in unknown workspaces. The key feature of this work was to propose an attraction-repulsion mechanism among the aerial and ground vehicles for a strategic and coordinated exploration. This work is at the early stages of experimentation and possess the potential for further enhancements which primarily include optimization of exploration algorithms and strategies. Moreover, this

work just incorporates two unmanned systems which can be expanded to exhibit a swarm intelligence behavior among UAVs or UGVs. We believe that the future of exploratory missions in unknown or hostile environments is very promising when it comes to unmanned systems.

6. Acknowledgements

This work is conducted under research group formed at University of Malaya, Kuala Lumpur in which the individuals from diverse universities and industries participate. The author gratefully acknowledges ACR Lab, Faculty of Engineering, University of Malaya and their members for providing assistance during assembling and testing.

References

1. H. Zhang, T. Yang, and Z. Su, "A formation cooperative reconnaissance strategy for multi-UGVs in partially unknown environment," Journal of the Chinese Institute of Engineers, vol. 46, no. 6, pp. 551-562, 2023/08/18 2023, doi: 10.1080/02533839.2023.2227879.
2. J. Nohel, P. Stodola, J. Zezula, P. Zahradníček, and Z. Flasar, "Area reconnaissance modeling of modular reconnaissance robotic systems," The Journal of Defense Modeling and Simulation, p. 15485129231210302, 2023, doi: 10.1177/15485129231210302.
3. A. Wang, F. Jing, X. Huang, C. Gong, and S. Xu, "Structure Keeping Control for Heterogeneous Formations Based on Consistency Theory and Graph Theory," in 2023 9th International Conference on Mechatronics and Robotics Engineering (ICMRE), 10-12 Feb. 2023 2023, pp. 139-145, doi: 10.1109/ICMRE56789.2023.10106518.
4. K. S. N. Ripon, E. Jakobsen, C. Tannum, and J. M. Montanier, "Assessing the effect of self-assembly ports in evolutionary swarm robotics," in 2016 IEEE Symposium Series on Computational Intelligence (SSCI), 6-9 Dec. 2016 2016, pp. 1-8, doi: 10.1109/SSCI.2016.7850181.
5. A. Eudes, S. Bertrand, J. Marzat, and I. Sarras, "Distributed Control for Multi-Robot Interactive Swarming Using Voronoi Partioning," Drones, vol. 7, no. 10, doi: 10.3390/drones7100598.
6. M. Aguilera, J. A. R. Linan, and L. M. Torres-Treviño, "Change of Behavior on Swarm Robotics Using Repulsion and Attraction," in 2015 Fourteenth Mexican International Conference on Artificial Intelligence (MICAI), 25-31 Oct. 2015 2015, pp. 169-172, doi: 10.1109/MICAI.2015.32.
7. R. Wiki. "About ROS." Wiki. http://wiki.ros.org/ (accessed January 21, 2024).
8. GazeboSim. "About - Gazebo Simulator." https://gazebosim.org/home (accessed January 21, 2024).
9. OpenRobotics. "About - Openrobotics." OpenRobotics. https://www.openrobotics.org/ (accessed January 21, 2024).

Authors Introduction

Mr. Anees ul Husnain

Anees is currently pursuing his Doctoral in Electrical Engineering at University of Malaya, Malaysia. He holds a master's degree in computer engineering from UET Taxila, Pakistan. He is working on autonomous path generation of UAVs to monitor fugitive emissions.

Ir. Dr. Norrima Mokhtar

Norrima Mokhtar received the Bachelor of Engineering (B. Eng) degree in Electrical Engineering from University of Malaya in 2000. She was appointed as a lecturer to serve the Department of Electrical Engineering, University of Malaya immediately after graduating with her Master of Engineering. As part of her career development, she received SLAB/SLAI scholarship to attain her Ph.D. in 2012. She is now serving Department of Electrical Engineering, University of Malaya.

Dr. Takao Ito

He is Professor of Management of Technology (MoT) in Graduate School of Engineering at Hiroshima University. He is serving concurrently as Professor of Harbin Institute of Technology (Weihai) China. He has published numerous papers in refereed journals and proceedings, particularly in the area of management science, and computer science. He has published more than eight academic books including a book on Network Organizations and Information (Japanese Edition). His current research interests include automata theory, artificial intelligence, systems control, quantitative analysis of inter-firm relationships using graph theory, and engineering approach of organizational structures using complex systems theory.

Dr. Siti Sendari

She is a lecturer at Universitas Negeri Malang, Indonesia, specializing in intelligent systems, robotics, and evolutionary computation. She holds the position of Head of Laboratory in the Department of Electrical Engineering at the Faculty of Engineering. Her work focuses on intelligent systems and related areas, contributing to the academic community.

Ir. Muhammad Farris Kyasudeen

He is associated with Universiti Teknologi MARA (UiTM) in Malaysia. He is part of the Faculty of Electrical Engineering at UiTM's branch in Pulau Pinang, where he holds the position of Ir. (Ir.) indicating his professional engineering qualification. He has also received recognition for winning third place in a competition related to Electrical Engineering at UiTM. His contributions include research on an Autonomous Ground Vehicle (COR-AGV) disinfectant system using far-UVC light exposure.

Mr. Muhammad Badri M Noor

He is associated with Ifcon Technology, as mentioned in the search results. He is linked to the University of Malaya, specifically in the Department of Computer System & Technology, Faculty of Science Computer & IT. Moreover, he has contributed to a research paper on Remote Desktop Power Management System using Single-Board Computer, showcasing his expertise in this field.

Dr. Heshalini Rajagopal

She received her PhD and Master's degree from the Department of Electrical Engineering, University of Malaya, Malaysia in 2021 and 2016, respectively. She received the B.E (Electrical) in 2013. Currently, she is an Assistant Professor in UCSI University, Kuala Lumpur, Malaysia. Her research interest includes image processing, artificial intelligence and machine learning.

Illumination Effects on Facial Expression Recognition using Empirical Mode Decomposition

Hasimah Ali, Wan Khairunizam Wan Ahmad

*Centre of Excellence Unmanned Aerial Systems (*COE-UAS*), Faculty of Electrical Engineering & Technology,
Universiti Malaysia Perlis, Arau, 02010, Malaysia*

Hariharan Muthusamy

Department of Electronics Engineering, National Institute of Technology Uttarakhand, India

Mohamed Elshaikh

*Department of Electronics Engineering & Technology (FTKEN), Universiti Malaysia Perlis, Arau, 02010,
Malaysi*a
E-mail: hasimahali@unimap.edu.my, hariharan@nituk.ac.id, khairunizam@unimap.edu.my, elshaikh@unimap.edu.my

Abstract

Facial expression recognition (FER) has been acknowledged as a significant modality that could bring facial expression into human-machine interaction and make the interaction more efficient. However, the ability of FER to operate in a fully automated and robust manner is still challenging. Illumination effects, for example, make the facial expression images always contaminated with different levels of ambient noise (such as brightness variation) in a cluttered background. Thus, this paper aims to investigate the illumination effects (brightness variations) on facial expression recognition using empirical mode decomposition reconstruction techniques. In this framework, firstly, the noisy facial expression images were simulated with the illumination effects using different brightness levels of 30%, 40%, 50%, 60%, and 70%. Then, the EMD will decompose the noisy facial expression images into a small set of intrinsic mode functions (IMF), namely IMF1, IMF2, IMF3, and residue. Based on property held by EMD, the signals are decomposed into several IMF components, each with a different time scale. Because the last several IMFs represent the majority of illumination effects, various reconstruction techniques for IMFs have been investigated at various brightness levels. Feature reduction techniques Principal component analysis (PCA) and linear discriminant analysis (LDA) have been employed to reduce the high-dimensional space of IMF features into low-dimensional IMF features. The reduced IMF reconstructions were then used as input to the k-nearest neighbour classifier to recognise the seven facial expressions. A series of experiments have been conducted on the JAFEE database using various reconstruction IMFs together with PCA plus LDA. Based on the results obtained, the reconstruction of IMF1 + IMF2 + IMF3 shows the highest accuracy in high illumination conditions, which is 99.06%.

Keywords: Illumination, Facial Expression Recognition, Empirical Mode Decomposition, PCA plus LDA.

1 Introduction

Facial expression is the most natural and important tool people have to show their emotional-state or deliver information. For decades, facial expression has been regarded as an important modality for conveying feelings and attitudes, and it has had an impact on physical communication. Specifically, facial expression is formed by contracting facial muscles [1]. Facial expression recognition (FER) is a function that humans and computers can perform by locating the face from its sources. Then, facial features will be analysed to classify the emotion [2]. A lot of information about human

behaviour can be effectively extracted from their facial expression. Facial expression recognition (FER) has been acknowledge as a significant modality that could bring facial expression into human machine interaction and makes the interaction more efficient. However, the ability of FER to operate in fully automated with robust is still challenging. Illumination effects for example, make the facial expression images always contaminated with different level of ambient noise (such as brightness variation) in clutter background. In previous work, limited study was conducted on illuminations effects for facial expression recognition. For example, [3] have combined the logarithm transform, discrete cosine

Hasimah Ali, M Hariharan, Wan Khairunizam Wan Ahmad, M Elshaikh

transform, and illumination compensation as a normalised DCT to eliminate illumination variation contained in the facial images of the JAFFE database. To recognise the facial expression, the extracted features from the pre-processed image were combined with a neural network classifier. However, analysis of the effects of illumination was not discussed. In [4], the authors utilized the independence component analysis (ICA) and locality-preserving projection to compare the effectiveness of those methods towards the illumination variations on the face recognition. They found ICA outperformed the results on YALEB database with 64 illuminations types. In [5], the authors utilized deep learning approach for emotion recognition under different pose and illumination variations. The Convolution Neural Network were constructed using 15 layers (three layers per emotion) that learn from facial expression database (CMU Multi-PIE) using input size of 32 x 32 in recognizing the five basic emotions. Although they achieved 96.55%, selecting optimum the input size considered to be computational cost. In [6] the authors utilized a deep stacked convolutional autoencoder in attempting reconstruct new input images with better illumination for facial expression recognition. The model used as pre-training in greedy layer-wise by learning the image representation and then map it to approximately reconstruction of input image. The network model learns to encode the input images and create a feature vector form relatively similar illumination, ignoring the level of luminance of the facial images. Even though the used of deep learning is emerging, however handcraft features of the image provides significant impact in understanding the behaviour of spatial details of the images. Therefore, in this study, we propose to investigate the illumination effect on facial expression recognition (FER) by using empirical mode decomposition (EMD) approach. In this framework, the original facial expression database will go through face detection in a manipulated illumination condition, and then various illumination coefficients will be analysed. The EMD will be used to decomposed the illuminated images (in our case brightness variations) into a set of intrinsic mode functions (IMFs). These IMFs will be used to further analyse the illumination effects containing in facial expression images. The significant IMFs will be reconstructed by removing the one contains illuminations. In order to reduce dimensionality of reconstructed IMF, PCA plus LDA will be adopted.

Finally, the reduced features of reconstructed IMFs will be classified using k-NN classifier to recognize the seven facial expression under the illumination effects.

2 Materials and Method

Fig. 1 shows the framework of the proposed system. It consists of five phases: facial expression image database, pre-processing, EMD-based feature extraction, IMF reconstruction, and classification. In this framework, firstly, the facial expression image was pre-processed to detect the local region of the face. Then, brightness variations were simulated on the detected face. By decomposing an image into a set of intrinsic mode functions via a sifting process on the simulated brightness of a facial expression database, the nonlinear technique, namely EMD, has been used as a feature extraction-method. To assess performance, an IMF reconstruction was performed on various IMFs. Later, feature reduction was applied to the reconstructed image to reduce its dimensionality before being subjected to the classifier. The detailed description of the process is highlighted as follows:
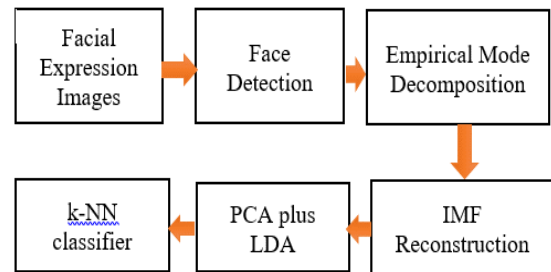
Fig. 1. Framework of the proposed method

2.1 *Facial Expression Database*

In this study, the Japanese Female Facial Expression (JAFFE) database has been used. It contains 213 grayscale facial expression images of ten subjects, all Japanese women, with seven facial expressions: neutral (30 images), angry (30 images), sad (30 images), happy (31 images), disgust (29 images), surprise (31 images), and fear (32 images).

2.2 Image Pre-Processing: Face Detection

Face detection is a technique used to detect and localize the face region from the unwanted background. The original size of JAFFE image is 256 x 256 pixel was cropped into 128 x 96 pixel that contains the face region. Fig. 2(a) and Fig. 2(b) show the example of the original and cropped image of seven facial expressions of a JAFFE subject, respectively.
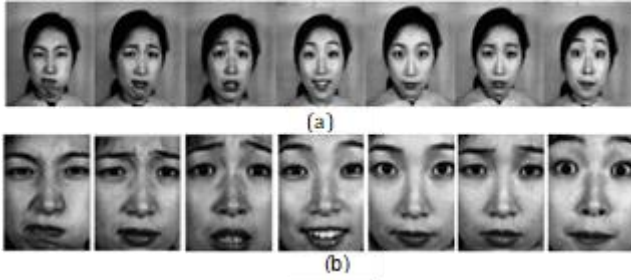
Fig. 2. Seven facial expression of JAFFE's subject: (a) original and (b) cropped image

2.3 Brightness Variation

In this study, illumination variation with different levels of brightness was used to simulate illumination effects that mimic real-world applications. The illumination variations were simulated by varying the brightness effect by controlling the intensity of each individual pixel inside the image. The percentage illumination level used in this project is 30%, 40%, 50%, 60%, and 70%. A percentage greater or less than those specified will result in either increased brightness or decreased darkness.

2.4 Feature Extraction: Empirical Mode Decomposition

Empirical mode decomposition (EMD) is a method of decomposing a natural signal without leaving the time domain. EMD decomposes any nonlinear signal into a small set of finite intrinsic mode functions (IMFs). Although EMD is focused on analysing one-dimensional signals, it can also be extended to bi-dimensional empirical mode decomposition (BEMD) to analyse two-dimensional signals while still having similar concepts and procedures [7], [8]. The main advantage of EMD is that it can extract data about local trends in the input signal by calculating and quantizing oscillations. Such oscillations can be quantized by local details and matching a local trend. The fundamental ideas of EMD are the implementation of the incorporation method such as cubic splines, sifting phase to extract the intrinsic

mode function, and numerical convergence criteria to stop the sifting phase.

In the EMD technique, there are two steps that must be met to extract the intrinsic mode functions (IMFs). To begin, the number of extremes and zero crossings must be equal, or at most one. Second, the mean of its upper and lower envelopes must not be more or less than zero. Those two steps satisfy the physically compulsory process for defining a significant instantaneous frequency. The EMD technique's entire procedure is as follows: for input signal s(t):

1. Determine the local maxima and local minima of s(t). Let $d_0(t) = s(t)$.
2. Incorporate the local maxima and the local minima to develop the upper envelope eu(t) and lower envelope el(t), correspondingly, using cubic B splines.
3. Calculate the mean of envelope m(t): $m(t) = (e_u(t) + e_l(t))/2$.
4. Determine the details using $d_1(t)=d_0(t)-m(t)$ (sifting phase).
5. Repeat steps 1-4 on the residual signal until the details signal $d_k(t)$ is included in the IMF: $c_1(t)=d_k(t)$
6. Procedures 1–5 should be repeated on the residual $rn(t)=s(t).C_n(t)$ to obtain all of the signal's IMFs ($c_1(t), c_2(t),..., c_n(t)$).

When the residual signal is either unaltered, monotone, or a function with just one extrema, the procedure is terminated. The advantage of the EMD approach is that once these IMFs are acquired, they may easily switch back and forth from them to the original signal. Similarly, the original signal may be reconstructed by combining all of the IMFs together, which can be adjusted to account for slight variations owing to the implant discovered in the method. EMD can be expressed in mathematically as:

$$X(n) = \sum_{n=1}^{N} C_n(t) + R_N(t) \qquad (1)$$

According to the technique, lower-order IMFs record rapid oscillation modes, whereas higher order IMFs record low oscillation modes. In this research, two-dimensional empirical mode decomposition (BEMD) is used to reduce illuminated face pictures into a limited collection of intrinsic mode functions (BIMFs). The BIMFs are features recovered at various scales or spatial frequencies using phase sifting.

2.5 Dimensionality reduction: PCA plus LDA

In this work, the PCA plus LDA has been employed as dimensionality reduction on the extracted IMFs features

2.5.1 PCA

Although the intrinsic mode functions effectively represent the different frequencies of oscillation for different expressions (local trends) in the image, the mode function responses can be large. A 128 x 96 image, for example, decomposes with the EMD resulting mode function into a 128 x 96 = 12,288-dimensional vector. To reduce the dimensions of the feature vector, we applied PCA to our mode function. Let consider, if $A = \{a_1, a_2, a_3 \ldots a_N\}$ is a vector of N images.

i. First, calculate the mean (μ) of the matrix, and covariance matrix (S) as:

$$S = \frac{1}{n}\sum_{i=1}^{n}(a_i - \mu)(a_i - \mu)^T \qquad (2)$$

ii. Calculate both eigenvalues (λ_i) and eigenvectors (v_i) of S as:

$$Sv_i = \lambda_i v_i, \qquad (3)$$

where i=1, 2, 3…n.

iii. Lastly, sort the eigenvectors in descending order that associated to the largest eigenvalues. The new projected space is given b:

$$y = W^T a, \qquad (4)$$

which W is represent the transformation matrix [9].

2.5.2 LDA

One of the main problem in pattern recognition is the curse of dimensionality. LDA is one of the method used to solve the issue of dimensionality by reducing the features from higher-dimensional space to lower-dimensional space. The goal of LDA is to shape the class scatter by maximizing the between-class scatter and minimising the within-class scatter [9]. In this framework, firstly we need to compute the between-class variance S_B as:

$$S_B = \sum_{i=1}^{n} N_i(\mu_i - \mu)(\mu_i - \mu)^T \qquad (5)$$

Second, within-class scatter, S_w need to be computed by $S_W = \sum_c \sum (a_k - \mu_i)(a_k - \mu_i)^T$. Finally, the lower-dimensional feature projection, W_{lda} can be determined via:

$$W_{lda} = \arg\max \left| \frac{W^T S_B W}{W^T S_W W} \right| = [w_1, w_2, w_3 \ldots w_m] \qquad (6)$$

in which w_i ($i = 1, 2, \ldots m$) is the set of S_B and S_w generalized eigenvectors based on the largest generalized eigenvalues λ_i of $S_B w_i = \lambda_i S_W w_i$.

2.5.3 PCA plus LDA

To avoid singularity of the S_w, [10] proposed PCA plus LDA. This framework involves two stages. First, original feature vector having f-dimensional feature space is mapped onto t-dimensional intermediate feature space in which ($t < f$) via PCA. Then, project the intermediate (t-dimensional) to k-dimensional feature space via LDA resulting new projected space as:

$$z_i = W_{pca}^T W_{lda}^T x_i \qquad (7)$$

where i = {1, 2, 3..N} is the number of input images.

3 Results and Discussion

In order to evaluate the effectiveness of the proposed system, the JAFFE database, which is publicly available, has been used in this experiment. The facial images obtained from successful face detection and applied with different levels of illumination will undergo feature extraction using the EMD technique in order to extract the features called intrinsic mode functions (IMFs) through sifting procedures. IMFs are extracted from high-dimensional data. As a result, feature reduction using PCA and LDA was used before being fed as input to the k-NN classifier to classify all seven emotions: anger, disgust, fear, neutral, happiness, sadness, and surprise.

3.1 Brightness Variation on Facial Expression Images

The facial images from JAFFE database are simulated with different level of illumination which is 30%, 40%, 50%, 60% and 70%. This is done by adding and subtracting the intensity of each pixel inside of the facial images. Fig. 3 shows the simulated illumination of facial expression images.
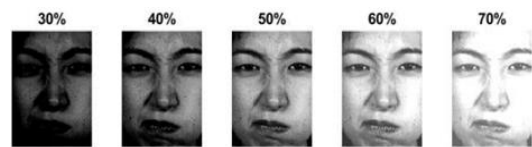
Fig. 3. Simulated illumination effects of facial expression images at different brightness level

3.2 Performance of applied EMD on expression images

The two-dimensional breakdown of a picture using the sifting process results in an interpretable display. Each IMF carries information on a certain scale that is efficiently segregated. Fig. 4 depicts the EMD decomposition of facial expression image into IMF1, IMF2, IMF3, and residual. Note that, the IMF1 contains local information about the smallest scales in the image, whereas the residue contains information about greater scales [7]. Specifically, IMF1 corresponds to the lowest time scale associated with the information's quickest time fluctuation. As the decomposition process progresses, the time scale increases, and the mean frequency of the mode decreases. As observed in Fig. 4, the set of IMF reveals the influent structure from the smallest to the largest relative to the input images. The IMF1 shows the most distinct facial features, including the boundary lines around the face region such as the mouth, nose, and eyes, which are crucial for expression recognition.
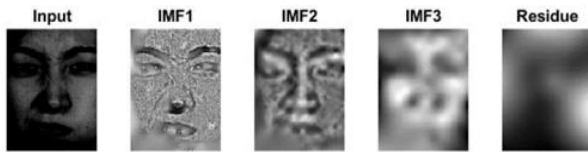
Fig. 4. EMD decomposition that associated with 30% brightness level

3.3 Performance Analysis on Facial Expression Recognition using Reconstructed Intrinsic Mode Function

According to [11], the majority of the illumination effects is represented by the last several IMFs. To evaluate how EMD can be employed to reduce the illumination effects in the proposed system. Fig. 5 shows the different combinations of the reconstructed mode functions. It should be noted that Figure 5(h) is a reconstruction of all IMF and a residue, and thus serves as a reference image. As can be seen in Fig. 5, the illumination effect was highly dominant in (b), (c), and (d). Whereas, (e), (f), and (g), the illumination effect appeared at a low degree. Meanwhile, in (a), which is IMF1+IMF2, the effects were almost gone. This observation agrees with [11] in which the several last mode functions are affected by illumination.
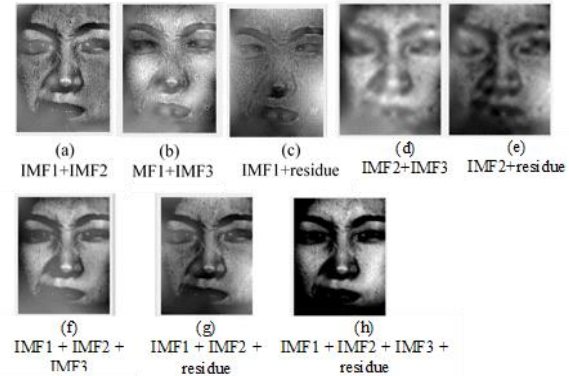
Fig. 5. Reconstruction of different mode functions (30% brightness level): (a) IMF1+IMF2, (b) IMF1+IMF3 and (c) IMF1+residue, (d) IMF2+IMF3, (e) IMF2+residue, (f) IMF1 + IMF2 + IMF3, (g) IMF1 + IMF2 + residue and (h) IMF1 + IMF2 + IMF3+residue (reference)

3.4 Performance Analysis of Applied PCA plus LDA on reconstructed IMFs

In this work, three reconstruction modes of IMFs have been investigated: IMF1+IMF2, IMF1+IMF2+IMF3, and IMF1+IMF2+IMF3+residue. Due to the higher dimensionality of feature space, the frameworks of PCA plus LDA were applied. The rationale for using PCA plus LDA is that, the feature vector space was reduced from a high-dimensional to a low-dimensional space by utilising the PCA plus LDA framework as mentioned in Section 2.5. Fig. 6 shows the distribution of reduced features due to IMF1+IMF2+IMF3+residue, whereas Fig. 7 shows the distribution of reduced features due to IMF1+IMF2+IMF3 at a 30% brightness level. Note that classes 1, 2, 3, 4, 5, 6, and 7 are denoted as *angry* (red), *disgust* (orange), *fear* (green), *happy* (cyan), *neutral* (blue), *sad* (purple), and *surprised* (pink), respectively. It can be seen in Fig. 6 that there is a low degree of overlap between the emotions of fear, anger, disgust, and neutral. Fig. 7 shows that the class emotions are well differentiated from one another.

Hasimah Ali, M Hariharan, Wan Khairunizam Wan Ahmad, M Elshaikh
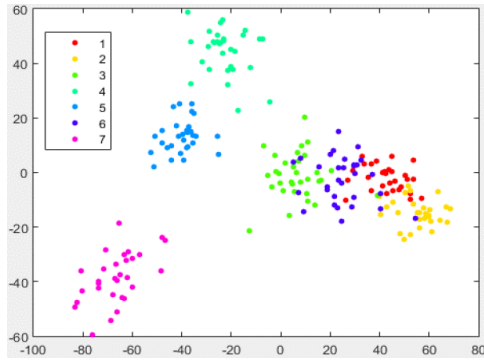
Fig. 6. Distribution of reduced features of **IMF1+IMF2+IMF3+residue** for seven facial expressions at 30% brightness level.
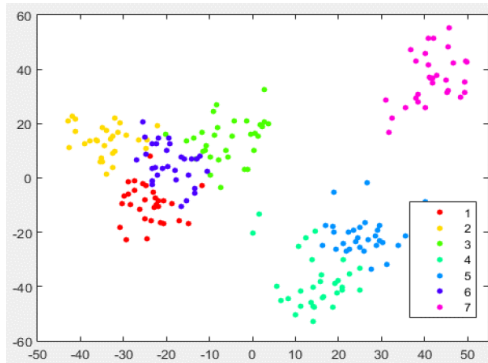
Fig. 7. Distribution of reduced features of **IMF1+IMF2+IMF3** for seven facial expressions at 30% brightness level.

3.5 Experimental Results

In this section, the reduced features using PCA plus LDA were applied to the reconstruction IMFs to recognize the seven facial expressions. The objective is to investigate the effectiveness of reconstruction IMFs in dealing with variations in illumination. Three different modes of IMF reconstruction were evaluated, namely IMF1+IMF2+IMF3+residue (used as reference), IMF1+IMF2+IMF3, and IMF1+IMF2 under different levels of illumination, which are 30%, 40%, 50%, 60%, and 70%. In this work, k-nearest neighbour was used as a classifier using a 10-fold cross validation strategy.

Fig. 8, Fig. 9, Fig. 10, Fig. 11, and Fig. 12 illustrate the recognition rate of IMF reconstruction of facial expression recognition at brightness levels of 30%, 40%, 50%, 60%, and 70%, respectively. As observed in Fig. 8, the reconstruction of IMF1+IMF2 shows the

highest recognition rate, which is 99.06% at a 30% brightness level. It seems that the combined IMF1+IMF2 is enough and would achieve similar recognition as in the reference. We agree that, the illumination effect may lie inside the IMF3 and residue. When the brightness level increased to 40% (Fig. 9), the combined IMF1+IMF2 again showed the highest recognition rate of 99.53%, which is slightly higher by 0.53% than the reference.

In Fig. 10, at a brightness level of 50%, the combined IMF1+IMF2+IMF3 shows the highest recognition of 99.53%, which is the same as the reference. Similar trends appear for brightness levels of 60% (Fig. 11) and 70% (Fig.12), where the combined IMF1+IMF2+IMF3 outperformed the results, which are 99.53% and 98.59%, respectively. This can be inferred from the fact that as the brightness level increases, the combined IMF without residue shows significant results. Thus, we agree that the majority of illumination effects may be represented by the last IMFs and residue.
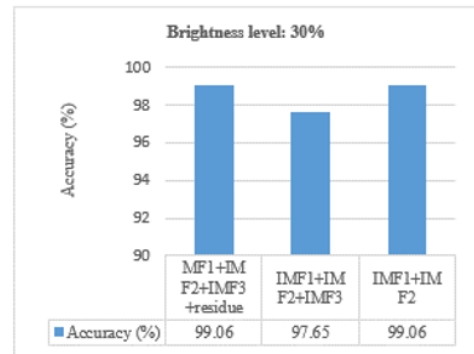
Fig. 8. Accuracy of FER based on reconstructed IMF at 30% brightness level.
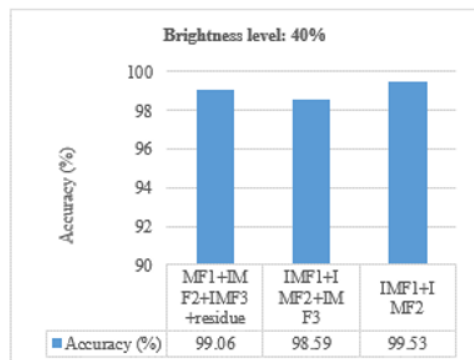
Fig. 9. Accuracy of FER based on reconstructed IMF at 40% brightness level.
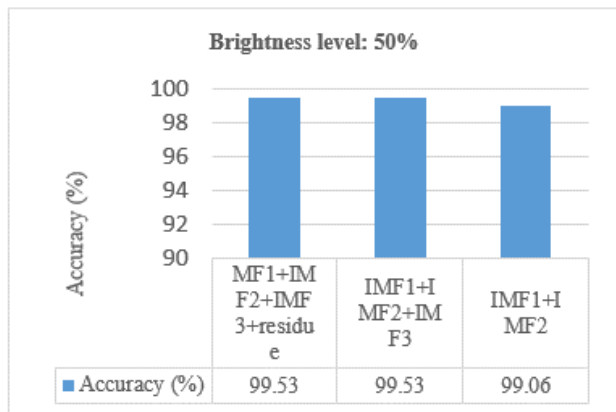
Fig. 10. Accuracy of FER based on reconstructed IMF at 50% brightness level.
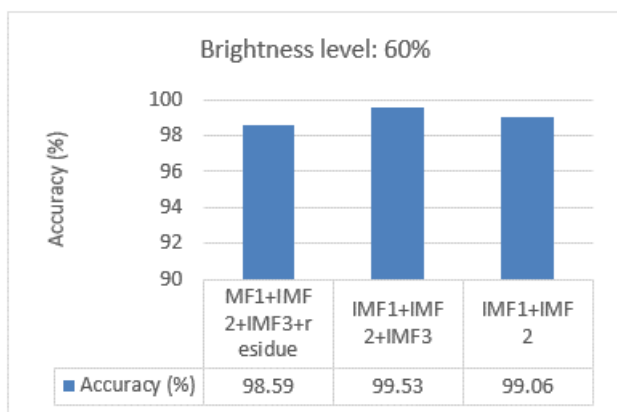
Fig. 11. Accuracy of FER based on reconstructed IMF at 60% brightness level.
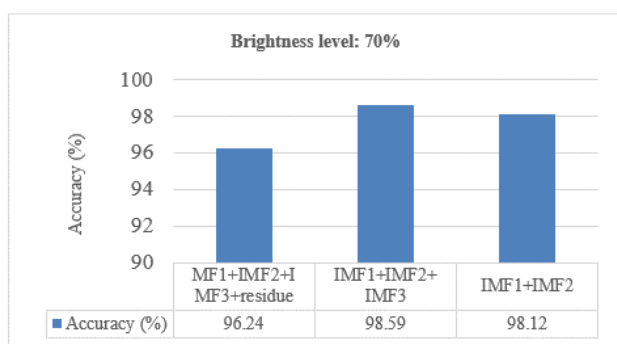
Fig. 12. Accuracy of FER based on reconstructed IMF at 70% brightness level.

4   Conclusion

This paper used an empirical mode decomposition approach to present the effects of illumination on facial expression recognition. Firstly, pre-processing of the facial image has been done via face detection to extract the local region of the face. Then, the face image was simulated to mimic the illumination effects by varying the brightness level between 30%, 40%, 50%, and 70%. Then, the simulated facial images were decomposed using empirical mode decomposition into a finite set of intrinsic mode functions. This work (IMF1, IMF2, IMF3, and residue) has been extracted via a sifting process. According to the findings, illumination effects may have been observed in the last few IMFs. Three reconstruction modes were considered for illumination analysis, which are the combined IMF1+IMF2, IMF1+IMF2+IMF3, and IMF1+IMF2+IMF3+residue (as reference). Based on the results obtained, at low brightness levels (30% and 40%), the combined IMF1+IMF2 gives the highest recognition rate of FER. However, as the brightness level was further increased to 50%, 60%, and 70%, the combined IMF1+IMF2+IMF3 gave the highest recognition rate. Therefore, we can conclude that as the brightness level increases, the combined IMF without residue shows significant results, and we agree that the majority of illumination effects may be represented by the last IMF and residue. However, more research on a larger dataset using different machine learning techniques is required to improve the robustness of FER to illumination effects.

Acknowledgment

References

1. Ali H, Hariharan M, Yaacob S, and Adom A H 2015 Facial emotion recognition using empirical mode decomposition *Expert Syst. Appl.* **42** p 1261.
2. Pantic M 2009 Facial Expression Recognition *Encycl. Biometrics* p 400.
3. Zankhana H S and Kaushik V 2015 Performance analysis of canny edge detection for illumination invariant Facial Expression recognition   International Conference on Industrial Instrumentation and Control (ICIC).
4. Fernandes S L and Bala G J 2013 Comparative study on ICA and LPP based face recognition under varying illuminations and facial expressions International

Conference on Signal Processing , Image Processing & Pattern Recognition

5. Palaniswamy S and Suchitra 2019 A robust pose & illumination invariant emotion recognition from facial images using deep learning for human-machine interface 2019 4th International Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS) p 1.

6. Ruiz-Garcia A, Palade V, Elshaw M and Almakky I 2018 Deep Learning for Illumination Invariant Facial Expression Recognition International Joint Conference on Neural Networks (IJCNN), p 1.

7. Nunes J C, Bouaoune Y, Delechelle E, Niang O, and Bunel P 2003 Image analysis by bidimensional empirical mode decomposition *Image Vis. Comput.* **21** p 1019

8. Bhuiyan S M A, Khan J F, and Adhami R R 2010 A bidimensional empirical mode decomposition method for color image processing Electrical Engineering Tuskegee *IEEE SiPS* p 272

9. Belhumeur P N, Hespanha J P and Kriegman D J 1997 Eigenfaces vs. Fisherfaces: Recognition using class special linear projection. *IEEE Trans. PAMI* **19**, p 711.

10. Pumlumchiak T. and Vittayakorn S. 2017 Facial expression recognition using local Gabor filters and PCA plus LDA 2017 9th International Conference on Information Technology and Electrical Engineering (ICITEE), pp. 1-6.

11. Qing, Jiang J, & Yang Z 2010 Empirical mode decomposition-based facial pose estimation inside video sequences *Optical Engineering*, *49*(3).

Authors Introduction

Ir. Dr. Hasimah Ali

She is senior lecturer at Faculty of Electrical Engineering Technology, UniMAP. She received her PhD in Mechatronic Engineering from UniMAP (2015), MSc in Mechatronic Engineering from IIUM (2008) and BEng in Mechatronic Engineering from IIUM (2004). Her research interests are signal and image processing, facial expression, ground penetrating radar, robotics gripper and IoT.

Prof Ir Dr. Wan Khairunizam Wan Ahmad

He received his B. Eng. degree in Electrical & Electronic Eng. from Yamaguchi University and Ph.D. in Mechatronic Eng. from Kagawa University, in 1999 and 2009 respectively. He is currently is a Prof. at Faculty of Electrical Engineering & Technology, University Malaysia Perlis and also serve as Director of Centre for Sustainable Academic Leadership Development (LEAD), UniMAP. He is member of Board of Engineer and Institute of Engineer, Malaysia. His research interest is in Human-Computer Interaction (HCI), Intelligent Transportation System, Artificial Intelligence and Robotics.

Assoc. Prof. Dr. Hariharan M.

He is currently working as an Associate Professor in the Department of Electronics Engineering, National Institute of Technology, Uttarakhand, India. He is a Senior Member of IEEE, USA, Member of IE, India and Member of IET, UK. He received Chartered Engineer status from Engineering Council, UK on February, 2016. His major research interests include Biomedical Signal and Image Processing, Speech Signal Processing, Signal Processing for Communications, AI/DL/ML Algorithms in VLSI:

Dr. Mohamed Elshaikh

He is a Senior Lecturer under Faculty of Electronic Engineering Technology (FTKEN), University Malaysia Perlis, Malaysia. He received his Ph.D. in Computer Engineering from University Malaysia Perlis, Malaysia, M.Sc Electrical and Electronics Engineering from University Technology Petronas (UTP), Malaysia and B.Sc Engineering Technology (Computer Engineering), University of Gezira, Sudan. His research mainly is on computer networking related.

Supercontinuum Generation Pump By a Molybdenum Disulfide Based Soliton Mode-Locked Fiber Laser

Aeriyn D. Ahmad
Photonics Engineering Laboratory, Department of Electrical Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia.

Norrima Mokhtar, Hamzah Arof, Sulaiman Wadi Harun
Department of Electrical Engineering, Faculty of Engineering, University of Malaya, Malaysia

Ahmad Haziq Aiman Rosol
*Department of Electronic Systems Engineering, Malaysia–Japan International Institute
of Technology, Universiti Teknologi Malaysia, Jalan Sultan Yahya Petra, 54100 Kuala Lumpur, Malaysia.*

Email: aeriyndwierni@gmail.com, norrimamokhtar@um.edu.my

Abstract

In this letter, a highly stable soliton mode-locked Erbium-doped fiber laser (EDFL) is passively obtained using a molybdenum disulfide (MoS_2) thin film as a saturable absorber (SA). The MoS_2 thin film obtained via electrochemical deposition technique is integrated into an EDFL cavity to generate mode-locked pulses operating at 1.88 MHz with a pulse duration of 3.03 ps. Supercontinuum (SC) light is generated using the proposed soliton mode pulses operating at 1560.4 nm as they are injected into a 100 m long highly nonlinear photonic crystal fiber (HN-PCF) after it is amplified to the output power of 17.8 dBm. The SC light operates in a wavelength range starting from 1360 nm to more than 1750 nm with the intensity above -35 dBm. The proposed supercontinuum laser can be seen as a promising light source for metrology and sensing applications.

Keywords: Fiber laser, supercontinuum, molybdenum disulfide, soliton.

1. Introduction

Super-continuum (SC) light sources combine the broadband attributes of lamps with the high brightness and spatial coherence of lasers. They have attracted extensive attention in recent years owing to their numerous applications in optical coherent tomography [1], optical communication [2], metrology [3] and sensing [4]. Broadband SC lasers are normally generated through a mode locked laser, which is coupled into a highly nonlinearity fiber (HNLF). The nonlinear mechanisms leading to the generation of SC include stimulated Raman scattering (SRS), four-wave mixing (FWM), self-phase modulation (SPM), cross-phase modulation (XPM), and dispersive-wave generation (DWG) [5]. Since the Raman Effect is self-phase-matched and shifts light to longer wavelength by emission of optical phonons, the SC spreads to longer wavelengths very efficiently. The short wavelength edge arises from four-wave mixing, and often the short wavelength edge is limited by increasing group velocity dispersion in the fiber.

Many works have been performed to understand the phenomenon as well as to implement the intended practical devices. To achieve a wider continuum and higher output power, mode-locked laser system that provides picosecond pulse train with high peak power must be employed as a pump source. Such systems utilizing passive saturable absorber (SA) as a mode-locker have been widely reported over the last decade.

2. Methodology

2.1. *Fabrication of MoS² SA*

Within the framework of this experimental investigation, we propose a methodology that leverages the electro-deposition technique. More precisely, the experiment is designed to employ the electrodeposition process for cathodically depositing a molybdenum sulphoselenide film. This deposition procedure will be implemented on tin oxide-coated conducting glass substrates, as well as silicon and/or metal substrates [6]. The electro-deposition

Aeriyn D. Ahmad, Norrima Mokhtar, Hamzah Arof, Sulaiman Wadi Harun, Ahmad Haziq Aiman Rosol.

technique has undergone thorough investigation, particularly in the fabrication of metallic alloy thin films. As compared to alternative methods, electrodeposition stands out for its scalability and cost-effectiveness. This method is notably advantageous due to its non-vacuum nature and its operation at room temperature [7]. A cyclic voltammetry analysis and the electrodeposition of the thin film are conducted using a three-electrode configuration. The electrolysis cell comprises an ITO-coated conductive film designated as the working electrode (WE), where the deposition of the molybdenum disulfide (MoS2) thin film takes place [8]. In contrast, a graphite rod serves as the counter electrode (CE), and a saturated calomel electrode (SCE) with an Ag/AgCl reference system functions as the reference electrode [9].

2.2. *Cavity Characterization*

In this area, we carry out an experimental demonstration of a simplified optical fiber-based supercontinuum source operating a simple mode-locked fiber laser saturating Molybdenum Disulfide (MoS^2) as a saturable absorber (SA).

The suggested supercontinuum (SC) laser system is illustrated in Fig. 1, comprising three main keys: the MoS^2-based mode-locked Erbium-doped fiber laser (EDFL), an optical amplifier, and a 100-meter-long highly nonlinear photonic crystal fiber (HN-PCF). The mode-locked EDFL utilized a 2.4-meter-long Erbium-doped fiber (EDF) pumped with a 980 nm laser diode serving as the gain medium. The SA was tested by encapsulating the newly developed MoS^2 thin film, obtained through the electrochemical deposition technique. We achieved the growth of the MoS^2 thin film on a transparent conductive indium tin oxide (ITO) film by functionalizing MoS^2 nano-flakes in the presence of monochloroacetic acid. This thin film was integrated into the EDFL cavity to function as a mode-locker.

To enhance the cavity nonlinearity and align the dispersion with the nonlinearity, a 70-meter-long standard single-mode fiber (SMF) with a group velocity dispersion (GVD) of -21.7 ps2/km was incorporated into the cavity. An isolator was employed to maintain unidirectional laser oscillation within the cavity. An 80:20 coupler was utilized to extract 20% of the output laser while retaining 80% of the light within the cavity to sustain the oscillation. The GVD values for both EDF and wavelength division multiplexer (WDM) were approximately 27.6 ps^2/km and -48.5 ps^2/km, respectively. The total cavity length amounted to 83 meters, with an anomalous net cavity dispersion of -1.59 ps^2. The output from the mode-locked laser was

subsequently amplified and introduced into the HN-PCF for generating supercontinuum light.
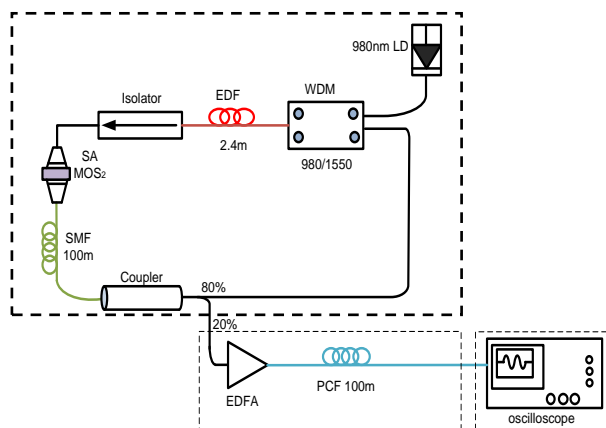
Fig 1. The SC laser system consisting of three main components namely MoS$_2$ based mode-locked laser, optical amplifier, and a HN-PCF spool.

2.3. *Mode-Locking of MoS^2*

With anomalous net cavity dispersion, fundamental mode locking in the soliton regime is achieved from the EDFL at the pump power of 76.8 mW. By adjusting the pump power up to the maximum pump power of 123.7 mW, the mode-locking state is still preserved. The output spectrum of the output pulses measured by an optical spectrum analyzer (OSA) is depicted in Fig. 2. It centered at 1560.4 nm with a 3 dB spectral width of 2.2 nm. The Kelly sidebands are clearly seen to appear at both sides of spectrum symmetrically, indicating the EDFL operates in conventional soliton mode-locking state. Fig. 3 illustrates the typical pulse train, which has uniform intensity with the interval of two pulses of 536 ns. The pulse period corresponds to the repetition rate of 1.88 MHz, which agrees with the cavity length. The radio-frequency spectrum is measured to study the stability of soliton pulse as shown in Fig. 4. The signal-to-noise ratio (SNR) at the fundamental repetition rate of 1.88 MHz is measured to be higher than 60 dB, which indicates the pulsed laser operates at high stability. Fig. 5 shows the autocorrelation curve, which was obtained by an autocorrelator to measure pulse duration. The full width at half maximum (FWHM) is 4.70 ps, which indicates that the pulse duration is 3.03 ps considering the pulse shape is Sech^2.
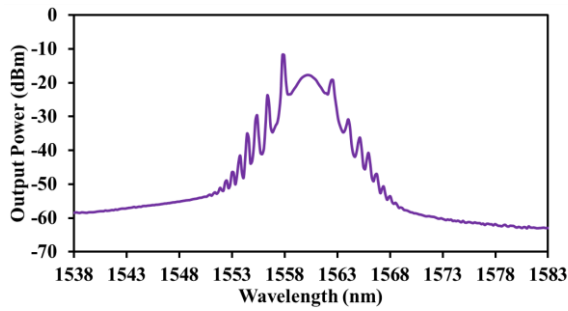
Fig 2. Output Spectrum taken from OSA
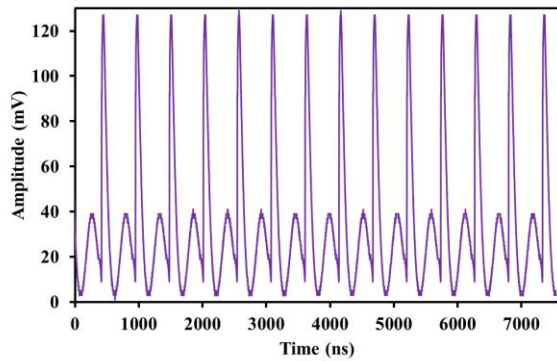
Fig 3. Output measured from oscilloscope
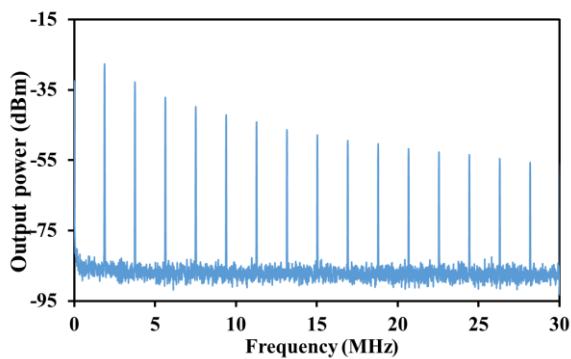
Fig 4. Radio Frequency Output at 30MHz span
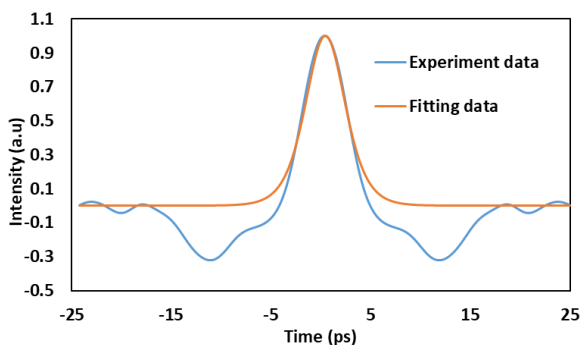
Fig 5. Autocorrelator

The average output power and pulse energy characteristics are shown in Fig. 6. As the pump power increases, both output power and pulse energy increase almost linearly. At the maximum pumping of 123.7 mW,

the output power and pulse energy are obtained at 0.79 mW and 0.42 nJ, respectively. To generate SC photons, the soliton laser is amplified by the optical amplifier and then launched into the HN-PCF. The amplifier boosts the signal up to the output power of 17.8 dBm so that it can initiate a spectral broadening in the so-called 'long pulse' regime. The HN-PCF used in the experiment has a zero dispersion at 1550 nm with á length of 100 m and nonlinearity coefficient of around $11W^{-1} km^{-1}$.
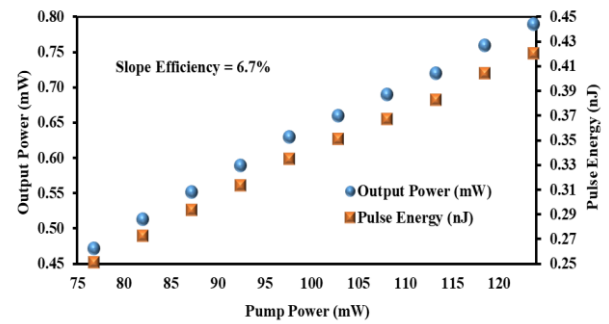
Fig. 6. The average output power and pulse energy at various launched pump power.

3. Results and Discussion

3.1. *Generation of Supercontinuum Light*

As the soliton picosecond pulses created through saturable absorption of MoS_2 are coupled into a HN-PCF, broadband SC photons are generated via various nonlinear mechanisms as shown in Fig. 7. We observe an SC light spanning from 1360 nm up to more than 1750 nm with output power of more than -35 dBm. As the amplified 1560.4 nm soliton pulse laser is injected into the PCF, the SC spreads to longer wavelengths very efficiently due to Raman effect, which is self-phase-matched to allow a shifting of light to longer wavelength by emission of optical phonons. The short wavelength edge arises from parametric four-wave mixing, which breaks up the higher-order solitons to produce frequencies at wavelengths shorter than the zero-dispersion wavelength.

The SC bandwidth can be further expanded by increasing the peak power of the injected pulses. This could be realized by improving the average power as well as compressing the pulse duration of the mode-locked laser. We expect a smaller pulse duration with the improvement of modulation depth and non-saturable loss of the MoS_2 SA. The use of an amplifier with higher saturated output power is also expected to further expand the bandwidth as well as improving the flatness and power of SC light.

Aeriyn D. Ahmad, Norrima Mokhtar, Hamzah Arof, Sulaiman Wadi Harun, Ahmad Haziq Aiman Rosol.
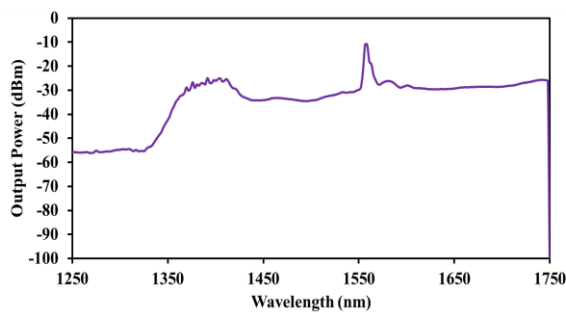
Fig. 7. SC generation due to the injection of the amplified soliton pulses into 100 m long HN-PCF.

From the broader bandwidth SC output, the potential application can be implemented are pulse generation, pulse amplification, pulse compression, metrology, spectroscopy, imaging, telecom and on-chip integration [10].

4. Conclusion

In conclusion, our research endeavors have yielded a successful demonstration of broad-spectrum supercontinuum (SC) light generation through the implementation of a newly developed mode-locked Erbium-doped fiber laser (EDFL) utilizing Molybdenum Disulfide (MoS^2) as a saturable absorber. The MoS2 thin film, synthesized via the electrochemical deposition technique, was seamlessly integrated into the EDFL cavity. This integration facilitated the generation of mode-locked soliton pulses operating at 1560.4 nm, characterized by a repetition rate of 1.88 MHz and a pulse duration of 3.03 picoseconds. The resultant pulses were subsequently amplified to an output power of 17.8 dBm and introduced into a 100-meter-long highly nonlinear photonic crystal fiber (HN-PCF) to induce SC light. The generated supercontinuum light spans a wavelength range from 1360 nm to beyond 1750 nm. This achievement holds significance in the context of advancing optical sources for applications in diverse fields such as telecommunications, metrology, and sensing.

References

1. Hartl, I., et al., Ultrahigh-resolution optical coherence tomography using continuum generation in an air–silica microstructure optical fiber. Optics letters, 2001. 26(9): p. 608-610.
2. Mori, K., et al., Supercontinuum lightwave source generating 50 GHz spaced optical ITU grid seamlessly over S-, C-and L-bands. Electronics Letters, 2003. 39(6): p. 544-546.
3. Woodward, J.T., et al., Supercontinuum sources for metrology. Metrologia, 2009. 46(4): p. S277.
4. Eslami Jahromi, K., et al., A broadband mid-infrared trace gas sensor using supercontinuum light source: applications for real-time quality control for fruit storage. Sensors, 2019. 19(10): p. 2334.
5. Dudley, J.M. and G. Genty, Supercontinuum light. Physics today, 2013. 66(7): p. 29-34.
6. Li, X. and H. Zhu, Two-dimensional MoS2: Properties, preparation, and applications. Journal of Materiomics, 2015. 1(1): p. 33-44.
7. Thomas, N., et al., 2D MoS2: structure, mechanisms, and photocatalytic applications. Materials Today Sustainability, 2021. 13.
8. Hossain, M.A., et al., Electrochemical deposition of bulk MoS2 thin films for photovoltaic applications. Solar Energy Materials and Solar Cells, 2018. 186: p. 165-174.
9. Lee Lian Hong, J., et al., Ultrafast Erbium-doped Fiber Laser Using Electrodeposition Coated MoS2 Thin Film As Saturable Absorber. Materials Research Express, 2023.
10. Bres, C.S., et al., Supercontinuum in integrated photonics: generation, applications, challenges, and perspectives. Nanophotonics, 2023. 12(7): p. 1199-1244.

Authors Introduction

Mrs. Aeriyn D. Ahmad

She received her Bachelor's degree in Electronic Engineering in 2017 from the Faculty of Engineering Technology, Universiti Teknikal Malaysia Melaka (UTeM) in Malaysia. She was again sponsored by Jabatan Perkhidmatan Awam (JPA) under Program Pelajar Cemerlang (PPC) by Malaysia Government for her Masters of Engineering Telecommunication. She is currently a PhD student in University of Malaya under Universiti Malaya Scholarship Scheme (UMSS).

Dr. Norrima Mokhtar

She received the Bachelor of Engineering (B. Eng) degree in Electrical Engineering from University of Malaya in 2000. She was appointed as a lecturer to serve the Department of Electrical Engineering, University of Malaya immediately after graduating with her Master of Engineering. As part of her career development, she received SLAB/SLAI scholarship to attain her Ph.D. in 2012. She is now serving Department of Electrical Engineering, University of Malaya.

Prof Dr Hamzah Arof

He received the B.Sc degree in Digital Signal Processing & Image Processing from Michigan State University, USA and Ph.D in same area at University of Wales, Cardiff UK. His area of expertise are time series economics, image processing, robotics, biomedical signals and photonics.

Prof Dr Sulaiman Wadi Harun

He received the B.E degree in Electrical and Electronics System Engineering from Nagaoka University of Technology, Japan in 1996, and M.Sc. and Ph.D. degrees in Photonics Technology from University of Malaya in 2001 and 2004, respectively. He has nearly 20 years of research experiences on the development of optical fiber devices including fiber amplifiers, fiber lasers, and fiber optic sensors. He also involved in exploiting new nanomaterials such as graphene, carbon nanotubes, black phosphorous, topological insulators for various fiber lasers and sensors applications. He has received about 10 research grants of value over RM4M from the Ministry of Education and the Ministry of Science, Technology, and Innovation. Prof. Harun has published more than 700 articles in ISI journals and his papers have been cited more than 7000 times with an h-index of 37, showing the impact on the community. Prof. Harun also worked as an adjunct Professor at Airlangga University, Indonesia and Ton Duc Thang University, Vietnam. He received a prestigious award of Malaysian Rising Star 2016 from the Ministry of Higher Education for his contribution in international collaboration. He is the Fellow of Malaysian Academic of Science, and also the founder and honorary advisor for the Optical Society of Malaysia.

Dr. Ahmad Haziq Aiman Rosol

He holds a Bachelor of Engineering (Hons.) in Electronic Communication from Universiti Teknologi MARA (UiTM) Shah Alam (2015), a Master of Science in Electrical Engineering from UiTM (2017), and a Ph.D. in Photonics from Universiti Malaya (2021). He is currently a Postdoctoral Fellow at the Malaysia-Japan International Institute of Technology (MJIIT), Universiti Teknologi Malaysia Kuala Lumpur, where he conducts research on innovative photonic technologies. His expertise lies in the areas of pulse fiber laser and nonlinear optics.

Adsorption Behavior of Arsenic and Selenium using NiZn Hydroxy Double Salts with Acetate, Chloride, Nitrate and Sulfate Anions

Kaoru Ohe*, Ryosuke Tabuchi, Tatsuya Oshima

Faculty of Engineering, University of Miyazaki, 1-1 Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan

*Corresponding author

Abstract

Arsenic(As) and selenium(Se) has become an increasingly serious water contamination worldwide, so the development of adsorbents to improve the adsorption performance of As and Se oxyanions is desired. In this study, NiZn adsorbents intercalated acetate, chloride, nitrate, and sulfate anions to interlayers of NiZn hydroxy double salts were prepared and their adsorption behavior of As and Se oxyanions was investigated. The Langmuir isotherm model, characteristic of monolayer adsorption, fit the experimental data well with $R^2 > 0.94$. The adsorption capacity(q_{max}) of Se(IV) was the other NiZn-AcO $\approx$ NiZn-Cl > NiZn-NO$_3$ > NiZn-SO$_4$ and that of Se(VI) was the order of NiZn-Cl > NiZn-AcO > NiZn-NO$_3$ > NiZn-SO$_4$. The q_{max} of as(V) on NiZn-Cl was about twice as large as that on NiZn-AcO, NiZn-NO$_3$ and NiZn-SO$_4$. NiZn-Cl was found to adsorb As(V) and Se(VI) oxyanions more efficient than NiZn-AcO.

Keywords: NiZn hydroxy double salts, arsenic, selenium, adsorption, Langmuir isotherm model

1. Introduction

Arsenic and selenium are known to be harmful to human body even at low concentration. The environmental standard values for arsenic [1] and selenium [2] in drinking water are set by the WHO guidelines as less than 0.01 mg/dm^3. Arsenic in nature has oxidation states of -3, 0, +3, and +5, and exists as inorganic and organic arsenic. Inorganic arsenic is known to more toxic than organic arsenic. Inorganic arsenic in natural water forms mainly oxyanions, arsenite (AsO_3^{3-}, As(III)) and arsenate (AsO_4^{3-}, As(V)). The acid dissociation constants of As(III) are pK_{a1}=9.23, pK_{a2}=12.13, and pK_{a3}=13.4, and there are four chemical forms, H_3AsO_3, $H_2AsO_3^-$, $HAsO_3^{2-}$, and AsO_3^{3-}[3]. The acid dissociation constants of As(V) are pK_{a1}=2.24, pK_{a2}=6.96, and pK_{a3}=11.5, and the chemical species are H_3AsO_4, $H_2AsO_4^-$, $HAsO_4^{2-}$, and AsO_4^{3-} [3]. Selenium solubilized and released into waste water exists in +4 and +6 oxidation states. Chemical species of selenium with these oxidation states are oxyanions selenite (SeO_3^{2-}, Se(IV)) and selenate (SeO_4^{2-}, Se(VI)) with high solubility and mobility in both aquatic and soil environment.. The acid dissociation constants of Se(IV) are pK_{a1}=2.62, pK_{a2}=8.23, and there are three chemical forms, H_2SeO_3, $HSeO_3^-$, and SeO_3^{2-} [4]. The acid dissociation constants of Se(VI) are pK_{a1}=1.7 and the chemical species are $HSeO_4^-$, and SeO_4^{2-} [4].

Hydroxyl double salts HDSs, one of inorganic anion-exchanger with a general formula $[M_A^{II}_{(1-x)}M_B^{II}_{2x}(OH)_2(Y^{z-})]_{2x/z}.nH_2O$. Specifically, the NiZn double basic salt with acetate anion (NiZn-AcO) Ni_{1-} $_xZn_{2x}(OH)_2(OCOCH_3)_{2x}.nH_2O$ has a structure in which 1/4 of Ni^{2+} in brucite-like layer of $Ni(OH)_2$ is removed from the brucite-like layer and two Zn^{2+} are located outside the layers, just above and below the Ni^{2+} vacancy [5], [6]. The layers have positive charges in excess, which are balanced by intercalation of acetate anion. The acetate anions bind directly to the tetrahedral sites along with three OH$^-$ groups from the layer to satisfy the coordination number of Zn^{2+}. The acetate anions are exchanged with guest ions in aqueous solution. Most of research on HDS is in the field of catalysis [7] and, to my knowledge, there are few applications in the aqueous environment. We reported that NiZn-AcO is an effective adsorbent for arsenic removal than MgAl Layered double hydroxide [8]. Kozai *et al.* showed that NiZn hydroxyl double salts is more effective than LDH for selenium removal using selective coefficients [9], but adsorption isotherms and adsorbent evaluation after adsorption selenium were not studied.

In this study, NiZn adsorbents with chloride, nitrate and sulfate anions in the interlayer of NiZn were prepared to improve for arsenic and selenium removal performance and their adsorption behavior for arsenic and selenium was evaluated.

2. Experimental

Kaoru Ohe, Ryosuke Tabuchi, Tatsuya Oshima

2.1. *Preparation of NiZn adsorbents*

NiZn-AcO, NiZn hydroxide double salt intercalating acetate anion (AcO) was first prepared by modifying the synthetic method reported by Rojas *et al*. [5]. NiZn-Cl adsorbent was prepared by an ion-exchange method. One gram of NiZn-AcO was added to 50 cm^3 of 1 M NaCl solution and stirred at 333 K for 24h. The supernatant was then wasted and fresh 1 M NaCl solution was added and stirred. This procedure was repeated three times. NiZn-NO$_3$ and NiZn-SO$_4$ were prepared in the same procedure as NiZn-Cl using 1 M Na NO$_3$ and 1 M Na$_2$SO$_4$, respectively. The resulting adsorbents were washed several times with deionized water, dried in a vacuum drying oven and characterized by XRD, FT-IR, and CHN elemental analysis.

2.2. *Adsorption test of arsenic and selenium*

All adsorption test was carried out by batch method at 303 K. 1.0×10^{-4} M metal solution was adjusted to a given pH using HCl solution or an aqueous NaOH solution in the pH-dependent adsorption test. In the adsorption isotherm experiments, metal concentrations were prepared at 2.0×10^{-4} - 4.0×10^{-3} M and adjusted to the prescribed pH with the aqueous NaOH solution. Ten mg of NiZn adsorbent and 15 cm^3 of metal solution was added to a sample tube and shaken in at 303 K, 120 rpm for 24 h. pH of the filtrate was measured. The metal concentrations of the solutions before and after adsorption equilibrium were measured using atomic absorption spectrophotometer and ICP emission spectrometer. The amount of adsorption(q) and the adsorption percentage ($A\%$) were calculated according to the following equations, respectively:

$$q = (C_{int} - C_{eq})\ v/w \qquad (1)$$
$$A = (C_{int} - C_{eq})\ /\ C_{int} \qquad (2)$$

where C_{int} and C_{eq} are initial and equilibrium concentration, respectively [mM], v is the volume of solution , w is the weight of adsorbent [g], q is the amount of arsenic or selenium adsorbed , A is the adsorption percentage [%]

Langmuir equation and its variant equation are given in Eq.(3) and (4), respectively.

$$q = K_L q_{max}\ C_{eq}\ /\ (1 + K_L C_{eq}) \qquad (3)$$
$$C_{eq}\ /\ q = C_{eq}\ /\ q_{max} + 1\ /\ K_L\ q_{max} \qquad (4)$$

where q_{max} is the maximum adsorption capacity [mmol/g] and K_L is the adsorption equilibrium constant [dm^3/mmol]

3. Results and Discussion

3.1. *Characterization of NiZn adsorbents*

NiZn adsorbents prepared were identified by XRD and FT-IR analysis. XRD diffraction pattern of NiZn-AcO is shown in Fig. 1. The diffraction peaks of 001, 002 and 003 are at $2\theta = 6.8$, 13.6, 20.6, respectively, in agreement with the literature values [5]. The interlayer distance of
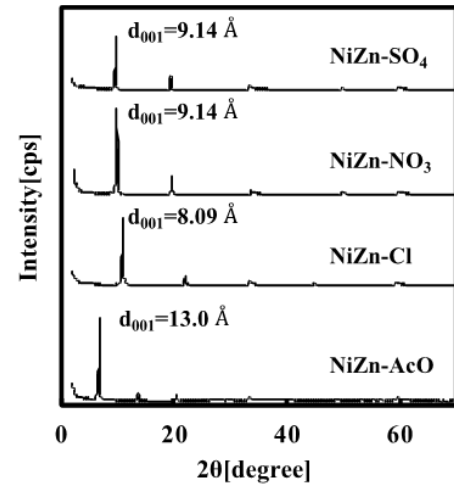
Fig. 1 XRD patterns of NiZn-AcO, NiZn-Cl, NIZn-NO$_3$ and NiZn-SO$_4$..

NiZn-AcO was determined to be 8.40 Å by subtracting the thickness of the basic layer, Ni(OH)$_2$, of 4.60 Å. Similarly, peaks derived from layered compounds were observed in the diffraction patterns of NiZn-Cl, NiZn-NO$_3$, and NiZn-SO$_4$, with d$_{001}$ values of 8.09 Å, 9.14 Å, and 9.14 Å, respectively, as shown in Fig. 1 [10].

Chemical formula of NiZn adsorbents was determined using atomic absorption analysis (Ni/Zn mole ratio), CNH elemental analysis (amount of acetate) and TG/DTA (amount of adsorption water) [11]. Table 1 shows the chemical formulas of the NiZn adsorbent, in which acetate anion in NiZn-AcO is ion-exchanged with Cl$^-$, NO$_3^-$ and SO$_4^{2-}$. The anion exchange capacity of NiZn-AcO is 2.45 meq/g. The amount of Cl$^-$, NO$_3^-$ and SO$_4^{2-}$exchanged were in the order of Cl$^-$ (2.22 meq/g) > SO$_4^{2-}$ (2.09 meq/g) > NO$_3^-$ (1.57 meq/g).

Table 1. Chemical formula of NiZn adsorbents

NiZn adsorbents	Chemical formula
AcO	Ni$_{0.67}$Zn$_{0.34}$(OH)$_2$(CH$_3$COO)$_{0.34}$ 0.65H$_2$O
Cl	Ni$_{0.67}$Zn$_{0.34}$(OH)$_2$(CH$_3$COO)$_{0.03}$(Cl)$_{0.31}$ nH$_2$O
NO$_3$	Ni$_{0.67}$Zn$_{0.34}$(OH)$_2$(CH$_3$COO)$_{0.12}$ (NO$_3$)$_{0.22}$ nH$_2$O
SO$_4$	Ni$_{0.67}$Zn$_{0.34}$(OH)$_2$(CH$_3$COO)$_{0.05}$ (SO$_4$)$_{0.15}$ nH$_2$O

3.2. *Adsorption isotherms of arsenic and selenium*

To evaluate the adsorption performance of NiZn adsorbents for As and Se oxyanions, adsorption isotherms of e and As oxyanions at 303 K were conducted. Langmuir isotherm model, characteristic of monolayer adsorption, since the amount increases with increasing concentration and approaches a constant value at high concentrations. The Experimental data for the adsorption of Se and As oxyanions using NiZn adsorbents fit Langmuir plot well with R^2 >0.94 (Fig. 2). The maximum adsorption capacity (q_{max}) and adsorption equilibrium constant (K_L) for Se(IV), Se(VI) and As(V) were

calculated from the Langmuir equation (Eq. (3) and (4)) and are shown in Table 2.
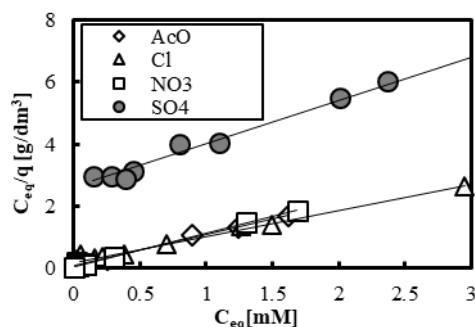
Fig. 2 Langmuir plots of Se(VI) on NiZn-AcO, -Cl, -NO$_3$ and -SO$_4$...

Table 2. Parameters of the Langmuir adsorption isotherms for Se(IV), Se(VI) and As(V)at 303 K.

Adsorbates	Adsorbents	pH$_{eq}$	q_{max} [mmol/g]	K_L [dm^3/mmol]	R^2
Se(IV)	NiZn-AcO	8.54±0.33	1.48	7.10	0.94
	NiZn-Cl	7.88±0.39	1.44	9.98	0.994
	NiZn-NO$_3$	7.67±0.59	1.54	7.47	0.971
	NiZn-SO$_4$	7.35±0.39	0.909	4.09	0.965
Se(VI)	NiZn-AcO	6.85±0.15	0.979	14.2	0.989
	NiZn-Cl	6.72±0.54	1.19	4.31	0.988
	NiZn-NO$_3$	6.60±0.42	0.983	12.8	0.997
	NiZn-SO$_4$	6.77±0.21	0.718	0.531	0.99
As(V)	NiZn-AcO	10.3±0.32	0.446	287	0.998
	NiZn-Cl	9.67±0.43	0.947	14.3	0.997
	NiZn-NO$_3$	10.0±0.36	0.524	77.6	0.999
	NiZn-SO$_4$	9.59±0.49	0.564	11.1	0.988

The maximum adsorption capacity (q_{max}) of Se(IV), Se(VI), and As(V) by NiZn-Cl, -NiZnNO$_3$, and SO$_4$ was almost the same or increased compared to that of NiZn-AcO, except for NiZn-SO$_4$. Since the dissolved species of Se(VI) or As(V) have a significant effect on qmax, the saturated adsorbed amount of Se(VI) or As(V) was converted to an equivalent unit by considering the mole fractions of the main dissolved species at each pH, which were obtained from the acid dissociation constants. As shown in Table 2, Se(VI) species are SeO$_4^{2-}$ from pH, the qmax of NiZn-AcO and NiZn-Cl are calculated to be 1.96 meq/g and 2.38 meq/g, respectively. Acetate in NiZn-AcO and NiZn-Cl The Se(VI) exchange percentage was estimated to be 80.0% and 107% by comparing the amount of acetate anion and Cl- (meq/g) in NiZn-AcO and NiZn-Cl, respectively, as shown in section 3.1.

When Se(IV) was examined similarly, the exchange rate of Se(IV) by NiZn-Cl (82.7%) was slightly lower than that of NiZn-AcO (96.7%) From the pH given in Table 2, the qmax of NiZn-AcO and NiZn-Cl were calculated to be 0.918 meq/g and 1.91 meq/g, respectively, considering the mole fraction of As(V) chemical species, and the

exchange percentage of As(V) was estimated to be 37.5% and 85.9%. These results indicate that NiZn-Cl can adsorb Se(VI) and As(V) more effectively than NiZn-AcO.

3.3. *Effect of contact time of Se(VI) adsorption*

The effect of contact time on Se(VI) adsorption by NiZn-AcO and NiZn-Cl is shown in Fig. 3. The time to reach equilibrium for Se(VI) adsorptions was 5 min which was extremely rapid. During the first 5 min 95.2% of Se(VI) was adsorbed on NiZn-AcO, whereas NiZn-Cl adsorbed 92.7% of Se(VI). These fast adsorption process has been considered ion exchange reactions involving, which is similar to ion exchange in LDHs [12].
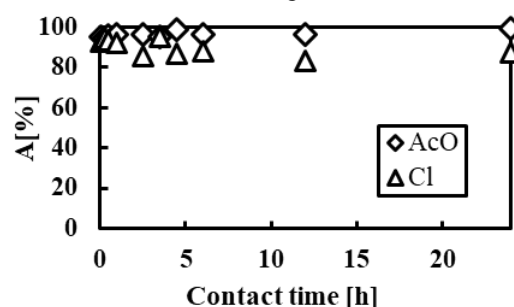
Fig. 3 Time course of Se(VI) adsorption on NiZn-AcO and NiZn-Cl...

3.4. *XRD results of NiZn adsorbent after Se(IV) and Se(VI) adsorption*

The basic brookite layer thickness does not change much during the process of anion exchange, but the interlayer distance varies with the placement and size of the intercalated anions. When inorganic anions are replaced by organic anions in the LDH interlayer, the LDH layer space increases significantly [12]. In the present study, the (003) diffraction peaks observed in
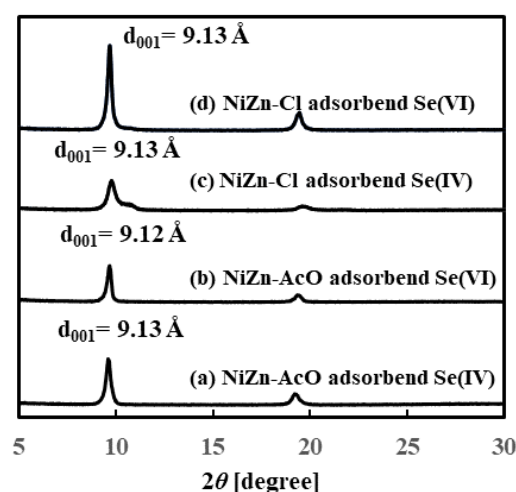
Fig. 4 XRD patterns of NiZn-AcO adsorbed (a) Se(IV) and (b) Se(VI)...and NiZn-Cl adsorbed (c) Se(IV) and (d) Se(VI)..

Kaoru Ohe, Ryosuke Tabuchi, Tatsuya Oshima

NiZn-AcO and NiZn-Cl have the d_{003} values of 13.0 and 8.09 Å, respectively (Fig. 1), and these peaks shifted after the adsorption of Se(IV) or Se(VI) (Fig. 4). and NiZn-Cl host layer thickness is 4.6 Å, the NiZn-AcO interlayer distance decreased significantly from 8.4 Å to 4.63 Å (Se(IV)), 4.52 Å (Se(VI)), and the NiZn-Cl interlayer distance increased from 4.49 to 4.53 Å (Se(IV), Se(VI)). Se(IV), Se(VI)). The interlayer distances after selenium adsorption were almost equal, suggesting that the interlayer acetate anions and Cl⁻ were replaced by Se(IV) or Se(VI).

4. Conclusions

NiZn adsorbents NiZn-Cl, NiZn-NO$_3$, and NiZn-SO$_4$ were prepared by ion exchange method using NiZn-AcO obtained by hydrothermal synthesis. Experimental investigating the adsorption of arsenic and selenium oxyanions from aqueous solutions using the NiZn adsorbents indicated that the arsenic and selenium oxyanions could be adsorbed on NiZn-AcO, NiZn-Cl, NiZn-NO$_3$, and NiZn-SO$_4$. Se(IV), Se(VI) and As(V) were found to be adsorbed on NiZn adsorbent by Langmuir isotherm model. The adsorption isotherm for As(V) shows that NiZn-Cl has twice higher adsorption capacity than NiZn-AcO. The adsorption of Se(IV) by NiZn-NO$_3$ was slightly increased compared to NiZn-AcO. The exchange of acetate anion in the NiZn interlayer with chloride ions improved the adsorption performance of As(V) and Se(VI).

References

1. B. K. Mandal and K. T. Suzuki, Arsenic round the world : A review, Talanta, 58(1) , 2002, pp.201-235.
2. S. O. Okonji, G. Achari and D. Pernitsky, Environmetanl inpacts of selenium contamination: A review on current-issues and remediation strategies in an aqueous system, *Water*, 13(11), 2021, 1473
3. D. Mohan and C. U. Pittman Jr., Arsenic removal from water/wastewater using adsorbents-A critical review, J. Hazardous Metals, 142(1-2), 2007, pp. 1-53.
4. S. Santos, G. Ungureanu, R. Boaventura and C. Botelho, Selenium contaminated waters: Am overview of analytical methods, treatment options and recent advances in sorption methods, Sci. Total Environ., 512-522, 2015, pp.246-260.
5. R. Rojas, M. A. Ulibarri, C. Barriga and V. Rives, Chromate intercalation in Ni-Zn layered hydroxide salts, Appl. Clay Sci., 49(3), 2010, pp.176-181.
6. R. Rojas, M. A. Ulibarri, C. Barriga and V. Rives, Intercalation of metal-edta complexes in Ni-Zn layered hydroxysalts and study of their thermal stability, Microporous and Mesoprous Mater., 112, 2008, pp. 262-272.
7. T. Hara, M. Ishikawa, J. Sawada, N. Ichikuni and S. Shimazu, Creation of Highly Stable Monomeric Pd(II) Species in an Anion-Exchangeable Hydroxyl Double Salt Interlayer: Application to Aerobic Alcohol Oxidation Under an Air Atmosphere, Green Chem., **11**, 2009, pp.2034-2040.
8. K. Ohe, R. Tabuchi, T. Oshima, T. Hara and S. Shimazu, Adsorptive removal of arsenic (III) and arsenic (V) from aqueous solution using nickel-zinc hydroxyle double salts, Kagaku gogaku Rombunshu, 45(2), 2019, pp.80-85.
9. N. Kozai, T. Ohnuki and S. Komarneni, Selenium Oxyanions: Highly Selective Uptake by a Novel Anion Exchanger, J. Mater. Res., 17, 2002, pp.2993-2996.
10. R. Rojas, C. Barriga, M. A. Ulibarri, P. Maletb and V. Rive, Layered Ni(II)-Zn(II) hydroxyacetates. Anion exchange and thermal decomposition of the hydroxysalts obtained, J. Mater. Chem, 12, 2002, pp.1071-1078.
11. N. Kozai, H. Mitamura, H. Fukuyama, F. Esaka and S. Komarneni, Synthesis and characterization of nickel-copper hydroxide acetate, NiCu(OH)$_3$.1(OCOCH$_3$)$_{0.9}$·0.9 H$_2$O, Microporous and Mesoprous Mater., 89, 2006, pp.123-131.
12. T. You, G. F. Bance and H. Zhao, Selenium adsorption on Mg-Al and Zn-Al layered double hydroxides, App. Clay Sci.20, 2001, pp.13-25.

Authors Introduction

Dr. Kaoru Ohe

She received her Ph.D. degrees from University of Miyazaki, Japan, in 2014. Currently she is an Associate Professor of Center for Science and Engineering Education, Faculty of Engineering. Her researches are separation engineering especially adsorption hazardous heavy metals and oxyanions.

Mr. Ryosuke Tabuchi

He received his Master's degree in Engineering in 2019 from Faculty of Engineering, University of Miyazaki in Japan.

Prof. Dr. Tatsuya Oshima

He received his Ph.D. degrees from Kyushu University, Japan, in 2001. Currently he is a Professor of the Department of Applied Chemistry, Faculty of Engineering. His researches are separation engineering (especially solvent extraction for precious metal ions) and formulation techniques for poorly water-soluble drugs and nutrients.

Parallel acoustic analysis based on the domain decomposition method with higher-order element

Amane Takei*, Makoto Sakamoto
Faculty of Engineering, University of Miyazaki, Japan

Akihiro Kudo
National Institute of Technology, Tomakomai college, Japan

*Corresponding author E-mail: takei@cc.miyazaki-u.ac.jp

Abstract

Large-scale analyses, using numerical models with over 10 trillion elements, are required for the analysis of a large space such as a concert hall with higher-frequency bands. Large spaces are often limited to low-frequency analysis. In this study, the number of elements is reduced by wave acoustic analysis using higher-order elements. Based on the results using higher-order elements, it is shown that it is possible to analyze a real environment model such as a live music club and a concert hall.

Keywords: large-scale simulation, acoustical sound field, higher-order element

1. Introduction

Estimation of the sound field is important for improving the quality of acoustic spaces such as concert halls and live music clubs [1]. Scale model experiments and the computer simulations are used for estimation of the sound field. Scale model experiments are used in many fields [2], however, creating models requires a lot of time and is expensive. On the other hand, computer simulation creates a model and sound in a virtual space. Therefore, in the computer simulation, it is easy to change the conditions of the analysis as well as to change materials and shapes. However, a large-scale analysis, using a numerical model with over 10 trillion elements, is required for the analysis of a large space with high frequency. n a large-scale finite element steady-state acoustic analysis, the iterative domain decom-position method [3] is proposed and applied as a parallelization technique. It is shown that the large-scale analysis becomes possible by the iterative domain decomposition method [4]. In this study, higher-order elements are introduced into a parallel finite ele-ment steady-state acoustic analysis method and greatly reduce the number of elements. Higher-order elements are not actively used because the matrix expands. In particular, there are some examples of higher-order elements, e.g., higher than the 3rd order ele-ment [5], however, there are few examples of acoustic analysis. As far as we know, there is no example showing the superiority of reducing the number of necessary elements by applying higher-order elements, especially in large-scale acoustic calculations using the domain decomposition method.

2. Finite element

2.1. Higher-order elements

To reduce the number of elements, 2nd and 3rd order elements are introduced. With higher-order elements, the number of nodes increases because the nodes are placed on the sides and faces of the element. Fig. 1 shows the nodal arrangement of the tetrahedras with 1st, 2nd, and 3rd order elements. Table 1 shows the shape function of each element [6].
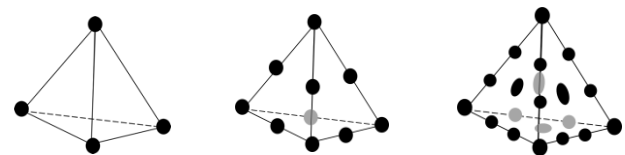
(a) 1st order elemenet   (b) 2nd order   (c) 3rd order
Fig. 1. Nodal arrangement of each element

Table 1. Shape function

1st order elm.	$N_i = L_i$	$i=0,1,...,3$
2nd order elm.	$N_i = L_i(2L_i - 1)$	$i=0, 1, ..., 3$
	$N_i = 4L_j L_k$	$i=4,...,9$ $j, k=0,...,3$
3rd order elm.	$N_i = \frac{1}{2}(3L_i - 1)(3L_i - 2)L_i$	$i=0,...,3$
	$N_i = \frac{9}{2}(3L_j - 1)L_j L_k$	$i=4,...,15$ $j, k=0,...,3$
	$N_i = 27L_j L_k L_l$	$i=16,...,19$

L : Volume coordinate variable
N : Shape function
i : Number of nodes

Amane Takei, Akihiro Kudo, Makoto Sakamoto

2.2. Helmholtz equation

In the 3-dimensional sound field, the wave equation for velocity potential is expressed by the following equation:

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} + \frac{\partial^2 \phi}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} = q \tag{1}$$

where ϕ is the velocity potential, c is the speed of sound, and q is the distribution function.

To consider the steady-state, the velocity potential is expressed by Eq. (2). Using Eq. (1) and Eq. (2), the Helmholtz equation is obtained:

$$\phi = \Phi e^{-j\omega t} \tag{2}$$

$$\frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} + \frac{\partial^2 \Phi}{\partial z^2} + \frac{\omega^2}{c^2}\Phi = q \tag{3}$$

where ω is the angular frequency.

The velocity potential of Eq. (3) and calculate the sound pressure using the following equation are obtained:

$$p = j\omega\rho\Phi \tag{4}$$

where j is the imaginary number, and ρ is the medium density.

2.3. Finite element formulation

To derive a weak form, the Galerkin method is applied to Eq. (3). By applying the finite element approximation and discretization, the following equation is obtained:

$$-k^2[M]\{\Phi\} + j\omega\rho[C]\{\Phi\} + [K]\{\Phi\} = \{q\} \tag{5}$$

where $[\,\cdot\,]$ is a matrix, $\{\,\cdot\,\}$ is a vector. In Eq. (5), the matrices $[M]$, $[C]$ and $[K]$ can be calculated using Eq. (6.1,2,3):

$$[M]_e = \iiint_{\Omega_e} \{N\}\{N\}^T d\Omega_e \tag{6.1}$$

$$[K]_e = \iiint_{\Omega_e} \nabla\{N\}\nabla\{N\}^T d\Omega_e \tag{6.2}$$

$$[C]_e = -\frac{1}{Z_n}\iint_{\Gamma_e} \{N\}\{N\}^T d\Gamma_e \tag{6.3}$$

where N is the shape function, $[M]_e$ and $[K]_e$ are the volume integrals, and $[C]_e$ is a surface integral to the sound-absorbing boundary surface. Z_n is a specific acoustic impedance.

2.4. Calculaion of element matrix

Let us consider the calculation $\{N\}\{N\}^T$ in Eqs. (6.1) and (6.3). This calculation uses an integration formula of the 3 or 2-dimensional finite element method. These integration formulas are shown in the following equations [7].

$$\iiint_{\Omega_e} L_1^k L_2^l L_3^m L_4^n dxdydz$$
$$= 6V_e \frac{k!\,l!\,m!\,n!}{(k+l+m+n+3)!} \tag{7}$$

$$\iint_{\Gamma_e} L_1^k L_2^l L_3^m dxdy = 2A_e \frac{k!\,l!\,m!}{(k+l+m+2)!} \tag{8}$$

On the other hand, the calculation of $\nabla\{N\}\nabla\{N\}^T$ in Eq. (6.2) shown in the following equation.

$$\frac{\partial f}{\partial x} = \frac{1}{6V_e}\left(b_1 \frac{\partial f}{\partial L_1} + b_2 \frac{\partial f}{\partial L_2} + b_3 \frac{\partial f}{\partial L_3} + b_4 \frac{\partial f}{\partial L_4}\right)$$

$$\frac{\partial f}{\partial y} = \frac{1}{6V_e}\left(c_1 \frac{\partial f}{\partial L_1} + c_2 \frac{\partial f}{\partial L_2} + c_3 \frac{\partial f}{\partial L_3} + c_4 \frac{\partial f}{\partial L_4}\right) \tag{9}$$

$$\frac{\partial f}{\partial z} = \frac{1}{6V_e}\left(d_1 \frac{\partial f}{\partial L_1} + d_2 \frac{\partial f}{\partial L_2} + d_3 \frac{\partial f}{\partial L_3} + d_4 \frac{\partial f}{\partial L_4}\right)$$

We convert the finite element equations of (5) to a matrix form as follows:

$$Ku = f. \tag{10}$$

2.5. Hierarchical domain decomposition method

The original analysis domain is first divided into parts, which are further decomposed into smaller domains called subdomains. This is called the hierarchical domain decomposition method (HDDM) [8], [9].

3. Numerical experiment

3.1. Verification by benchmark problem

Fig.2 shows that test model for simulation. The model is AHLV100 that is known as a reference model in code_Aster. This is also described in the ADVENTURE_sound manual as a sample. This simulation was done to confirm the use of transient analysis in ADVENTURE_Sound. To evaluate the accuracy of the acoustic analysis code, an acoustic benchmark problem is used. The analysis uses the test model AHLV100 of Code_Aster [10], which is known as a representative benchmark problem among acoustic problems (Fig. 4).

This model is an acoustic tube that has a length of 1 [m], a height of 0.1 [m] and a width of 0.2 [m]. It has a vibration boundary at the left end and a sound absorption boundary at the right end. The other faces are given rigid boundaries. The specific acoustic impedance Z_n=445.9 [kg/m^3 · s] is given as a sound absorption boundary condition.

The accuracy is calculated from the average error of four points on the sound absorption boundary. The formula for calculating the theoretical solution is as follows:

$$p(x,y,z) = \rho c V_n exp(-ikx) \tag{11}$$

where ρ is the medium, c is the speed of sound, and V_n is the particle velocity.
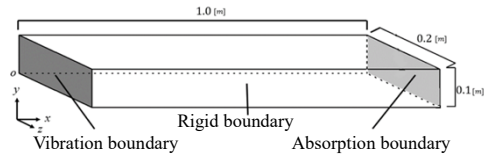
Fig. 2. Acoustic benchmark problem AHLV100

3.2. Performance evaluation

The performance of higher-order elements is evaluated based on the number of elements, accuracy rate, and memory usage. The performance evaluation conditions are: 4.0 [kHz] for frequency, air for medium, and 343 [m/s] for sound velocity. The analysis uses a PC cluster composed of 5 PCs (40 cores) equipped with a multicore CPU (Intel Core i7-9700K, 3.6GHz 8core, 32GB of memory).

Numerical results for the error rate, plotted against the number of elements, are shown in Fig. 3. Results for the memory, plotted against the number of elements, are shown in Fig. 4.
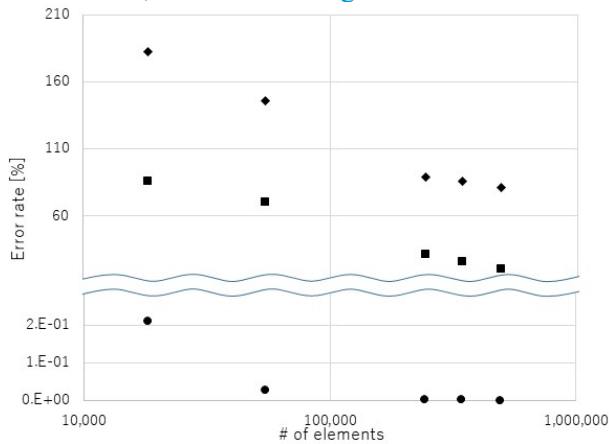
Fig. 3. Error rate plotted against the number of elements
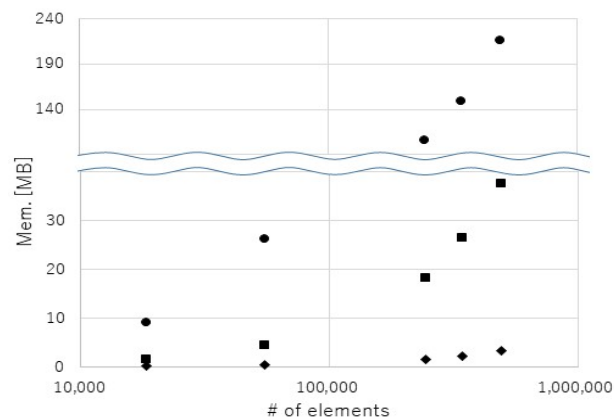
Fig. 4. Memory usage plotted against the number of elements

The accuracy increases dramatically when a higher-order element is applied. In particular, the accuracy changes more rapidly as the order increases.

3.3. Analyses using real environment models

To confirm the effectiveness of higher-order elements, the sound field of the real environment model is analyzed. The model used for analysis is concert hall model that is shown in Fig. 5.
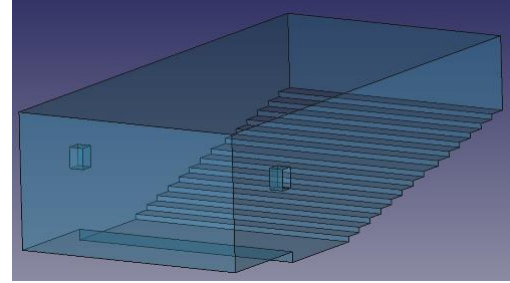
(a) Concert hall 3D CAD model
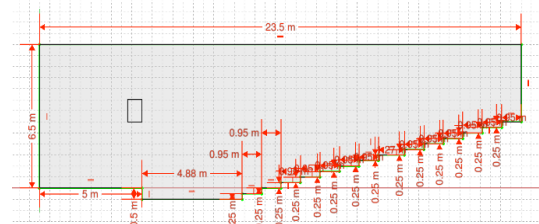
(b) Concert hall model dimensions
Fig. 5. Concert hall model

The concert hall model is constructed based on real spaces [11].

The concert hall model is 11.5 [m] wide, 7 [m] high, and 23.5 [m] deep. The sound source is set as a pair of speakers at either end of the stage. Oakwood flooring, a wooden stage, and a glass wool wall at the back of the hall are used as sound-absorbing boundary conditions. The sound field is analyzed by applying a 400 [Hz] sound to these models.

The results of these analyses are shown in Table 2. The visualization results are shown in Fig. 6, and the convergence histories of the iterative method (COCG) are shown in Fig. 7. The concert hall analysis using the 1st and the 2nd order elements are excluded from the evaluation because the number of elements exceeds 100 million. The data I/O library currently applied is 32 bit. This library cannot use data sizes that exceed 100 million elements. A 64-bit I/O library is currently under development.

Table 2. Numerical results of the concert hall model

Element type	Number of elements	Number of nodes	Elapsed Time [sec]	Memory requirement [MB/core]
1st order	823,285,162			
2nd order	102,910,645			
3rd order	878,624	4,051,235	53.97	391.97

Amane Takei, Akihiro Kudo, Makoto Sakamoto

(a) Visualization results (X-Y)

(b) Visualization results (X-Z)
Fig. 6. Concert hall model

Fig. 7. Converegence history of the COCG
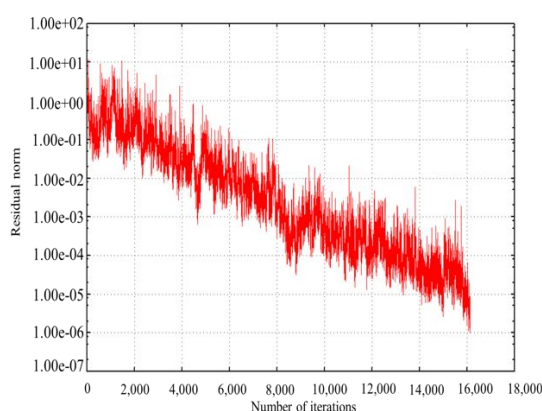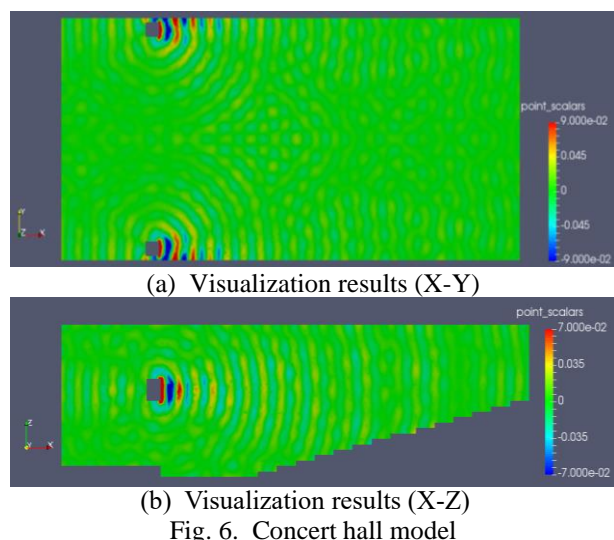
4. Conclusion

This paper described a large-scale acoustic analysis method using a domain decomposition method and the introduction of higher-order elements. The performance of proposed method was evaluated with higher-order elements using of AHLV100. The error rate, number of elements, and memory usage were the considered as evaluation criterions. It was shown that the calculation efficiency improved in higher-order elements. In particular, in the 3rd order element, the calculation efficiency was vastly improved. Furthermore, the accuracy of higher-order elements was verified. The 2nd order element and the 3rd order element were compared in terms of the number of elements and calculation time. Additionally, real environment model was analyzed, namely, a small concert hall model. It was shown that the real environment models could be successfully analyzed by using higher-order elements.

References

1. T. Okuzono, M. Shadi and K. Sakagami, "Potential of Room Acoustic Solver with Plane-Wave Enriched Finite Element Method", Appled. Science, Vol.10, No.6, 2020.
2. K. Suzuki, Y. Yamada, S. Koyanagi, T. Hidaka, "Basic study on improvement of precision of measurement of room acoustic characteristics using scale model and 3D sound field auralization", Acoustical Society of Japan (ASJ) Academic journal, Vol. 74, No. 5,
3. K. Kowalczyk, M. Walstijn, "Room Acoustics Simulation Using 3-D Compact Explicit FDTD Schemes", IEEE Transactions on Audio, Speech, and Language Processing, Vol.19, No.1, pp.34-46, 2011.
4. Y. Yasuda, T. Oshima, T. Sakuma, A. Gunawan, T. Masumoto, "Fast multipole boundary element method for low-frequency acoustic problems based on a variety of formulations", Journal of Computational Acoustics, Vol.18, No.4, pp.363-395, 2010.
5. K. Okuzono, T. Otsuru, R. Tomiku, N. Okamoto, "Fundamental accuracy of time domain finite element method for sound-field analysis of rooms", Applied Acoustics, Vol. 71, pp. 940-946, 2010.
6. K. Ueno Ed., Science of concert hall –Harmony of shape and sound- (in Japanese, CORONA PUBLISHING Co., Ltd. 2012).
7. M. Koshiba, Fundamentals of the finite element method for light and waves (in Japanese, MORIKITA PUBLISHING Co., Ltd, 1990).
8. J. Mandel, "Balancing domain decomposition", Communications on Numerical Methods in Engineering, Vol.9, pp.233-241, 1993.
9. M. Ogino, A. Takei, S. Sugimoto, S. Yoshimura, "A numerical study of iterative substructuring method for finite element analysis of high frequency electromag-netic fields", Computers and Mathematics with Applications, Vol. 72, Issue 8, pp. 2020-2027, 2016
10. F. Stifkens, G. Rousseau, "Code_Aster Manuel de Validation", No. V8. 22. 100, 1998.
11. STAR PINE'S CAFÉ, Homepage: http://mandala.gr.jp/SPC/, 2019.

Authors Introduction

Prof. Amane Takei

He is working as Associate Professor for Department of Electrical and systems Engineering, University of Miyazaki, Japan. His research interest includes high performance computing for computational electromagnetism, iterative methods for the solution of sparse linear systems, domain decomposition methods for large-scale problems. Prof. Takei is a member of IEEE, an expert advisor of IEICE, a delegate of the Kyushu branch of IEEJ, a member of JSST。 .

Prof. Makoto Sakamoto

He is presently a professor in the Faculty of Engineering, University of Miyazaki. His first interests lay in hydrodynamics and time series analysis, especially the directional wave spectrum. He is a theoretical computer scientist, and his current main research interests are automata theory, languages and computation. He is also interested in digital geometry, digital image processing, computer vision, computer graphics, etc.

Prof. Akihiro Kudo

He received Ph.D. degree from Nagaoka University of Technology He is a professor in the Department of engineering for innovation , National Institute of Technology, Tomakomai college. He is a member of Acoustical society of Japan, Information and Communication Engineers (IEICE).

Sound field evaluation on the acoustical experiment using non-steady state analysis

Akihiro Kudo*
National Institute of Technology, Tomakomai college, Japan

Makoto Sakamoto, Amane Takei
Faculty of Engineering, University of Miyazaki, Japan

*Corresponding author

Abstract

Identifying acoustic effects of sound waves emitted by a sound source on the listener is important for understanding the localization of sound as perceived by the listener. In this presentation, we will analyze the effect of the presence of a chair on the psychoacoustic effect using a practical model simulating a psychoacoustic experiment consisting of a loudspeaker, a listener, and a chair. Non-steady state analysis is required to analyze changes in acoustic effects over time. To analyze the interference of sound waves in the vicinity of the listener, a short Gaussian pulse-like solitary wave is set up on the loudspeaker's vibrating surface. Since sound waves in the audible band are used for these simulations, a mesh size of less than 25 mm is used. The laboratory is 3m of depth and width, and 2.5m of height, which means that the computational whole target consists of over 15 million elements. This simulation requires a large number of calculations, so a parallel computing method with ADVENTURE_Sound that is an open-source software for sound simulation, is used to run the simulation. From the simulation results of the non-stationary analysis and consider the effect of the presence of the chair on the sound image localization.

Keywords: large-scale simulation, acoustical sound field, non-steady state analysis

1. Introduction

In the experimental environments, there are objects to disturb sound wave propagation: wall, floor, ceil and chair, etc. Since the effects of these objects change the sound waves reaching the listener's ears, the accuracy of sound image localization may also change. Our goal is to clarify the effect of these environments on the sound image localization characteristics. To evaluate the time-related effects of objects around the listener on the sound waves reaching the listener's ears, this paper examines a non-steady state sound field simulation is performed.

2. Adventure Sound
2.1. Brief introduction

ADVENTURE_Sound is an open-source software for sound simulation. This allows us to simulate parallel finite element analysis software for the sound fields in huge space. Analyses are performed by solving the large-scale linear system with parallel computing based on the Mandel's domain decomposition method (DDM) [1].

2.2. Simulation procedures with ADVENTURE_Sound

The procedure to use ADVENTURE_Sound is as follows; 1) generate the mesh data, 2) add boundary conditions (BCs), 3) subdivide the mesh data into subdomains, 4) simulate the sound field distributions, 5) convert to the results to vtk data for visualization. On the fourth procedure, it needs to calculate Helmholtz equation in the time domain. For this, Newmark β method is adopted to the equation. As the boundary conditions, input wave function and first and second derivatives of that function.

3. Preliminary test to simulate non-steady state analysis

3.1. Test model

Fig.1 shows that test model for simulation. The model is AHLV100 that is known as a reference model in code_Aster. This is also described in the ADVENTURE_sound manual as a sample. This simulation was done to confirm the use of transient analysis in ADVENTURE_Sound.

Akihiro Kudo, Makoto Sakamoto, Amane Takei
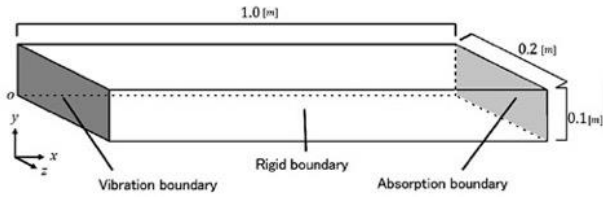
Fig. 1. Test model

3.2. Specs of the parallel computer

Our parallel computer is called as TOMAHAWK.
Table 1 shows that the detailed specs of the parallel computer.

Table 1. Specs of parallel computer

CPU	Intel Xeon E5-2650L,1.8GHz with 16 cores L2 20480[KB]
Main memory/node[GB]	32
Number of nodes	8
Total cores	128
OS	CentOS 6.2
Compiler	gcc-4.4.7 mipch-3.2

3.3. Simulation conditions

Table 2 shows that analysis conditions.

Table 2. Analysis conditions for the test model

Mesh size[m]	0.012
Number of elements	140, 604
Absorption boundary [kg/m^{2}s]	445.9
Medium(air) r_0[kg/m^3]	1.3
Speed of sound c_0[m/s]	343
Time resolution Δt [μs]	1.0
Duration time to simulate[μs]	4000
Velocity potential $\{\Phi\}_t$ Full width at half maximum[ms] Cut-off frequency f_c[Hz]	Gaussian pulse 0.264 646

For boundary conditions as input wave, the single wave is useful to analysis the transient sound field, because sound reflections are observed clearly. In this study, a gaussian pulse as velocity potential is adopted as follows,

$$\{\Phi\}_t = e^{-\{\alpha(t-t_0)\}^2} \ ... \qquad (1)$$

In the Eq. (1), decay parameter α is set to 3141, delay parameter t_0 is set to 0.001[s].
An input wave for boundary conditions is shown in Fig. 2.
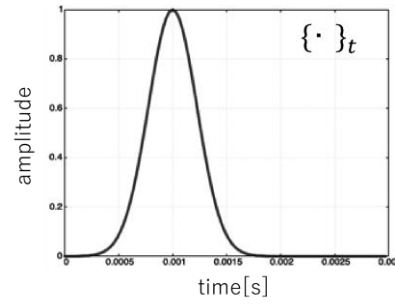
Fig. 2. An input wave for boundary conditions

3.4. Simulation results

Fig.3 shows that the simulation results at some time points. In this figure, bright red color indicates that high intensity of the vector potential that is proportional to the sound pressure.
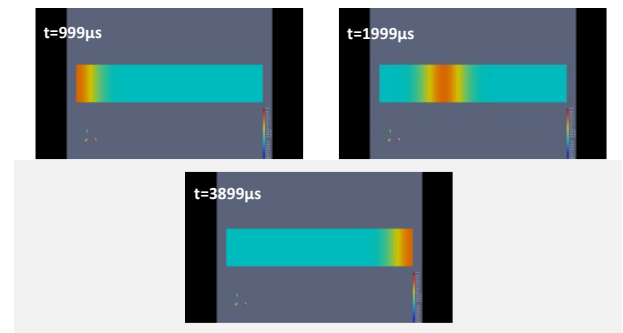
Fig. 3. Simulation results at some time points

3.5. Discussions

From this figure, a single waveform with continuous and positive intensity was observed and the waveform moves from left side to right. Because the sound wave takes 2900μs to reach the right side and length of AHLV100 is 1m, estimated sound speed is 345m/s. This estimate is consistent with the theoretical value. The simulation results show that the use of ADVENTURE_Sound in non-steady state acoustical analysis is appropriate.

4. Simulation of acoustical experimental environment

4.1. Simulation model

Fig. 4 is the dimensions on the side view of simulation model. Table 3 and 4 are dimensions of target and analysis conditions for acoustical experimental environment, respectively.
In this simulation model, there are one loudspeaker and snowman pair that simulates an acoustical experimental arrangement. The snowman simulates a listener's head and torso [2]. In addition, this model also includes a chair to hold the listener in position. These are practical and typical acoustical experimental environment. By

analyzing the sound waves at a position corresponding to the ear of the Snowman model, it is possible to evaluate the effect of sound reflections from the chair as an obstacle on the sound image localization.

Fig. 4. Dimensions on the side view of simulation model

Table 3. Dimensions of the target

Room, width [m] ×depth[m]×height [m]	3×3×2.5
Room volume [m^3]	22.5
Loudspeaker, width [m] ×depth [m] ×height [m]	0.1×0.12×0.12
Diameter of the listener head, torso [m]	0.177, 0.3 [2]
Distance from loudspeaker to listener [m]	1.5
Backrest, width [m] ×depth [m] ×height [m]	0.5×0.1×0.65
Seat, width [m] ×depth [m] ×height [m]	0.5×0.5×0.1

Table 4. Analysis conditions for the acoustical experimental environment

Mesh size [m]	0.025
Number of elements	14,497,814
Room wall, floor, ceiling [kg/m^{2}s]	445.9
Loudspeaker, snowman, chair [kg/m^{2}s]	4.56× 10^9
Medium(air) r_0 [kg/m^3]	1.3
Speed of sound c_0 [m/s]	343
Time resolution Δt [µs]	1.0
Duration time to simulate [µs]	8000
Velocity potential $\{\Phi\}_t$	Gaussian pulse
Full width at half maximum [ms]	0.264
Cut-off frequency f_c [Hz]	646

4.2. Simulation results

Fig. 5 is the simulation results on the xy-plane at y=1.5m. Since gaussian pulse is given to the surface of the loudspeaker, the sound waves are spread over. As sound wave travels from the loudspeaker, the sound directivity can be observed on x-axis.

Fig. 6 is the simulation results on the xy-plane at z=1.0m. The distance from the loudspeaker to the listener is 1.5m, so it takes 4.4ms for the sound wave to reach the listener.

From these results, the sound wave reaches to the listener at approx. 5ms and the time for gaussian pulse to show its maximum value is 1ms, because it takes for approx. 4ms, the simulation corresponds to the theoretical value.

Fig. 5. Simulation results on the xz-plane at y=1.5m

Fig. 6. Simulation results on the xy-plane at z=1.0m

4.3. Discussions

Fig. 7 is the snapshots on the xz-plane at y=1.5m.
From 6.7ms to 7.3ms, an increased sound pressure was observed bottom of the listener's torso due to the sound waves reflecting off the seat of the chair.
The sound wave reaches the listener at approx. 5.2ms and the sound wave reflected by the chair returns to the listener at approx. 6.4ms, so this duration time is 1.2ms. This might affect the spaciousness of sound image although not directional accuracy of the sound image localization, because this time interval exceeds the

Akihiro Kudo, Makoto Sakamoto, Amane Takei

maximum time difference 0.6ms that humans use to localize a sound image [3].

Fig. 7. Snapshots on the xz-plane at y=1.5m

5. Conclusion

Purpose of this paper is to clarify the effect of these environments on the human auditory characteristics, especially sound image localization ability. The simulation target is loudspeaker and snowman model with chair, was adopted as a typical acoustical experimental environment. The simulation results showed that 1) An increased sound pressure was observed bottom of the listener's torso due to the sound waves reflecting off the seat of the chair, 2) The time taken for sound wave reaching the listener to reflect off the chair and return to the listener, was 1.2ms.

Since the time required for sound image localization is up to 0.6ms and the sound wave delay caused by chair reflection is shorter than this time, the reflection on chair may affect the spread of the sound image more than the directional accuracy of sound image localization.

Acknowledgements

References

1. J. Mandel, "Balancing domain decomposition", Communications on Numerical Methods in Engineering, Vol.9, pp.233-241, 1993.
2. S. Shimada, K. Sugiyama, H. Hokari, Wave equation and Head related transfer functions model (in Japanese, Kameda book service, 2011).
3. B. C. J. Moore, An introduction to the Psychology of Hearing (ELSEVIER ACEDEMIC PRESS, 2007)

Authors Introduction

Prof. Akihiro Kudo

He received Ph.D. degree from Nagaoka University of Technology He is a professor in the Department of engineering for innovation, National Institute of Technology, Tomakomai college. He is a member of Acoustical society of Japan, Information and Communication Engineers (IEICE).

Prof. Makoto Sakamoto

He is presently a professor in the Faculty of Engineering, University of Miyazaki. His first interests lay in hydrodynamics and time series analysis, especially the directional wave spectrum. He is a theoretical computer scientist, and his current main research interests are automata theory, languages and computation. He is also interested in digital geometry, digital image processing, computer vision, etc.

Prof. Amane Takei

He is working as Associate Professor for Department of Electrical and systems Engineering, University of Miyazaki, Japan. His research interest includes high performance computing for computational electromagnetism, iterative methods for the solution of sparse linear systems, domain decomposition methods for large-scale problems. Prof. Takei is a member of IEEE, an expert advisor of IEICE, a delegate of the Kyushu branch of IEEJ, a director of JSST.

A Deep Insight method with Morphological Analysis

Toyoaki Tomimoka
Graduate School of Engineering, Miyazaki University,
Japan, hm18026@student.miyazaki-u.ac.jp

Satoshi Ikeda*
Department of Computer Science and System Engineering, Miyazaki University,
Japan, bisu@cs.miyazaki-u.ac.jp

Makoto Sakamoto
Department of Computer Science and System Engineering, Miyazaki University,
Japan, saka2000@cc.miyazaki-u.ac.jp

Takao Ito
Graduate School of Engineering, Hiroshima University,
Japan, itotakao@hiroshima-u.ac.jp

*Corresponding Author

Abstract

DeepInsight method, which interprets non-image data as image data, has garnered significant attention; however, a universal approach for this method has yet to be standardized. Our objective was to formulate a comprehensive method for DeepInsight, encompassing three key steps: strategic variable placement, effective feature extraction, and optimal model construction. Consequently, we successfully developed a model capable of predicting the Nikkei Stock Average with an accuracy of approximately 60%, marking a significant stride towards establishing the versatility of the DeepInsight method.

Keywords: DeepInsight method, t-SNE, Convolutional Neural Network, Morphological analysis.

1. Introduction

1.1 Research background

Machine learning faces challenges when dealing with high-dimensional data, known as the curse of dimensionality [1]. This issue arises because as the number of dimensions increases, the potential combinations of data grow exponentially, leading to computational challenges and inadequate learning results. To address this, two approaches can be employed:

1) Dimensionality Reduction: One strategy is to reduce dimensionality while preserving information. t-SNE [2] is a common technique that expresses data distances as conditional probabilities, keeping similar high dimensional data close in lower dimensions.

2) Feature Combination Restrictions: Another approach involves limiting combinations of features. Deep Learning [3] achieves this by dividing the problem into smaller regions, but it is most effective with data exhibiting strong pixel correlations, such as images. DeepInsight method [4], has emerged to extend deep learning to non-image data, although a standardized procedure for its application is yet to be established.

With the rise of social media, text mining utilizing platforms like SNS has gained prominence. Predicting stock prices has seen success by leveraging emotions expressed on SNS. Researchers are exploring deep learning analysis of news articles to build more accurate models, especially as sentiment analysis on platforms like

Toyoaki Tomimoka, Satoshi Ikeda, Makoto Sakamoto, Takao Ito

SNS has proven effective. The broader adoption of these techniques holds promise for improving predictions and insights across various domains.

1.2   Research purpose

The aim of this research is to transform the three stages involving proper variable placement, feature extraction, and model construction into a versatile DeepInsight method. By utilizing the proposed method, we aim to accurately predict tomorrow's stock price movements and validate the effectiveness of the approach.

2．Research method

2.1   Summary of t-SNE

The t-SNE is a dimension reduction algorithm designed to reduce high-dimensional data to two or three dimensions. The algorithm represents the proximity between points x_i and x_j on a high dimension as a joint probability distribution p_{ij}:

$$p_{ij} = \frac{p_{j|i} + p_{i|j}}{2n} \ (i \neq j), \ \ p_{ii} = 0.$$

Here $p_{j|i} = \dfrac{\exp\left(\dfrac{-\left\|x_i - x_j\right\|^2}{2\sigma_i^2}\right)}{\sum_{k \neq i} \exp\left(\dfrac{-\left\|x_i - x_k\right\|^2}{2\sigma_i^2}\right)}$ for $i \neq j$.

And the proximity between points y_i and y_j in a low dimension as a joint probability distribution q_{ij}:

$$q_{ij} = \frac{\left(1 + \left\|y_i - y_j\right\|^2\right)^{-1}}{\sum_{k \neq i}\left(1 + \left\|y_i - y_j\right\|^2\right)^{-1}} \ (i \neq j).$$

The objective is to match the distance relationships between data points in high dimensions with those in low dimensions after dimension reduction. This is achieved by setting set $p_{i|j} = q_{i|j}$, and the Kullback-Leibler (KL) divergence is employed to measure the distance between $p_{i|j}$ and $q_{i|j}$ with the goal of minimizing the loss function:

$$C = \sum_i KL(P_i \| Q_i) = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}}$$

Finally, stochastic gradient descent is used to minimize the loss function. The gradient is given by

$$\frac{\delta C}{\delta y_i} = 4\sum_j (p_{ji} - q_{ji})(y_i - y_j)\left(1 + \left\|y_i - y_j\right\|^2\right)^{-1}$$

Using this gradient, is gradually updated using the renewal formula:

$$Y^{(t)} = Y^{(t-1)} + \eta\frac{\delta C}{\delta Y} + \alpha\left(Y^{(t-1)} - Y^{(t-2)}\right).$$

Here, t denotes the iteration, η is the learning rate, and α is the momentum term.

2.2   Learning Image Data with t-SNE using CNN

A Convolutional Neural Network (CNN) comprises fully connected, convolution, and pooling layers. In the fully connected layer, units are interconnected, with weights depicted in Fig. 2.1. Inputs and outputs to each unit are one-dimensional vectors, and the output unit's value is determined by multiplying input values by connection weights and adding bias.
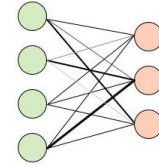
Fig. 2.1 Fully Connected layer weights

In the Convolutional layer, the outcomes of convolution operations for each filter are generated and serve as one unit in the subsequent layer. To derive the output, akin to the Fully Connected layer, compute the inner product for each filter, incorporate the bias, and apply the activation function. The Pooling layer condenses information to convert the input data into a more manageable format. Fig. 2.2 illustrates Max-pooling, wherein the maximum value is selected from color-coded small areas.
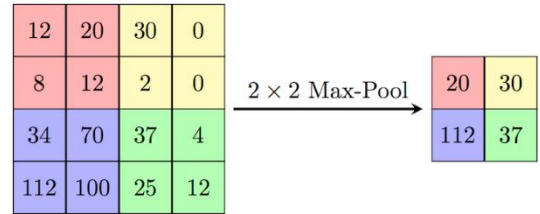
Fig. 2.2 Max-pooling

2.3   Morphological analysis
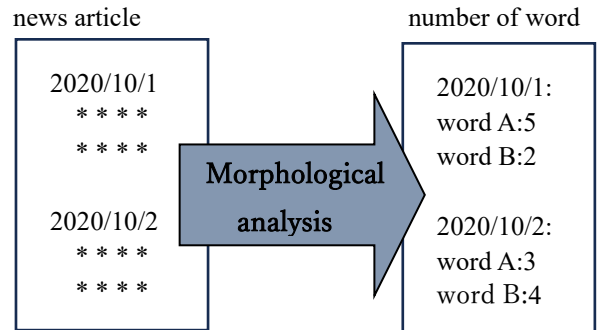
Fig. 2.3 Example of Morphological analysis

Morphological analysis dissects sentences into morphemes, assigning parts of speech. Using a wordbook, the least-cost method selects optimal words based on part-of-speech

connection and word appearance. This process is applied to news articles to count occurrences of words essential for stock price prediction, as illustrated in Fig. 2.3.

3. Experiment

3.1 Create input data

In the experiment, two datasets were used:
1) Daily stock prices of the 225 stocks constituting the Nikkei Stock Average.
2) Word count in news articles.

For 2), news articles related to stocks from the Kabutan website [5] during the period from 2020 to 2021 were used. Morphological analysis was applied to the news articles to count the occurrences of words considered effective in predicting stock prices (e.g., "weak yen"). A specialized dictionary (with 563 registered words) tailored for stock price prediction was employed for word counting. The data for stock prices and word counts in news articles were consolidated into a single CSV datasheet (Table 3.1).

Table3.1 CSV datasheet

	10/1	10/2	・・	10/10
company1	a1	b1	・・	j1
company2	a2	b2	・・	j2
・・	・・	・・	・・	・・
company225	a225	b225	・・	j225
word1	A1	B1	・・	J1
word2	A2	B2	・・	J2
・・	・・	・・	・・	・・

The first row of Table 3.1 represents the dates for which data was obtained. In this study, one sheet was created for a continuous 10-day period. Lowercase letters (a1, a2, etc.) represent daily stock prices for individual companies, while uppercase letters (A1, A2, etc.) represent the word count for each date. A total of 222 sheets of data were created for the entire data collection period. Using t-SNE, the data from each of the 225 sheets were transformed into image data composed of 783 points (225 for stock prices + 563 for words), as illustrated in Fig. 3.1. These images will be used as training and testing data for the Convolutional Neural Network (CNN).
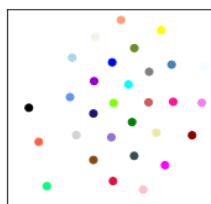
Fig. 3.1 An example of t-SNE transformation on CSV data

Each generated image is labeled with 0, 1, or 2. Here, 0 signifies that the next day's Nikkei Stock Average will remain unchanged (range from -100 yen to +100 yen), 1 indicates an increase (range from more than +100 yen), and 2 indicates a decrease (range from less than -100 yen).

3.2 Data processing

Firstly, split all the data into training and test sets. Allocate 80% of the total data to the training set and 20% to the test set to ensure unbiased data representation. Subsequently, normalize the input image data. Given that color values are typically expressed as {0, 1, ..., 255}, normalizing the data to the range [0, 1] can be achieved by dividing each value by 255. Finally, apply a one-hot vector transformation to the label data of the input images, converting them into a binary class matrix.

3.3 Evaluation method

Display the accuracy rates for each training and test dataset to assess the progress of learning and whether DeepInsight method is effective for prediction.

3.4 Experimental results

The number of training data is 175, and the number of test data is 45. Among the 45 test data, the percentage of '0' was 16%, the percentage of '1' was 53%, and the percentage of '2' was 31%, indicating that the percentage of '0' was relatively small, with '1' accounting for more than half.

Fig. 3.2 Experimental result

In Fig. 3.2, where the horizontal axis represents the epoch (number of executions) and the vertical axis represents accuracy (correct answer rate), the accuracy rate of the training data reached nearly 100% at 25 epochs. The test data achieved its highest accuracy rate of 59.6% at 27 epochs, but further increasing the number of epochs did not lead to a higher accuracy rate. Additionally, at epoch 10,

Toyoaki Tomimoka, Satoshi Ikeda, Makoto Sakamoto, Takao Ito

the accuracy rates of the training data and test data were similar, but thereafter, a significant difference emerged.

Examining the conditional probabilities in Table 3.2, out of 45 test data items, 53% (24 items) were predicted as '1', and among those, 58% (14 items) were correctly predicted as '1'. In contrast, 31% (14 items) were predicted as '2', and among those, 50% were correctly predicted as '2'.

Table3.2　Conditional probability

	0	1	2
0(16%)	14%(1)	71%(5)	14%(1)
1(53%)	16%(4)	58%(14)	25%(6)
2(31%)	14%(2)	36%(5)	50%(7)

4.Consideration

Our experiment aimed to turn the three steps of appropriate variable placement, feature extraction, and appropriate model construction into a general-purpose DeepInsight method. Appropriate variable placement was achieved by using t-SNE to convert effective prediction words obtained through morphological analysis of stock price statistical data and news articles into image data. Feature extraction utilized CNN, an effective method for image recognition in machine learning. Despite struggling with suppressing overfitting, we achieved a maximum accuracy of approximately 60%. Assuming the model was used for actual trading, this resulted in an approximately 11% increase in profits over one month, surpassing the average annual interest rate for stock investments (3-10%).

These results indicate that the "appropriate variable placement," "feature extraction," and "appropriate model construction" approach is effective in DeepInsight. Future challenges include increasing the types of label data and improving overall accuracy.

References

1. R.EBellman, "Dynamic Programming", Princeton University Press,1957
2. G.E.Hinton, Laurens van Maaten, "Visualizing Data using t-SNE",Journal of Machine Learning Research 9 (2008) 2579-2605
3. G.E.Hinton, S.Osindero, Y.W.Teh, "A Fast Learning Algorithm for Deep Belief Nets", NeuralComputation 18(7), 2006,4pp.1527-1554
4. Tatsuhiko Tsunoda, "DeepInsight: A methodology to transform a non-image data to an image for convolution neural network architectu", Scientific Repotrts. 10.1038/s41598-019-47765-6
5. "株探"

Authors Introduction

Mr. Toyoaki Tomioka

He is a student of the Mechanical and Information Systems Course, graduate school of Engineering, University of Miyazaki.

Prof. Satoshi Ikeda

He received PhD degree from Hiroshima University. He is an associate professor in the Faculty of Engineering, University of Miyazaki. His research interest includes graph theory, probabilistic algorithm, fractal geometry and measure theory.

Prof. Makoto Sakamoto

He is presently a professor in the Faculty of Engineering, University of Miyazaki. His first interests lay in hydrodynamics and time series analysis, especially the directional wave spectrum. He is a theoretical computer scientist, and his current main research interests are automata theory, languages and computation. He is also interested in digital geometry, digital image processing, computer vision, computer graphics, etc.

Prof. Takao Ito

He is Professor of Management of Technology (MOT) in Graduate School of Engineering at Hiroshima University. He is serving concurrently as Professor of Harbin Institute of Technology (Weihai) China. He has published numerous papers in refereed journals and proceedings, particularly in the area of management science, and computer science. He has published more than eight academic books including a book on Network Organizations and Information (Japanese Edition). His current research interests include automata theory, artificial intelligence, systems control, quantitative analysis of inter-firm relationships using graph theory, and engineering approach of organizational structures using complex systems theory.

Support for Museum Exhibition of Small Fungi using AR Technology

Kakeru Takemura, Ota Hamasuna, Fumito Hamakawa
Graduate School of Engineering, University of Miyazaki, Japan

Satoshi Ikeda, Kaoru Ohe, Amane Takei, Makoto Sakamoto*
Faculty of Engineering, University of Miyazaki, Japan

Shuichi Kurogi
Miyazaki Prefectural Museum of Nature and History, Japan
Email: hm18022@student.miyazaki-u.ac.jp, saka2000@cc.miyazaki-u.ac.jp

*Corresponding author

Abstract

There are 25 species of bioluminescent fungi that have been confirmed in Japan, and 12 species have been confirmed in Miyazaki Prefecture. Those fungi are very small. Therefore, it is difficult to observe the structure of fungi exhibited in museums with the naked eye. The purpose of this research is to display 3DCG models of mushrooms using AR (Augmented Reality) technology in order to facilitate observation of these small mushrooms exhibited in museums. Two hundred visitors to the museum were asked to use the application and complete a survey. In order to measure the ease of observation of different app functions and mobile devices used in the survey, we divided the respondents into four groups.

Keywords: AR, Exhibition Support, Museum, bioluminescent fungi

1. Introduction

Augmented reality is a technology that superimposes digital information on the real world. In recent years, AR technology can be used not only on head mount display, but also on mobile devices such as smartphones. This study will examine the use of AR technology to support observation in museums. There are 25 species of bioluminescent fungi that have been confirmed in Japan, and 12 species have been confirmed in Miyazaki Prefecture. Those fungi are very small. Therefore, it is difficult to observe the structure of fungi exhibited in museums with the naked eye. The purpose of this research is to display 3DCG models of fungi using AR (Augmented Reality) technology in order to facilitate observation of these small fungi exhibited in museums.

2. AR Application Development
2.1. Development Environment

An AR application was created for use with smartphones and Tablets. The development environment is shown (Table 1). Blender is a comprehensive 3DCG software. It was used to create a 3DCG model of fungi [3], [4]. Unity was used to develop applications for mobile terminals [1], [2].

Table 1. Development Environment

OS	Windows 10 Pro
Software	Blender 2.93
	Unity 2020.3.22f1
	Visual Studio 2019
Smart phone	Galaxy S8
Tablet	Yoga Tab 11

2.2. Target fungi

The bioluminescent fungi, *Favolaschia peziziformis* [Fig. 1], *Cruentomycena orientalis* [Fig. 2] and *Favolus manipularis* [Fig. 3], were the subjects of this study.

2.3. 3DCG of fungi

The CG model was created based on the images provided.

Fig. 1 *Favolaschia peziziformis*

Kakeru Takemura, Ota Hamasuna, Fumito Hamakawa, Satoshi Ikeda, Kaoru Ohe, Amane Takei, Makoto Sakamoto, Shuichi Kurogi
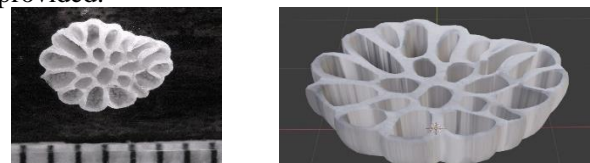
Fig. 2 *Cruentomycena orientalis*

Fig. 3 *Favolus manipularis*

2.4. AR application

An image that serves as a marker is loaded, and an AR 3DCG model of a mushroom is displayed on top of it. When the marker disappears from the screen, the displayed 3DCG model also disappears, and when the same or another marker is scanned, the 3DCG model is displayed again. This application was used as the basic application. In addition to the basic application, another application was prepared with a function that tilts the mushrooms in the same direction when the user moves his/her finger while touching the AR-displayed mushrooms on the screen. A demonstration of the application in use is shown below.

3. Evaluation experiment

3.1. Experimental environment

The evaluation experiment was conducted at the Miyazaki Prefectural Museum. A special exhibition was being held and a place for observation using the application was set up next to the mushroom exhibit. Visitors were asked to cooperate in the evaluation experiment.

3.2. Experimental procedure

The experiment should take no more than 5 minutes. After the observation using the app was over, the participants were asked to fill out a questionnaire. The experiment was divided into four groups using smartphones and tablets, basic apps and apps with additional features. A total of 200 people, 50 in each group, were surveyed. How the experimental procedure is shown.

i. Explain that specimens of glowing mushrooms are on display.
ii. Explain to users how to use the application.
iii. Users use the app to make observations.
iv. Users are free to finish their observations.

Show how the Questionnaire items [Table 2].

Table 2. Questionnaire items

1. Observation was fun.
2. I interested in fungi.
3. New discoveries were made.
4. I understand the structure.
5. It was easy to observe.
6. AR experience was good.
7. Operation was easy.
8. I want to use it in other exhibits.
9. I got the emotion of observing native fungi.
10. I wanted to touch the fungi.
11. Did you notice the hole?
12. Did you notice the rounded bottom?
13. Did you notice the reddish color?
14. Did you notice how the back of the fungus's umbrella is shaped like folds?
15. Did you notice the reticulated underside of the fungus umbrella?

3.3. Questionnaire Results

The questionnaire included questions to answer about the experience of observing the mushrooms using the application and whether or not they noticed any characteristics of the fungi. The questionnaire was based on the MES [5]. Ten questions on a 5-point scale were asked, five questions asking if the mushrooms were observed for their characteristics, and one open-ended question. The results of the questionnaire are shown below [Table 3], [Table 4].

Table 3. Experiences observed using the app (average)

item	Basic apps	Additional apps	Smartphone	Tablet
1	4.9	4.92	4.92	4.9
2	4.57	4.58	4.54	4.61
3	4.72	4.71	4.67	4.76
4	4.51	4.47	4.94	4.48
5	4.76	4.78	4.82	4.72
6	4.91	4.87	4.91	4.88
7	4.75	4.63	4.77	4.61
8	4.8	4.68	4.73	4.74
9	4.44	4.49	4.53	4.4
10	4.45	4.45	4.39	4.51

Table 4. About the characteristics of fungi

item		11	12	13	14	15
Basic apps	Noticed	88 %	68 %	89 %	58 %	66 %
	Didn't notice.	12 %	32 %	11 %	42 %	34 %

Additional apps	Noticed	87%	82%	92%	80%	85%
	Didn't notice.	13%	18%	8%	20%	15%
Smartphone	Noticed	84%	69%	85%	63%	71%
	Didn't notice.	16%	31%	15%	37%	29%
Tablet	Noticed	91%	80%	96%	75%	80%
	Didn't notice.	9%	20%	4%	25%	20%

Since all items were 3 or higher in the experience observed using the app, it can be said that the app is an aid to observation. In comparing the results between the basic app and the app with additional functions, there was a clear difference between Q7 and Q8. The respondents' opinion that simple functions were easier to operate and use was based on the belief that they wanted ease of use rather than in-depth observation. For the question about mushroom characteristics, differences were found at 12, 14, and 15. Apps with additional features for areas that cannot be observed without looking directly at the mushroom from the side, such as the back of the mushroom, were a good result. Comparisons will be made between the use of smartphones and tablets. Differences were found in Q4, 5, 7, 9, and 10. Smartphones can be handled with one hand, making it easier to operate and observe from various angles, which is thought to have led to better understanding of mushrooms. Tablets were found to provide a sense of immersion due to their large screens. In the questions about mushroom characteristics, there was a 10% difference in Q12, 13, and Q14, with tablet users being able to observe mushrooms in more detail.

4. Conclusion

We were able to confirm that AR technology supports observation in museums. The apps with additional functions were found to allow more detailed observation of the mushrooms. As for the difference between the different devices, it was also found that the tablet device allowed for more detailed observation of the mushrooms. However, the basic function apps were better in terms of ease of use, while smartphones were rated higher in terms of ease of observation. Therefore, it is better to use a smartphone when carrying a mobile device for observation when used in a museum. In this way, it is important to use different app functions and mobile devices depending on the situation in which they are used. In this study, we supported the observation of small objects, but in the future we would like to consider supporting the observation of very large objects that are difficult to observe.

References

1. 多田憲孝，"Unity AR Foundation による AR アプリ開発入門"，インプレス R&D(in Japanese), 2021.
2. 橋本直, "AR プログラミング -Processing レシピ-",オーム社(in Japanese), 2012.
3. 伊丹シゲユキ，"入門 Blender 2.9 ~ゼロから始める 3D 制作",秀和システム(in Japanese), 2020.
4. Benjamin，"Blender 2.9 3DCG モデリング・マスター",ソーテック社(in Japanese), 2021.
5. Mohd Kamal Othman, Engaging Visitors in Museums with Technology: Scales for Measurement of Visitor and Mulrimedia Guide Experience, Human-Computer Interaction-INTERACT, 2011, pp. 92-99.

Authors Introduction

Mr. Kakeru Takemura

He is currently enrolled in the Master's course in Mechanical Information Systems, Graduate School of Engineering, University of Miyazaki. His current research theme is museum support using AR technology.

Mr. Ota Hamasuna

He is a master student at Department of Computer Science and System Engineering, University of Miyazaki. His current research interests are augmented reality, computer graphics, and so on

Mr. Fumito Hamakawa

He is a master student at Department of Computer Science and System Engineering, University of Miyazaki. His current research interests are computer graphics, cellular automaton simulation, and so on.

Prof. Satoshi Ikeda

He received PhD degree from Hiroshima University. He is an associate professor in the Faculty of Engineering, University of Miyazaki. His research interest includes graph theory, probabilistic algorithm, fractal geometry and measure theory.

Dr. Kaoru Ohe

2.   She received her Ph.D. degrees from University of Miyazaki, Japan, in 2014. Currently she is an Associate Professor of the Center for Science and Engineering Education, Faculty of Engineering. Her research is separation engineering especially adsorption hazardous heavy metals and oxyanions

Prof. Amane Takei

He is working as Associate Professor for Department of of Electrical and systems Engineering, University of Miyazaki, Japan. His research interest includes high performance computing for computational electromagnetism, iterative methods for the solution of sparse linear systems, domain decomposition methods for large-scale problems. Prof. Takei is a member of IEEE, an expert advisor of IEICE, a delegate of IEEJ, a director of JSST.

Prof. Makoto Sakamoto

He is presently a professor in the Faculty of Engineering, University of Miyazaki. His first interests lay in hydrodynamics and time series analysis, especially the directional wave spectrum. He is a theoretical computer scientist, and his current main research interests are automata theory, languages, and computation. He is also interested in digital geometry, digital image processing, computer vision, computer graphics, virtual reality, etc.

Mr. Shuichi Kurogi

He is a curator of Miyazaki Prefectural Museum of Nature and History. His research interests are wild plants, fungi and lichens in Miyazaki Prefecture. His current research interests are luminous mushrooms, an endangered plant. Member of Japanese Society for Plant Systematics and Mycological Society of Japan.

Automatic Selection of High-Grade Dried Shiitake Mushrooms using Machine Learning

Ota Hamasuna, Kakeru Takemura, Kodai Hasebe, Fumito Hamakawa, Bidesh Biswas Biki
Graduate School of Engineering, University of Miyazaki, Japan

Satoshi Ikeda, Kaoru Ohe, Amane Takei, Makoto Sakamoto*
Faculty of Engineering, University of Miyazaki, Japan

Kazuhide Sugimoto
SUGIMOTO Co., Ltd., Japan

*Corresponding Author

abstract

Miyazaki Prefecture is blessed with the rich nature of the Kyushu Mountains, where many mushrooms have grown wild since ancient times, and the production of dried shiitake mushrooms ranks second in Japan after Oita Prefecture. SUGIMOTO Co., Ltd., which will be cooperating with us this time, is a long-established company established in 1970, and at times it is necessary to sort over 1 ton. However, since selection is still done manually, it is very difficult for employees to do the selection. The goal of this research is to determine the quality of shiitake mushrooms using deep learning and video image processing.

Keywords: image classification, video processing, CNN, OpenCV

1. Introduction

SUGIMOTO Co., Ltd. collects dried shiitake mushrooms directly from approximately 600 producers in Takachiho. Shiitake mushrooms are at their peak in spring and autumn, and at peak times, more than 1 ton of shiitake mushrooms can be brought in a day [1]. However, shiitake mushrooms are still sorted manually, and sorting this much requires a lot of effort. The goal of this research is to use deep learning and video processing to determine the quality of donko, a high-quality dried shiitake mushroom.

2. Development environment

In creating this program, we developed it in the following environment (Table 1). In the experiment, a stand was fixed to a conveyor belt, and the smartphone camera was attached to the end of the stand with the camera facing down.

Table 1 Development environment

OS	Windows10
language	Python
camera	moto g30
software	iVCam
software	OpenCV
	PIL
	NumPy
	Pandas

3. Method

3.1. Target of selection

The following two targets are selected. The one on the left (Fig. 1) is a good shiitake mushroom, and the one on the right (Fig. 2) is a bad one.

Fig. 1 Good item   Fig. 2 Bad item

The following four main examples of bad items are listed below. Fig. 3 shows a state where the pileus of a shiitake mushroom part is broken. Fig. 4 shows a state in which the pileus of a shiitake mushroom is greatly deformed. Fig. 5 shows a state where there are holes caused by insect damage. Fig. 6 shows the pileus of a shiitake mushroom part being rubbed.

Fig. 3 Cracked   Fig. 4 Deformed

Fig. 5 Hole   Fig. 6 Rubbed

Ota Hamasuna, Kakeru Takemura, Kodai Hasebe, Fumito Hamakawa, Bidesh Biswas Biki, Satoshi Ikeda, Kaoru Ohe, Amane Takei, Makoto Sakamoto, Kazuhide Sugimoto,
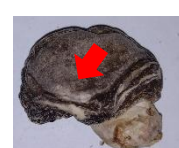
The example of a bad item mentioned this time is the front side with a pileus of a shiitake mushroom on top. Although there are characteristics of bad items on the back side as well, this experiment focused on the front side. The reason why we focused on the front side is that it takes a huge amount of time to collect samples from the front and back sides, and the camera can only capture either the front or back side due to the sorting process on the conveyor belt.

3.2. CNN

In the experiment, CNN was used as an algorithm to determine whether an image was a good item or a bad item. Explain about CNN. CNN (Convolution Neural Network) is a network often used in image recognition research. This CNN is characterized by being constructed by stacking layers with several unique functions, such as convolutional layers and pooling layers [1]. Currently, it is attracting increasing attention as it is being used in a variety of fields. The image in Fig. 7 shows the flow of handwritten digit classification using CNN. First, handwritten digits are given as an input image, the features of the image are extracted by convolution, and the extracted features are aggregated in a pooling layer. Finally, the fully connected layer transforms the output into a one-dimensional form and outputs each selection result as a probability.
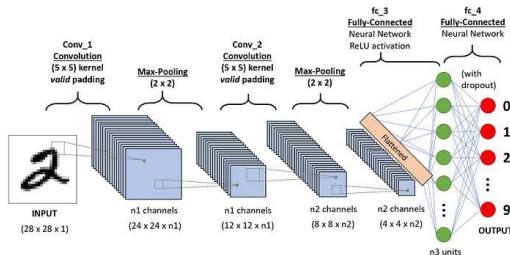
Fig. 7 Classification of handwritten digits [2]

3.3. Video image processing

In this research, we extract still images of shiitake mushrooms from moving images and select them using a model learned by CNN. The video image used in the experiment was taken from directly above the conveyor belt, and the shiitake mushrooms were flowing from left to right. Fig. 8 shows an image of the program being executed.
A still image extracted from a moving image is displayed at the top left of the screen. Extract three still images of the shiitake mushroom using OpenCV. The extraction location is when the shiitake mushrooms flow to the left, center, or right side of the screen. The three images are selected by the model, and the one with a majority of selection results is the final result. At the bottom center of the screen, a "○" is displayed if the item is good, and an "×" is displayed if the item is bad. If the test is in progress, "・・・" is displayed. The number of good

items and bad items is displayed at the top right of the screen. The count is performed when the final result is obtained.

Fig. 8 Program running

4. Experiment content

An evaluation experiment was conducted to verify the usefulness of the developed sorting system. They sort out 17 good and 23 bad mushrooms, and tally up whether the final result matches the shiitake mushrooms that were flushed. The number of matches is defined as the number of correct answers, and the accuracy is defined as follows. Since the final result is updated sequentially until the flow is finished, the last displayed final result is taken as the detection result. The number of cards trained on the CNN was 668 for good and 759 for bad, split 8:2 between training and testing.

$$accuracy = \frac{Number\ of\ correct\ answers}{Number\ of\ shiitake\ mushrooms\ shed} \times 100(\%)$$

5. Experimental result

As a result, the accuracy was low, with 41.1 % of good items and 91.3 % of bad items. Most of the shiitake mushrooms thrown away by the model were sorted out as bad shown in Table 2.

Table 2 Program execution results

	Good item	bad item
number of flows	17	23
The number of correct answers	7	21
accuracy	41.1%	91.3%

The figure below (Fig. 9) shows the selection results of the CNN model in the still image state. Although the accuracy of sorting out bad items as bad items was high, the accuracy in sorting out good items as good items was poor.
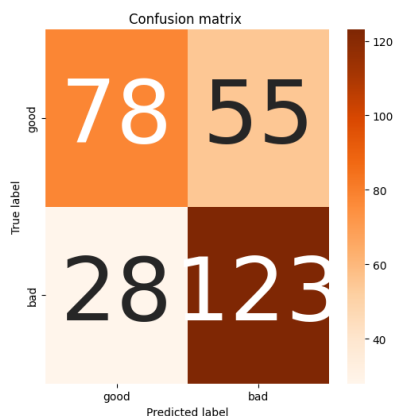
Fig. 9 Results of still image selection using CNN

6. Conclusion

This research is still in its infancy, and the experimental results were very poor. First, there is the problem that it is difficult to distinguish between good items at the CNN learning stage. It was difficult for me to sort by just looking at still images. When actually sorting, not only the front side but also the back side is important. Since this study was an experiment on only the front side, information on the back side was missing, and we believe that the accuracy is reduced. As a result, the accuracy of moving images deteriorated even more than the accuracy of still images. The reasons for the lower accuracy in the video images are that in the case of the video images, the shiitake mushrooms swayed on the conveyor belt, making it impossible to obtain accurate images, and that the images to be trained by the CNN and the video images on the conveyor belt had different lighting and size of the shiitake mushrooms within the frame. Future issues include the following.

- Increasing the number of images for learning
- Collect images of the back side of shiitake mushrooms
- Organizing the environment in which still images are collected and the environment in which video images are taken
- Reconstruction of CNN model

Acknowledgment

I would like to thank my supervisor, Professor Masato Sakamoto, for his guidance in preparing this thesis. We would like to thank Kazuhide Sugimoto of Sugimoto Shoten Co., Ltd. for providing information on shiitake mushrooms for this study.

References

1. AIsmiley 編集部.” 画像認識でよく聞く「CNN」とは？仕組みや特徴を 1 から解説 (in Japanese)” .AISmiley.2023.[Online] https://aismiley.co.jp/ai_news/cnn/,(accessed on 2023/12/5).

2. Saha, S.” A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way “.Medium.2018.[Online]https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53,(accessed on 2023/12/05)

Authors Introduction

Mr. Ota Hamasuna
He is a master student at Department of Computer Science and System Engineering, University of Miyazaki. His current research interests are augmented reality, computer graphics, and so on.

Mr. Kakeru Takemura
Currently enrolled in the Master's course in Mechanical Information Systems, Graduate School of Engineering, University of Miyazaki. His current research theme is museum support using AR technology.

Mr. Kodai Hasebe
He is a master student at Department of Computer Science and System Engineering, University of Miyazaki. His current research interests are image processing, machine learning, and so on.

Mr. Fumito Hamakawa
He is a master student at Department of Computer Science and System Engineering, University of Miyazaki. His current research interests are computer graphics, cellular automaton simulation, and so on.

Mr. Bidesh Biswas Biki
He is master's student in Data Science at Department of Statistics from TU Dortmund. He is an enthusiast of Machine Learning and Deep Learning. He has done some research work and individual projects on supervised learning especially on "Classification" and "Regression" techniques. Also has done some work on Image Processing and NLP.

Ota Hamasuna, Kakeru Takemura, Kodai Hasebe, Fumito Hamakawa, Bidesh Biswas Biki, Satoshi Ikeda, Kaoru Ohe, Amane Takei, Makoto Sakamoto, Kazuhide Sugimoto,

Prof. Satoshi Ikeda

He received PhD degree from Hiroshima University. He is an associate professor in the Faculty of Engineering, University of Miyazaki. His research interest includes graph theory, probabilistic geometry, fractal geometry and measure theory.

Prof. Kaoru Ohe

She received her Ph.D. degrees from University of Miyazaki, Japan, in 2014. Currently she is an Associate Professor of the Center for Science and Engineering Education, Faculty of Engineering. Her research is separation engineering especially adsorption hazardous heavy metals and oxyanions

Prof. Amane Takei

He is working as Associate Professor for Department of Electrical and systems Engineering, University of Miyazaki, Japan. His research interest includes high performance computing for computational electromagnetism, iterative methods for the solution of sparse linear systems, domain decomposition methods for large-scale problems. Prof. Takei is a member of IEEE, an expert advisor of IEICE, a delegate of IEEJ, a director of JSST.

Prof. Makoto Sakamoto

He is presently a professor in the Faculty of Engineering, University of Miyazaki. His first interests lay in hydrodynamics and time series analysis, especially the directional wave spectrum. He is a theoretical computer scientist, and his current main research interests are automata theory, languages and computation. He is also interested in digital geometry, digital image processing, etc.

Mr. Kazuhide Sugimoto

He was Born in Takachiho town, Miyazaki Prefecture. After working in sales in the food service and apparel industries, he joined SUGIMOTO Co., Ltd., a wholesaler of dried shiitake mushrooms produced in Takachiho town, in 2011 after the Great East Japan Earthquake. In response to the current harsh situation, such as aging contract farmers and sluggish demand for dried shiitake mushrooms, he decided to protect the producers by promoting new business development, which he had experienced in sales. New items developed using shiitake mushrooms from Takachiho Township have become a standard item at supermarkets and department stores outside the prefecture. In March 2020, he was appointed as Representative Director, and in 2021, he was selected as a Small and Medium Enterprise Supporter and GFP Ambassador.

Predicting High Volatility Cryptocurrency Prices using Deep Learning

Tsutomu Ito
National Institute of Technology, Ube College, Japan

Kodai Hasebe, Fumito Hamakawa, Bidesh Biswas Biki
Graduate School of Engineering, University of Miyazaki, Japan

Satoshi Ikeda, Amane Takei, Makoto Sakamoto
Faculty of Engineering, University of Miyazaki, Japan

Md Riajuliislam
Department of Statistics, TU Dortmund, Germany

Sabrina Bari Shital
Computer Science and Engineering, Daffodil International University, Bangladesh

Takao Ito*
Hiroshima University, Japan

*Corresponding Author E-mail: itotakao@hiroshima-u.ac.jp>

Abstract

Even if you want to make a profit from cryptocurrency, you are worried that you will lose money, and it is difficult to afford it. There are a vast number of papers that study such unpredictable price fluctuations of cryptocurrency. Currently, it is mainstream to use learning deep to predict the price of cryptocurrency. The goal of this research is to predict the price of cryptocurrency over the long-term using deep learning. The algorithms used are LSTM, GRU, and Bi-LSTM. The targeted cryptocurrencies are Bitcoin, Ethereum, Litecoin, and Cardano. Finally, we will compare it with previous research and verify the performance of our model.

Keywords: Price prediction, Cryptocurrency, LSTM, GRU, Bi-LSTM

1. Introduction

Cryptocurrency is still not very familiar to Japan and may feel like a distant entity, but it is actually something that will affect our lives in the near future. The main uses of Bitcoin are investment and speculative management. Cryptocurrency, which started out as a niche product used by small online groups, has now become mainstream, attracting the attention of both financial professionals and the general public. With a market capitalization reaching billions of dollars, this market has also become a new arena for speculators [1]. Cryptocurrency assets lack intrinsic value, regulatory oversight, lack of fixed-term investment funds, thin order books, short-term investment methods, and collective psychology, causing cryptocurrencies to be highly volatile [2]. Therefore, it is effective to perform technical trading on products with high volatility, and if you trade well, you can generate large profits [3]. Therefore, it is useful to predict the price movements of cryptocurrency using DNN (deep neural network). The goal of this research is to use DNNs such as LSTM, GRU, and Bi-LSTM to predict cryptocurrency such as BTC, ETH, LTC, and ADA with higher accuracy than previous research.

2. Method
2.1. *Cryptocurrency*

Cryptocurrency is a currency that is exchanged only as electronic data, does not have the power of compulsory currency imposed by the state like legal currency, and is primarily used for transactions over the Internet. Since the advent of Bitcoin, which began operations in 2009, derivative cryptocurrency called altcoins have been created one after another, and with the appearance of cryptocurrency exchanges that exchange legal currency and cryptocurrency, holding cryptocurrency has become easier. It spread rapidly [4]. There are various types of cryptocurrencies such as BTC, ETH, LTC, NEM, Ethereum Classic, LISK, etc., and the number is still increasing. The targets to be predicted in this research are

Tsutomu Ito, Kodai Hasebe, Fumito Hamakawa, Bidesh Biswas Biki, Satoshi Ikeda, Amane Takei, Makoto Sakamoto, Md Riajuliislam, Sabrina Bari Shital, Takao Ito

BTC (Bitcoin), ETH (Ethereum), LTC (Litecoin), and ADA (Cardano).

2.2. Algorithm

We will explain the three algorithms used in this research to predict cryptocurrency: LSTM, GRU, and Bi-LSTM. First, I will explain LSTM (Long Short-Term Memory). Although RNN (Recurrent Neural Network), which is a neural network that remembers past states, is suitable for analyzing time-series data, it has not been able to properly process long-term data [5]. LSTM overcomes this weakness and can maintain long context relationships. Data is controlled by three gates centered on the memory cells. The three gates are called forget gate, input gate, and output gate, and each uses a sigmoid function to determine whether or not to pass data. Fig. 1 shows the contents of LSTM.
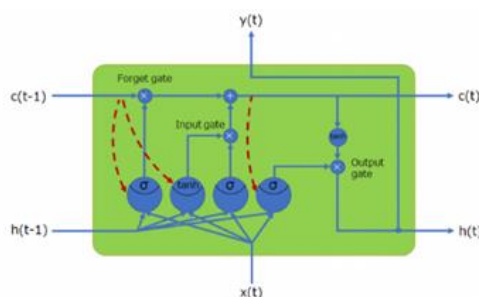
Fig.1. The structure of LSTM [6].

GRU (Gated Recurrent Unit) is said to have similar performance to LSTM, and requires less calculation than LSTM, allowing for faster learning. LSTM is a cell that makes it possible to learn long-term features, which was impossible with RNN, but it has the problem of high computational cost. High calculation costs are not desirable, not only in machine learning. The solution is to eliminate memory cells and reduce the number of gates to reduce computational costs. In LSTM, two states, the storage cell, and the output value, are passed on to the next cell, but GRU combines these states into one. In addition,

LSTM requires one gate controller for input gate, forget gate, and output gate, but in GRU, forget gate and input gate are operated by one controller. Fig. 2 shows the contents of the GRU.
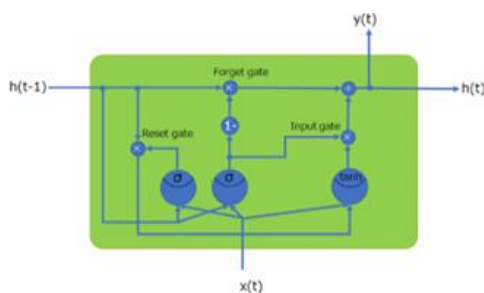
Fig.2. The structure of GRU [6].

The structure of Bi-LSTM (Bidirectional LSTM) ~~has a structure that has~~ with a forward LSTM and a reverse LSTM is shown in Fig. 3. LSTM only flows in the forward direction, but when Bi-LSTM predicts data, for example, given x_n a signal to be predicted, prediction is performed $x_{END} \sim x_{n+1}$ using the Forward Layer in $x_0 \sim x_{n-1}$ the forward direction and the Backward Layer in the reverse direction. It has the characteristic of doing. This allows it to handle more data than LSTM when the same amount of data is used for learning. However, since it handles a large amount of data, it takes longer to process than LSTM [7].
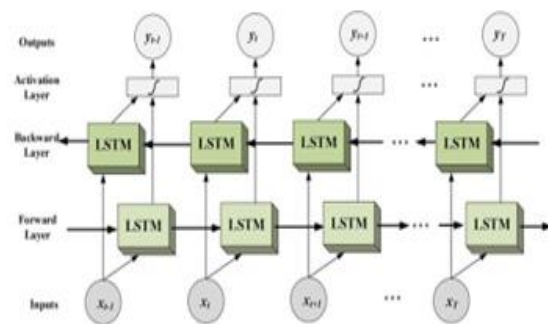
Fig.3. The structure of Bi-LSTM [8].

3. Experiment Content

The conditions for conducting the experiment are described below.

- Three algorithms were used in the experiment: LSTM, GRU, and Bi-LSTM.
- The four cryptocurrencies targeted for prediction are BTC, ETH, LTC, and ADA.
- There are two evaluation indicators: RMSE and MAPE.

3.1. *Development environment*

The model was trained and executed using Google Colab, which allows experiments to be performed for free. The number of CPU cores is 2, the number of threads is 4, the RAM is 12.7GB, and the disk is 225.8GB. The deep learning models LSTM, GRU, and Bi-LSTM used in the experiment were implemented in Python version 3.10.12. We used python libraries such as Sklearn and Keras, which are used for deep learning, and numpy and pandas, which are used for numerical calculations and analysis.

3.2. *Data*

The dataset was collected from Yahoo Finance (https://finance.yahoo.com/). We collected the prices of four types of cryptocurrencies on a daily basis: Bitcoin (BTC), Ethereum (ETH), Litecoin (LTC), and Cardano (ADA). The training data consists of January 1, 2018 to June 8, 2022 (80% of the data), and the test data consists of June 9, 2022 to November 30, 2023 (20% of the data). ~~be done. Split the~~ The data proportion of 80:20 for training and testing is designed in this research. The

detailed information of data specifications is listed in Table 1.

Table 1. Data specifications.

Attribute	Explanation	Type
Date	Date of transaction	Date
Open	First traded price	Continuous
High	Highest traded price	Continuous
Low	Lowest traded price	Continuous
Close	Last traded price	Continuous
Adj Close	The closing price before the split is the adjusted price after the split.	Continuous
Volume	Quantity of trades completed during the period	Continuous

3.3. *Evaluation index*

Root-Mean-Square Error (RMSE) and Mean-Absolute-Percentage error (MAPE) are used to evaluate the algorithm in this study. The smaller the RMSE and MAPE values, the better the prediction model performance. In Eq. (1) and Eq. (2), n is the number of elements, y_i is the correct value, and $\hat{y}_i$ is the predicted value.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad (1)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \qquad (2)$$

4. Experimental result

Table 2 shows the results predicted according to the conditions in Chapter 3. Evaluation values are rounded down to the fifth decimal place. Bold letters in the evaluation value indicate the lowest value (the highest value in terms of accuracy). The results showed that the accuracy of Bi-LSTM was better than other algorithms for all cryptocurrencies.

Table 2. Performance results of this research model.

Currencies	Methods	RMSE	MAPE
BTC	LSTM	807.7135	0.0332
	GRU	873.7594	0.0362
	Bi-LSTM	**770.3047**	**0.0314**
ETH	LSTM	21.0737	0.0123
	GRU	33.0912	0.0192
	Bi-LSTM	**15.1817**	**0.0092**
LTC	LSTM	1.5991	0.0188
	GRU	1.7605	0.0207
	Bi-LSTM	**1.1928**	**0.0130**
ADA	LSTM	0.0203	0.0578
	GRU	0.0236	0.0695
	Bi-LSTM	**0.0172**	**0.0526**

5. Comparative verification

In the comparative verification, we compare the performance of the model for predicting the price of cryptocurrency in this study with other models in previous studies [9], [10]. Here, we will verify whether the model proposed in this study is effective. When making comparisons, we will verify that the period of experimentation is consistent with previous research. Table 3 is a comparison between this research model and the previous study [9] model, and Table 4 is a comparison between this study and the previous study [10]. The accuracy of Bi-LSTM in this study was good in all comparisons, indicating that the model in this study is effective.

Table 3 Comparison of this study and previous study [9] .

Currency	Studies	Methods	RMSE	MAPE
BTC	Our model	LSTM	1184.7059	0.0426
		GRU	1094.3126	0.0422
		Bi-LSTM	**647.2073**	**0.0221**
	[9]	LSTM	1447.648	0.03059
		ARIMA	1288.5	0.03479
		SARIMA	1802.31	0.04665

Table 4 Comparison of this study and previous study [10] .

Currency	Studies	Methods	RMSE	MAPE
BTC	Our model	LSTM	836.2219	0.0318
		GRU	901.6092	0.0312
		Bi-LSTM	**752.4224**	**0.0290**
	[10]	LSTM	1031.340	0.0397
		GRU	1274.171	0.057
		Bi-LSTM	1029.362	0.036
ETH	Our model	LSTM	59.4498	0.0339
		GRU	51.6724	0.0293
		Bi-LSTM	**45.7691**	**0.0253**
	[10]	LSTM	148.522	0.297
		GRU	98.314	0.148
		Bi-LSTM	83.953	0.124
LTC	Our model	LSTM	3.620	0.049
		GRU	5.453	0.078
		Bi-LSTM	**1.988**	**0.020**
	[10]	LSTM	9.668	0.064
		GRU	8.122	0.046
		Bi-LSTM	8.025	0.041

6. Conclusion

In this study, we used three algorithms, LSTM, GRU, and Bi-LSTM, to predict the prices of four cryptocurrency: BTC, ETH, LTC, and ADA. The model was evaluated using the evaluation indicators RMSE and MAPE. The research results showed that Bi-LSTM made the most accurate predictions for all cryptocurrency, followed by LSTM and GRU. Finally, we conducted a comparative verification with previous studies and showed that this study is more effective than previous studies. Future works could include applying reinforcement learning to previous studies [11] and predicting price fluctuations based on the number of tweets [12].

Tsutomu Ito, Kodai Hasebe, Fumito Hamakawa, Bidesh Biswas Biki, Satoshi Ikeda, Amane Takei, Makoto Sakamoto, Md Riajuliislam, Sabrina Bari Shital, Takao Ito

References

1. Bergsli, L. et al. "Forecasting volatility of Bitcoin". Research in International Business and Finance. 2022, vol.59, pp1-30. (accessed on 2023-11-16) [Online]https://doi.org/10.1016/j.ribaf.2021.101540

2. S, Ninomiya. "専門的見解：なぜ仮想通貨市場はボラティリティが高いのか (in Japanese)" CoinPost.2018, [Online] https://qr.paps.jp/Xsj3v (accessed on 2023/12/07)

3. Hideki, N. "強化学習を利用した暗号資産の価格予測 (in Japanese)". Master's thesis, Department of Data Science, Graduate School of Data Science, Shiga University. 2022. https://shiga-u.repo.nii.ac.jp/record/14567 /files/%E4%BF%AE%E5%A3%AB%E8%AB%96%E6%96%87%E5%85%A8%E6%96%87R4nagasawa.pdf (accessed on 2023-11-8)

4. CoinDesk. "暗号資産／仮想通貨とは何か？ クリプトとは？ 初心者にもわかりやすく解説 (in Japanese)". CoinDesk. 2023. [Online]https://www.coindeskjapan.com/learn/cryptocurrency/ (accessed on 2023-11-9)

5. Yusuke, T. "RNN と LSTM(Long Short Term Memory)の違いと特徴 (in Japanese)". AUC+X.2019. [Online]https://www.acceluniverse.com/blog/developers/2019/07/lstm.html (accessed on 2023- 11-9)

6. AFIRobotsBlog. "直感で理解する LSTM・GRU 入門 - 機械学習の基礎をマスターしよう！(in Japanese)". 2023. [Online]https://developers.agirobots.com/jp/lstmgruentrance-noformula/ (accessed on 2023-11-9)

7. Kai, O. "ディープラーニングを用いた低 SNR 下での生体信号抽出 In vivo Signal Extraction Under Low SNR by Deep Learning (in Japanese) ". 338th Research Meeting of the Tohoku Branch of the Society of Instrument and Control Engineers, vol.4, pp1-8. https://www.topic.ad.jp/sice/htdocs/papers/338/338-4.pdf(accessed on 2023-11-9)

8. Tracyrenee." What is the difference between Tensorflow's LSTM and BiLSTM?".2023. https://pub.aimind.so/what-is-the-difference-between-tensorflows-lstm-and-bilstm-d7b1f5d8c772(accessed on 2023- 11-9)

9. Dinshaw, C. et al. "Statistical Scrutiny of the Prediction Capability of Different Time Series Machine Learning Models in Forecasting Bitcoin Prices". 2022 IEEE 4th International Conference on Cybernetics, Cognition and Machine Learning Applications (ICCCMLA), Goa, India, 2022, pp. 329-336, doi: 10.1109/ICCCMLA56841.2022.9989057. (accessed on 2023-11-16)

10. Seabe, P. et al. "Forecasting Cryptocurrency Prices Using LSTM, GRU, and Bi-Directional LSTM: A Deep Learning Approach". Fractal and Fractional 7, no. 2: 203. [Online]https://doi. org/10.3390/fractalfract7020203 (accessed on 2023-11-16)

11. Hideki, N. "強化学習を利用した暗号資産の価格予測 (in Japanese)". Department of Data Science, Graduate School of Data Science, Shiga University. 2023. https://shiga-u.repo.nii.ac.jp/records/14567 (accessed on 2023-11-16)

12. Abraham, J. et al."Cryptocurrency Price Pr ency Price Prediction Using T ediction Using Tweet Volumes and olumes and Sentiment Analysis". SMU Data Science Review, 2018, vol.1, no.3. [Online]https://scholar.smu.edu/datasciencereview/vol1/iss3/1 (accessed on 2023-11-16)

Authors Introduction

Prof. Tsutomu Ito

He is Associate Professor of the Department of Business Administration at National Institute of Technology, Ube College, Japan. His current research interests include internet of things (IoT), mechanical engineering, artificial intelligence (AI), automata theory, quantitative analysis of Japanese Keiretsu. Dr. Ito earned his doctor degree of Engineering from Hiroshima University, Japan in 2018.

Mr. Kodai Hasebe

He is a master student at Department of Computer Science and System Engineering, University of Miyazaki.His current research interests are image processing, machine learning, and so on.

Mr. Fumito Hamakawa

He is a master student at Department of Computer Science and System Engineering, University of Miyazaki. His current research interests are computer graphics, cellular automaton simulation, and so on.

Mr. Bidesh Biswas Biki

He is master's student in Data Science at Department of Statistics from TU Dortmund. He is an enthusiast of Machine Learning and Deep Learning. I have done some research work and individual projects on supervised learning especially on "Classification" and "Regression" techniques. Also has done some work on Image Processing and NLP.

Prof. Satoshi Ikeda

He received PhD degree from Hiroshima University. He is an associate professor in the Faculty of Engineering, University of Miyazaki. His research interest includes graph theory, probabilistic algorithm, fractal geometry and measure theory.

Prof. Amane Takei

He is working as Associate Professor for Department of Electrical and systems Engineering, University of Miyazaki, Japan. His research interest includes high performance computing for computational electromagnetism, iterative methods for the solution of sparse linear systems, etc. Prof. Takei is a member of IEEE, an expert advisor of IEICE, a delegate of the Kyushu branch of IEEJ, a director of JSST.

Prof. Makoto Sakamoto

He is presently a professor in the Faculty of Engineering, University of Miyazaki. His first interests lay in hydrodynamics and time series analysis, especially the directional wave spectrum. He is a theoretical computer scientist, and his current main research interests are automata theory, languages, and computation. He is also interested in digital geometry, digital image processing, computer vision, etc.

Mr. Md Riajulii

He is master's student in Data Science at Department of Statistics from TU Dortmund. He is an enthusiast of Machine Learning and Deep Learning. He has done some research work and individual projects on supervised learning especially on "Classification" and "Regression" techniques. Also has done some work on Image Processing and NLP.

Ms. Sabrina Bari Shital

She is a computer engineer with a background in Computer Science and Engineering from Daffodil International University. Her expertise encompasses machine learning, deep learning, and databases. She is eager to contribute her knowledge to the evolving landscape of technology.

Prof. Takao Ito

He is Professor of Management of Technology (MOT) in Graduate School of Engineering at Hiroshima University. His current research interests include automata theory, artificial intelligence, systems control, quantitative analysis of interfirm relationships using graph theory, and engineering approach of organizational structures using complex systems theory.

Simulation of Weathering Representation using Vertex and UV Information

Tsutomu Ito
National Institute of Technology, Ube College, Japan

Fumito Hamakawa, Kodai Hasebe
Graduate School of Engineering, University of Miyazaki, Japan

Satoshi Ikeda, Amane Takei, Makoto Sakamoto
Faculty of Engineering, University of Miyazaki, Japan

Takao Ito*
Hiroshima University, Japan

*Corresponding Author E-mail: itotakao@hiroshima-u.ac.jp>

Abstract

In recent years, three-dimensional computer graphics (3DCG) technology has been developed. In addition, much research has been done on weathering representations such as rust and moss for realistic representation. However, when outputting simulation results on an image, a large number of images are required to display different simulation results in a three-dimensional space. In this paper, a simulation method using vertex information of a 3D model and simple images is proposed. In this method, when the number of vertices is sufficient, the simulation result output does not use images, thus reducing the data increase. This figure is one of the experimental results.

Keywords: Computer graphics, Vertices, Rust, Weathering, Visual simulation

1. Introduction

In recent years, three-dimensional computer graphics (3DCG) have been used in films, games, VR-based education, and medicine. In this context, there is a need to make 3DCG look more realistic. However, it is not subject to interferences such as environment or time and therefore does not degrade as in reality. This is why research into weathering expressions for CG is being conducted to this day.

This study proposes a method to implement weathering representations such as rust and moss using vertex information and UVs, instead of outputting simulation results as images for textures. The simulation results were also applied to a 3DCG model in practice.

2. Research Background

To date, a plethora of research has been done on weathering representations related to computer graphics. However, most of them require the simulation results to be output as images for textures. However, when simulations are performed on multiple models, images are required for the number of models, resulting in a large amount of data. With the aim of being flexible enough to deal with existing models and real-time representations, a simulation method using vertex information and UVs of 3DCG models was developed in this study.

In this study, the UV map used for the material representation of the original 3DCG model is called the base UV map, and the UV map for simulation is called the simulation UV map.

Tsutomu Ito, Fumito Hamakawa, Kodai Hasebe, Satoshi Ikeda, Amane Takei, Makoto Sakamoto, Takao Ito

3. Development Environment

The development environment is shown in Table 1, with C# used as the programming language for Unity and Python for Blender.

Table 1. Development Environment.

OS	Windows10
Software	Blender 3.6
	Unity 2022.3.14fl
	Visual Studio 2019
	Adobe Illustrator 2023

4. Simulation Method

4.1. *Adjacency list*

In this study, simulations are performed using vertex information in 3DCG. Vertex adjacency information was obtained, and an adjacency list was created to represent the phenomenon of rust and moss propagating locally in weathering simulations. The method for creating the neighborhood list used vertex information constituting polygons and edges, although the method of searching for neighborhood information differs from software to software. The adjacency list of vertexes v_i at index number i is Fig. 1.

vertex	adjacent to
v_0	v_1, v_2
v_1	v_0, v_2
v_2	v_0, v_1
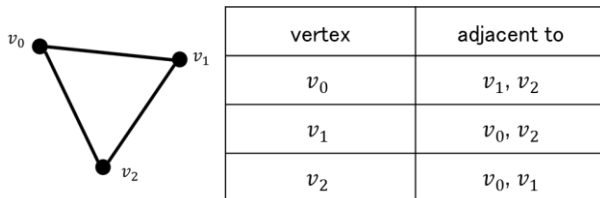
Fig. 1. Adjacency list at vertex.

4.2. *Initial condition setting*

The susceptibility to rust and moss varies with the environment, such as humidity and sunlight, but also with the material of the object, just as there is a difference in the probability of rust between metal and plastic. Few previous studies have considered the material of the object and conducted simulations, and most of them assume that the entire object is made of the same material. In this study, the user can control the erosion probability of the initial state like a mask image by painting directly using the base UV map.

4.3. *Dynamic environment acquisition*

Most of the previous studies assume a situation where the object is in the same environment and in the same shape at every step. However, some objects may change dynamically, such as changing orientation over time or increasing obstacles. Therefore, in this study, by incorporating vertex information into the local transition rules, real-time changes such as changes in orientation and up/down can be managed.

4.4. *UV-based erosion representation*

Determine UV coordinates from the degree of rust and moss erosion at the vertices using the adjacency list. The color of the eroded object is also determined by the user entering a gradient image. In addition, as multiple UV maps are used, they can be used in conjunction with the base colors of existing 3DCG models as shown in Fig. 2 and Fig. 3.
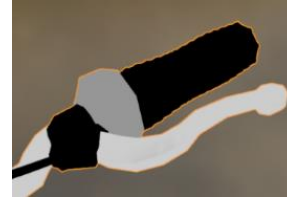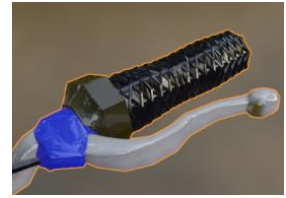
Fig. 2. Base color.　　Fig. 3. Initial condition.

5. Experimental Results

This paper describes some of the experimental results, mainly related to rust formation. The transition conditions for rust generation are as follows, based on previous studies [1], [2], [3].

- Rust progression from 0 to the maximum rust value R_{MAX} determined by the user.
- Rust progression is assigned to each vertex.
- Vertex v as the initial rust occurrence has a low probability of rust progression R_v increases by 1.
- Vertex adjacent to a vertex with progression greater than 1 has its R_v increased by 1 by a function that takes gravity and probability into account.
- Vertexes above a certain level of progression are peeled off.
- UV placement of vertex v at index number i in step t is determined by $R_{v_i}^t / R_{MAX}$

In this experiment, simulated UVs are applied to the mask textures in Fig. 4 and the color textures in Fig. 5. The mask texture is a grayscale image and can be used to express textures such as the metallic feel and roughness of the surface of the 3DCG model by inputting the simulation results into software shaders. Besides, it can

also be used to control the range of application of noise textures to represent uneven and rusty surfaces or the base color of an existing 3DCG model. Color textures represent color scales. In addition, alpha values can be set, and the expression of peeling and holes above a certain value is implemented.

Fig. 4. Mask texture.　　Fig. 5. Color texture.

5.1. *Base experimental*

In this experiment, simulations were performed on a cube that was tilted. The results, shown in Fig. 6 and Fig. 7, show rust propagating according to gravity. In addition, the strongly eroded areas have holes and peeling. The rust-free area in Fig. 6 has no surface irregularity due to the noise texture in Fig. 7, indicating that the mask texture controls the texture of the 3DCG model.
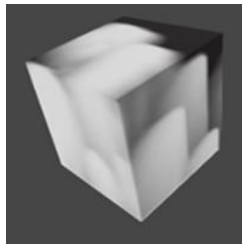
Fig. 6. Ex1 Mask result.　　Fig. 7. Ex1 Color result.
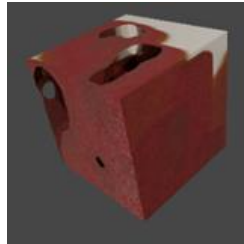
5.2. *Verification of randomness*

In this experiment, the same 3DCG model, the same transition rules, and the same textures as in 5.1 were used in the simulation. The results in Fig. 8 and Fig. 9 differ from those in 5.1, indicating that the same image and the same model can output different patterns of results depending on the probability.
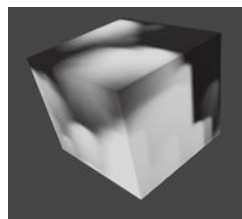
Fig. 8. Ex2 Mask result.　　Fig. 9. Ex2 Color result.
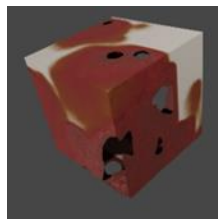
5.3. *Shape change*

In this experiment, only the tilt of the object was changed from 5.1 to the simulation. 5.1 shows that rust

propagates in the left front direction because the object is tilted in the left front direction. However, in this experiment, the rust propagated in the right-most direction as the object was tilted in the right-most direction as illustrated in Fig. 10 and Fig. 11.
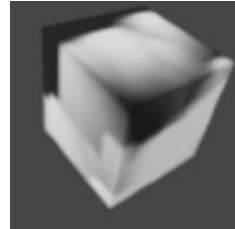
Fig. 10. Mask result in　　Fig. 11. Color result in
different shapes.　　　　　different shapes.

5.4. *Initial condition*

In this experiment, the objects and initial conditions were changed and simulated. The initial state is shown in Fig. 3. Fig. 12 and Fig. 13 show that rust only occurs in the most rust-prone white area in the initial state, and no rust occurs in the most rust-prone black area.

Fig. 12. Mask result in　　Fig. 13. Color result in
another object.　　　　　another object.

6. Consideration

From 5.1 and 5.2, it was shown that rust propagates in the direction of gravity depending on probability as in previous studies, and that the location of rust and the direction of rust propagation change depending on probability. This indicates that the same mask texture and color texture can output different patterns of results for the same object of the same shape existing at the same time. This means that the simulation results do not need to be output to an image, which may reduce the amount of data when multiple patterns of weathering representation are desired for the same object. In addition, since the simulation results are the erosion degree for each vertex, they can be saved by outputting them to a CSV file or other format.

From 5.3, it was shown that the simulation results can be output in response to tilt even when the object shape is changed. This is because the vertex information is

Tsutomu Ito, Fumito Hamakawa, Kodai Hasebe, Satoshi Ikeda, Amane Takei, Makoto Sakamoto, Takao Ito

referenced in the transition rule, and it can respond to changes in the tilt and size of the object during simulation. In addition to vertex position information, vertex distance and the number of neighbors can also be referenced. However, if the neighbor information changes, such as when a vertex is deleted or added, it is necessary to reacquire the neighbor information.

From 5.4, it is shown that it is possible to limit the impact of the simulation by setting initial conditions. This allows the base UV map and simple painting to control the strength of the influence of different metals and materials, without the In this study, only rust was mentioned, but by changing the texture and transition rules for each, the propagation of moss can be represented as shown in Fig. 14.

Fig. 14. Simulation results assuming moss in another object.

7. Conclusion

In this study, we performed weathering simulations using vertex adjacency and UV information. As in the

previous study, expressions such as rust and moss propagation could be made without outputting them to the texture. However, the number of vertices needs to be increased for more realistic expressions and detailed patterns. Therefore, if real-time changes are to be observed, it is necessary to optimize the number of vertices and to limit the number of objects to be simulated at the same time.

For future works, it is necessary to pursue the improvement of transition rules and the fields in which this vertex technology can be utilized.

References

[1] 鈴木雄太ほか,"3D オートマトンを用いた鉄の腐食による形状変化",2017 年映像情報メディア学会年次大会(in Japanese), 21C-2, 2 pages (2017).

[2] 田邉竜馬ほか,"本錆と副錆の挙動を考慮した錆の経年変化テクスチャ生成", ITE Technical Report(in Japanese), Vol.38, No.16, pp.95-98 (2014).

[3] 金澤功尚ほか,"オブジェクトの形状を考慮した錆によるエイジング画像シミュレーション",画像電子学会誌 Vol.46, No.4, pp.547-558 (in Japanese) (2017).

Authors Introduction

Prof. Tsutomu Ito

He is Associate Professor of the Department of Business Administration at National Institute of Technology, Ube College, Japan. His current research interests include internet of things (IoT), mechanical engineering, artificial intelligence (AI), automata theory, quantitative analysis of Japanese Keiretsu. Dr. Ito earned his doctor degree of Engineering from Hiroshima University, Japan in 2018.

Mr. Fumito Hamakawa

He is a master student at Department of Computer Science and System Engineering, University of Miyazaki. His current research interests are computer graphics, cellular automaton simulation, and so on.

Mr. Kodai Hasebe
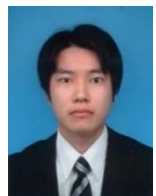
He is a master student at Department of Computer Science and System Engineering, University of Miyazaki. His current research interests are image processing, machine learning, and so on.

Prof. Satoshi Ikeda

He received PhD degree from Hiroshima University. He is an associate professor in the Faculty of Engineering, University of Miyazaki. His research interest includes graph theory, probabilistic algorithm, fractal geometry and measure theory.

Prof. Amane Takei

He is working as Associate Professor for Department of Electrical and systems Engineering, University of Miyazaki, Japan. His research interest includes high performance computing for computational electromagnetism, iterative methods for the solution of sparse linear systems, domain decomposition methods for large-scale problems. Prof. Takei is a member of IEEE, an expert advisor of IEICE, a delegate of the Kyushu branch of IEEJ a director of JSST.

Prof. Makoto Sakamoto

He is presently a professor in the Faculty of Engineering, University of Miyazaki. His first interests lay in hydrodynamics and time series analysis, especially the directional wave spectrum. He is a theoretical computer scientist, and his current main research interests are automata theory, languages, and computation. He is also interested in digital geometry, digital image processing, computer vision, computer graphics, virtual reality, etc.

Prof. Takao Ito

He is Professor of Management of Technology (MOT) in Graduate School of Engineering at Hiroshima University. His current research interests include automata theory, artificial intelligence, systems control, quantitative analysis of interfirm relationships using graph theory, and engineering approach of organizational structures using complex systems theory.

Investigating the Engineering Interventions in the Conservation of Malaysia Heritage Structures: A Review on Preserving Historical Edifices Through Advanced Civil Engineering Techniques.

Muhammad Azizi Azizan
Faculty of Civil Engineering & Technology, Universiti Malaysia Perlis, Perlis, Malaysia
Centre of Excellence for Unmanned Aerial System, Universiti Malaysia Perlis, Perlis, Malaysia

Nurfadzillah Ishak
Faculty of Civil Engineering & Technology, Universiti Malaysia Perlis, Perlis, Malaysia
Centre of Excellence for Unmanned Aerial System, Universiti Malaysia Perlis, Perlis, Malaysia

Hazry Desa
Faculty of Electrical Engineering & Technology, Universiti Malaysia Perlis, Perlis, Malaysia
Centre of Excellence for Unmanned Aerial System, Universiti Malaysia Perlis, Perlis, Malaysia

E-mail: aziziazizan@unimap.edu.my

Abstract

This review delves into the realm of Malaysia's heritage conservation, spotlighting the transformative impact of advanced civil engineering techniques. Through the integration of Fiber-Reinforced Polymers (FRP) and nanotechnology, historical edifices are fortified, seamlessly blending modern engineering with architectural elegance. Non-destructive testing (NDT) methods, including ground-penetrating radar and Finite Element Analysis (FEA), empower conservationists with deep insights into structural intricacies, guiding targeted interventions. In the digital sphere, 3D laser scanning captures intricate details, while Virtual Reality (VR) simulations facilitate immersive exploration and informed decision-making. Beyond preservation, these technologies foster public engagement, ensuring a collective understanding of Malaysia's cultural heritage. This harmonious fusion of tradition and cutting-edge engineering ensures the enduring legacy of Malaysia's architectural treasures.

Keywords: Heritage conservation, Non-destructive testing (NDT), Finite Element Analysis (FEA), 3D laser scanning, Virtual Reality (VR)

1. Introduction

Heritage structures stand as invaluable remnants of our cultural legacy, embodying the architectural achievements and historical significance of bygone eras [1]. However, the preservation of these structures poses intricate challenges that demand innovative engineering solutions [2], [14]. As urban landscapes evolve and modernization accelerates, the need to safeguard these historical edifices has never been more pressing [3]. This review embarks on a comprehensive exploration of the engineering interventions essential in the conservation of heritage structures, delving into the intricate amalgamation of tradition and technology [4], [6]. Through a meticulous analysis of advanced civil engineering techniques, this study aims to shed light on the nuanced methods employed to preserve the structural integrity, cultural heritage, and aesthetic essence of these buildings. By examining the successes, challenges, and lessons learned from previous conservation projects, this research endeavors to offer valuable insights into the sustainable preservation of our architectural heritage, ensuring that these treasures endure for generations to come. In this pursuit, study will navigate the intersection of history, culture, and engineering prowess, striving to strike an intricate balance between preservation imperatives and contemporary demands [5].

2. Preserving Historical Edifices in Malaysia

Conducting a thorough exploration of Malaysia's rich historical and cultural heritage is vital. This involves studying the evolution of architectural styles and understanding the cultural significance of historical edifices within Malaysian society [6], [23]. The architectural diversity in Malaysia, influenced by Malay, Chinese, Indian, and colonial elements, forms a critical backdrop for conservation efforts [7]. Malaysia's architectural heritage is a tapestry woven from centuries of cultural interactions and influences. The evolution of

architectural styles in Malaysia reflects the rich history of the nation, characterized by the convergence of Malay, Chinese, Indian, and colonial elements. Traditional Malay architecture, exemplified by intricate wooden carvings and steeply pitched roofs, showcases indigenous craftsmanship [8], [19]. The Chinese influence is evident in the intricate details of temples and clan houses, emphasizing symmetry and ornate decorations [9]. Indian architectural traditions, characterized by vibrant colors and sculpted facades, have left an indelible mark on religious structures across the country [10], [26]. Additionally, the colonial era introduced a blend of European styles, seen in buildings with ornate facades, arched windows, and grand pillars, reflecting the British, Dutch, and Portuguese colonial influences [11].

Understanding the unique architectural amalgamation is crucial in preserving historical edifices. Each cultural influence represents a chapter in Malaysia's history, encapsulating the nation's multicultural identity. Preservation efforts must delicately balance these diverse elements, ensuring that the architectural heritage reflects the harmonious coexistence of cultures within Malaysian society [12]. The selection process involves collaboration with local heritage organizations and experts. It's crucial to choose diverse edifices, considering regional representation and architectural styles. Identifying overlooked structures that are significant to local communities ensures a holistic approach to preservation [13].

3. Advanced Civil Engineering Techniques in Conservations Work

The preservation of historical edifices necessitates the integration of advanced civil engineering techniques, marking a paradigm shift in heritage conservation practices. Traditional methods, while valuable, often face limitations in addressing the complex challenges posed by aging structures and evolving environmental conditions. Advanced civil engineering techniques encompass a spectrum of innovative approaches, from cutting-edge materials science to sophisticated computational modeling and non-destructive testing methods. These techniques play a pivotal role in the structural assessment, restoration, and long-term sustainability of historical buildings, ensuring the preservation of cultural heritage for future generations.

3.1 Advanced Structural Assessment

Non-Destructive Testing (NDT) is a techniques like ground-penetrating radar (GPR) and ultrasonic testing enable engineers to assess internal structures without damaging the building fabric [14]. Ground-penetrating radar (GPR) operates on the principle of emitting high-frequency radio waves into the structure. These waves penetrate the materials and reflect back when encountering boundaries between different materials or voids. By analyzing the time taken for the signals to return and their strength, engineers can create detailed subsurface images without invasive measures.

Finite Element Analysis (FEA) is another civil engineering computational methods such as FEA provide detailed insights into stress distributions and structural vulnerabilities, aiding in targeted interventions [2], [15]. FEA tools enable engineers to visualize the impact of interventions on the structure's aesthetics. Conservationists can assess how alterations might affect the visual aspects of the edifice, ensuring that preservation efforts align with historical significance and architectural beauty. FEA can validate restoration efforts. By comparing the digital model of the restored structure with historical data, conservationists can ensure that the restoration accurately reflects the original design, maintaining the structure's historical authenticity [16]. Incorporating Finite Element Analysis into heritage conservation efforts empowers engineers and conservationists with detailed, data-driven insights. It enables targeted, efficient interventions while preserving the historical and aesthetic essence of these invaluable structures.

3.2 Innovative Materials and Techniques

Fiber-Reinforced Polymers (FRP) composites offer lightweight and high-strength solutions for structural reinforcement, preserving the original aesthetics [17]. Fiber-Reinforced Polymers (FRP) are composite materials made of a polymer matrix reinforced with fibers such as glass, carbon, or aramid. These materials offer exceptional strength-to-weight ratios and durability, making them ideal for structural reinforcement applications. FRP composites are used to reinforce historical structures without significantly altering their original appearance. For instance, FRP strips can be bonded to beams and columns, providing additional strength without compromising the aesthetic integrity [18]. In earthquake-prone regions, FRP materials are employed to enhance the seismic resilience of historical buildings. They can be strategically placed to reinforce vulnerable areas, ensuring the structure's stability during seismic events [19].

Nanotechnology Applications Nanomaterials, such as self-healing coatings, enhance the durability and resilience of historical structures against environmental

factors [20]. Nanotechnology involves the manipulation of materials at the nanoscale, where their properties exhibit unique characteristics. In heritage conservation, nanomaterials are utilized for their ability to enhance the durability, resilience, and protective qualities of historical structures. Nanomaterials, such as microcapsules containing healing agents, are embedded in coatings applied to building surfaces. When the structure experiences minor cracks due to environmental stresses, these capsules rupture, releasing the healing agents and effectively repairing the damage [21]. Hydrophobic nano coatings repel water and prevent moisture infiltration, safeguarding historical structures from water-related deterioration. These coatings can be applied to various surfaces, including stone, wood, and metal, preserving the integrity of the building materials [22].

The integration of FRP composites and nanotechnology applications represents a forward-thinking approach in heritage conservation, combining structural reinforcement with advanced protective coatings. These techniques not only strengthen historical structures but also contribute to their long-term resilience against various environmental challenges, ensuring the preservation of cultural heritage assets.

3.3 Preservation through Digital Technologies

3D Laser Scanning: High-resolution 3D scans aid in documenting intricate architectural details, facilitating accurate restoration [23]. 3D laser scanning involves using lasers to create highly detailed, three-dimensional representations of physical objects, including historical buildings. Laser scanners emit laser beams that bounce off surfaces and return to the scanner. By measuring the time, it takes for the laser to return, the scanner can create precise 3D models. 3D laser scanning captures intricate architectural details, including ornate carvings, moldings, and textures, with unparalleled accuracy. This documentation serves as a digital record of the structure, aiding in historical analysis and preservation efforts [24]. Laser scanning provides precise measurements, allowing conservationists to assess dimensions, angles, and spatial relationships within the building. These measurements are crucial for restoration work, ensuring that new elements align perfectly with the original design [25]. 3D models generated through scanning enable virtual preservation. Conservationists can digitally archive buildings, allowing future generations to explore and study historical structures in virtual environments [26].

Virtual Reality (VR) Simulations: VR technologies allow conservationists to visualize proposed changes and assess their impact on the historical context [1]. Virtual Reality (VR) simulations immerse users in computer-generated environments. Conservationists use VR technologies to create virtual replicas of historical structures, enabling interactive exploration and analysis. Conservationists use VR simulations to visualize proposed alterations or restoration efforts. This immersive experience allows stakeholders to assess the visual impact of changes before implementation, ensuring that modifications align with the historical context below. VR simulations enhance public engagement by allowing virtual tours of historical sites. This interactive experience fosters public interest, awareness, and appreciation for cultural heritage, promoting advocacy for conservation efforts [28]. VR simulations serve as educational tools, enabling students, architects, and conservationists to virtually explore historical buildings.

The integration of advanced civil engineering techniques in heritage conservation signifies a progressive approach toward preserving our architectural legacy. By leveraging cutting-edge methods and materials, conservationists can ensure the structural integrity, authenticity, and longevity of historical edifices, contributing significantly to the safeguarding of cultural heritage for generations to come. This evolving field continues to redefine the boundaries of conservation practices, marking a new era in the intersection of engineering innovation and cultural preservation.

4. Conclusion

This comprehensive exploration into the conservation of Malaysia's rich heritage structures illuminates the pivotal role played by advanced civil engineering techniques in preserving these historical gems. Fiber-Reinforced Polymers (FRP) emerge as lightweight, yet robust solutions, seamlessly merging modern engineering with historical aesthetics. Collaborating with nanotechnology, self-healing coatings and hydrophobic nano coatings fortify structures against environmental ravages, ensuring their endurance for posterity.

Non-destructive testing (NDT) methods like ground-penetrating radar (GPR) and ultrasonic testing, coupled with Finite Element Analysis (FEA), provide profound insights into the structural intricacies of historical edifices. These techniques empower conservationists with precise stress distributions, vulnerability identifications, and predictive capabilities, guiding meticulous interventions that preserve both the architectural integrity and historical

Muhammad Azizi Azizan, Nurfadzillah Ishak, Hazry Desa

resonance.

In the digital realm, 3D laser scanning meticulously captures architectural nuances, laying the foundation for precise restoration. Virtual Reality (VR) simulations then breathe life into historical spaces, enabling immersive exploration and informed decision-making. Beyond conservation, these technologies foster public engagement, ensuring a collective understanding and appreciation of Malaysia's cultural heritage.

As custodians of Malaysia's architectural legacy, conservationists stand at the nexus of tradition and technology. Armed with these advanced tools, they meticulously balance preservation imperatives and contemporary demands. Through this synergy, Malaysia's heritage structures not only withstand the tests of time but also continue to inspire, educate, and enrich future generations, ensuring the vibrant preservation of the nation's cultural heritage.

5. Acknowledgement

References

[1] Volk, R., Stengel, J., & Schultmann, F. (2014). Building Information Modeling (BIM) for existing buildings—Literature review and future needs. Automation in construction, 38, 109-127. DOI: https://doi.org/10.1016/j.autcon.2013.10.023

[2] Azpeitia Santander, A., Azkarate Garai-Olaun, A., & De la Fuente Arana, A. (2018). Historic urban landscapes: A review on trends and methodologies in the urban context of the 21st century. *Sustainability*, *10*(8), 2603. DOI:https://doi.org/10.3390/su10082603

[3] Ukabi, E. B., & Akçay, A. Ö. (2023). Conserving the historical identity of north Nicosia Walled City: Exploring design approaches and implications from 1983 to 2003. *Buildings*, *13*(9), 2199. DOI:https://doi.org/10.3390/buildings13092199

[4] Roger, J. A. (2022). *Navigating the Currents and Countercurrents of Southern New Zealand's Human/River Relationship: An Environmental History of Rivers in Otago and Southland, 1890-1920* (Doctoral dissertation, University of Otago). DOI: https://doi.org/10.9734/ajrcos/2023/v16i4386

[5] Henderson, J. C., Liew, G., Ong, J., & Quek, B. (2013). The use of urban built heritage for boutique hotels: Examples from Malaysia and Singapore. Tourism Review International, 17(2), 63-74. DOI: https://doi.org/10.3727/154427213X13728688260758

[6] Henderson, J. C. (2001). Conserving colonial heritage: Raffles hotel in Singapore. International Journal of Heritage Studies, 7(1), 7-24. DOI: https://doi.org/10.1080/13527250119383

[7] Lu, D. Architecture in the Age of Playfulness: Mapping a Framework for Global Historiography. In *The Routledge Companion to Contemporary Architectural History* (pp. 1-80). Routledge. DOI: https://doi.org/10.4324/9781315674469

[8] Rao, X., Zhou, J., Ding, K., Wang, J., Fu, J., & Zhu, Q. (2022). Research on the cultural tracing of the patriarchal clan system of traditional buildings in the Eastern Zhejiang Province, China, based on space syntax: The case study of Huzhai in Shaoxing. *Sustainability*, *14*(12), 7247. DOI: https://doi.org/10.3390/su14127247

[9] Khan, N. R. (2021). Art and architectural traditions of India and Iran: Commonality and diversity. Routledge India. DOI: https://doi.org/10.4324/9781003229421

[10] Botti, G. (2023). Towards a Quantitative Turn: Frontiers of Architectural Historiography. In Designing Emerging Markets: A Quantitative History of Architectural Globalization (pp. 19-53). Singapore: Springer Nature Singapore. DOI: https://doi.org/10.1007/978-981-99-1552-1_2

[11] Rugkhapan, N. T. (2022). Between Toponymy and Cartography: An Evolving Geography of Heritage in George Town, Malaysia. In New Directions in Linguistic Geography: Exploring Articulations of Space (pp. 327-354). Singapore: Springer Nature Singapore. DOI: https://doi.org/10.1007/978-981-19-3663-0_13

[12] Berkes, F. (2007). Community-based conservation in a globalized world. *Proceedings of the National academy of sciences*, *104*(39), 15188-15193. DOI: https://doi.org/10.1073/pnas.0702098104

[13] El Masri, Y., & Rakha, T. (2020). A scoping review of non-destructive testing (NDT) techniques in building performance diagnostic inspections. *Construction and Building Materials*, *265*, 120542. DOI: https://doi.org/10.1016/j.conbuildmat.2020.120542

[14] Castellazzi, G., D'Altri, A. M., de Miranda, S., & Ubertini, F. (2017). An innovative numerical modeling strategy for the structural analysis of historical monumental buildings. Engineering Structures, 132, 229-248. Clark, S. (2021). FEA Validation in Architectural Restoration. Journal of Architectural Conservation, 27(1), 45-58. DOI: https://doi.org/10.1016/j.engstruct.2016.11.032

[15] Arias, P., Armesto, J., Di-Capua, D., González-Drigo, R., Lorenzo, H., & Perez-Gracia, V. (2007). Digital photogrammetry, GPR and computational analysis of structural damages in a mediaeval bridge. *Engineering Failure Analysis*, *14*(8), 1444-1457. DOI: https://doi.org/10.1016/j.engfailanal.2007.02.001

[16] Sippach, T., Dahy, H., Uhlig, K., Grisin, B., Carosella, S., & Middendorf, P. (2020). Structural optimization through biomimetic-inspired material-specific application of plant-based natural fiber-reinforced polymer composites (NFRP) for future sustainable lightweight architecture. *Polymers*, *12*(12), 3048. DOI: https://doi.org/10.3390/polym12123048

[17] Hollaway, L. C. (2010). A review of the present and future utilisation of FRP composites in the civil infrastructure with reference to their important in-service properties. *Construction and building materials*, *24*(12), 2419-2445. DOI: https://doi.org/10.1016/j.conbuildmat.2010.04.062

[18] Freddi, F., Galasso, C., Cremen, G., Dall'Asta, A., Di Sarno, L., Giaralis, A., ... & Woo, G. (2021). Innovations in earthquake risk reduction for resilience: Recent advances and challenges. *International Journal of Disaster Risk Reduction*, *60*, 102267. DOI: https://doi.org/10.1016/j.ijdrr.2021.102267

[19] Idumah, C. I., Obele, C. M., Emmanuel, E. O., & Hassan, A. (2020). Recently emerging nanotechnological advancements in polymer nanocomposite coatings for anti-corrosion, anti-fouling and self-healing. *Surfaces and Interfaces*, *21*, 100734. DOI:https://doi.org/10.1016/j.surfin.2020.100734

[20] Zhu, D. Y., Rong, M. Z., & Zhang, M. Q. (2015). Self-healing polymeric materials based on microencapsulated healing agents: From design to preparation. *Progress in Polymer Science*, *49*, 175-220. DOI:https://doi.org/10.1016/j.progpolymsci.2015.07.002

[21] Ruffolo, S. A., & La Russa, M. F. (2019). Nanostructured coatings for stone protection: An overview. *Frontiers in Materials*, *6*, 147. DOI: https://doi.org/10.3389/fmats.2019.00147

[22] Xu, Z., Wu, L., Shen, Y., Li, F., Wang, Q., & Wang, R. (2014). Tridimensional reconstruction applied to cultural heritage with the use of camera-equipped UAV and terrestrial laser scanner. *Remote sensing*, *6*(11), 10413-10434. DOI:https://doi.org/10.3390/rs61110413

[23] Vileikis, O., Cesaro, G., Santana Quintero, M., Van Balen, K., Paolini, A., & Vafadari, A. (2012). Documentation in world heritage conservation: Towards managing and mitigating change–The case studies of Petra and the Silk Roads. *Journal of Cultural Heritage Management and Sustainable Development*, *2*(2), 130-152. DOI:https://doi.org/10.1108/20441261211273635

[24] Muchiri, P., Pintelon, L., Gelders, L., & Martin, H. (2011). Development of maintenance function performance measurement framework and indicators. *International Journal of Production Economics*, *131*(1), 295-302. DOI:https://doi.org/10.1016/j.ijpe.2010.04.039

[25] Trillo, C., Aburamadan, R., Mubaideen, S., Salameen, D., & Makore, B. C. N. (2020). Towards a systematic approach to digital technologies for heritage conservation. Insights from Jordan. *Preservation, Digital Technology & Culture*, *49*(4), 121-138. DOI: https://doi.org/10.1016/j.ijpe.2010.04.039

[26] Chandler, T., Richards, A. E., Jenny, B., Dickson, F., Huang, J., Klippel, A., ... & Prober, S. M. (2022). Immersive landscapes: modelling ecosystem reference conditions in virtual reality. *Landscape Ecology*, 1-17. DOI: https://doi.org/10.1007/s10980-021-01313-8

[27] Lefebvre, M., Espinosa, M., Gomez y Paloma, S., Paracchini, M. L., Piorr, A., & Zasada, I. (2015). Agricultural landscapes as multi-scale public good and the role of the Common Agricultural Policy. Journal of Environmental Planning and Management, 58(12), 2088-2112. DOI: https://doi.org/10.1111/conl.12398

[28] Najafi, P., Mohammadi, M., van Wesemael, P., & Le Blanc, P. M. (2023). A user-centred virtual city information model for inclusive community design: State-of-art. *Cities*, *134*, 104203. DOI: https://doi.org/10.1016/j.cities.2023.104203

Authors Introduction

Dr. Muhammad Azizi Bin Azizan

He received his PhD in Civil Engineering from the Universiti Malaysia Perlis. He is currently a Senior Lecturer in the same institution. He is Head of Project Integration & Management (PIM) at Centre of Excellence for Unmanned Aerial System (COEUAS), Universiti Malaysia Perlis.

Dr. Nurfadzillah Binti Ishak

She received her PhD in Building Engineering from the Universiti Malaysia Perlis. She is currently a Senior Lecturer in the same institution. She is Head of Built Environment Intelligent (BELL) at Centre of Excellence for Unmanned Aerial System (COEUAS), Universiti Malaysia Perlis.

Prof. Dr. Hazry Bin Desa

He obtained his PhD in Materials Science and Production Engineering (Robotics) from Oita University and currently holds the position as a Head at the Centre of Excellence for Unmanned Aerial Systems (COE-UAS) at Universiti Malaysia Perlis (UniMAP).

Drones and Data: A Comprehensive Exploration of UAVs in Data Mining

Muhammad Azizi Azizan
Faculty of Civil Engineering & Technology, Universiti Malaysia Perlis, Perlis, Malaysia
Centre of Excellence for Unmanned Aerial System, Universiti Malaysia Perlis, Perlis, Malaysia

Nurfadzillah Ishak
Faculty of Civil Engineering & Technology, Universiti Malaysia Perlis, Perlis, Malaysia
Centre of Excellence for Unmanned Aerial System, Universiti Malaysia Perlis, Perlis, Malaysia

Hazry Desa
Faculty of Electrical Engineering & Technology, Universiti Malaysia Perlis, Perlis, Malaysia
Centre of Excellence for Unmanned Aerial System, Universiti Malaysia Perlis, Perlis, Malaysia
E-mail: aziziazizan@unimap.edu.my, nurfadzillah@unimap.edu.my, hazry@unimap.edu.my

Abstract

Encapsulates a comprehensive investigation into the symbiotic relationship between Unmanned Aerial Vehicles (UAVs) and data mining, as encapsulated in the discourse titled "Drones and Data." This exploration delves into the multifaceted applications and transformative impact of UAV technology within the data mining landscape. The examination begins by elucidating the pivotal role of UAVs, highlighting their mobility, accessibility, and capability to collect data from challenging or remote environments. As technology evolves, UAVs have emerged as versatile platforms that redefine the data is collected and analysed across diverse sectors. The narrative unfolds through distinct dimensions, encompassing precision agriculture, environmental monitoring, infrastructure inspection, mining and exploration, disaster response, and urban planning. The technological transitions facilitated by UAVs, emphasizing the integration of machine learning algorithms, cloud-based data processing, and the collaborative synergy between stakeholders. These advancements position UAVs as transformative tools that not only enhance the efficiency of information acquisition but also open avenues for innovative solutions and insights. Therefore, this study a glimpse into the intricate web of applications and technological advancements at the intersection of UAVs and data mining. It serves as a scholarly guide, navigating the reader through the evolving landscape of "Drones and Data," UAVs play a central role in unlocking unprecedented insights and efficiencies, reshaping the future of data mining.

Keywords: Acceptance, Readiness; Unmanned Aerial Vehicle; Data Mining

1. Introduction

In recent years, the integration of Unmanned Aerial Vehicles (UAVs) in data mining applications has witnessed a remarkable surge, reshaping the landscape of information acquisition and analysis. UAVs, commonly known as drones, have become pivotal tools in various industries, offering unique advantages that propel the field of data mining into new frontiers [1]. As technology continues to advance, UAVs have emerged as versatile platforms capable of transforming how we collect and analyze data. Their applications span a wide spectrum, from environmental monitoring and agriculture to infrastructure inspection and disaster management. In the realm of data mining, UAVs play a crucial role in augmenting our ability to gather information with unprecedented efficiency and accuracy [2].

2. Unmanned Advantages of UAVs in Data Mining
2.1. *Mobility*

UAVs provide unparalleled mobility, enabling access to areas that are challenging or impossible for traditional data collection methods. Whether navigating dense forests, inspecting infrastructure, or monitoring vast agricultural landscapes, UAVs offer a dynamic and flexible approach to data acquisition [3].

2.2. *Accessibility*

The ability of UAVs to access hard-to-reach or remote locations is a game-changer in data mining. They overcome geographical constraints and offer a cost-effective alternative to manned missions, allowing for more frequent and widespread data collection [4].

2.3. *Data Collection in Challenging Environments*
UAVs excel in collecting data from environments that pose logistical or safety challenges. From rugged

terrains to disaster-stricken areas, UAVs can swiftly and efficiently capture valuable information, contributing to a more comprehensive understanding of diverse landscapes [5].

2.4. *High-Resolution Imaging*

Equipped with advanced imaging technologies, UAVs capture high-resolution aerial imagery, facilitating detailed spatial analysis. This capability is particularly beneficial in applications such as land-use mapping, environmental modelling, and infrastructure planning [6].

2.5. *Real-time Monitoring*

UAVs enable real-time data acquisition and monitoring, allowing for quick decision-making in dynamic situations. This responsiveness is crucial in applications like disaster response, where timely information can save lives and mitigate the impact of emergencies [7].

The synergy between UAVs and data mining opens up new avenues for exploration and discovery. This introduction sets the stage for a comprehensive examination of how UAVs are reshaping the field, emphasizing their mobility, accessibility, and prowess in collecting data from challenging or remote environments. As a delve deeper into the subsequent sections, study will unravel the multifaceted applications and the transformative impact of UAVs in the realm of data mining [8].

3. UAV Readiness Index (URI) for Data Mining: Navigating the Technological Landscape

The integration of Unmanned Aerial Vehicles (UAVs) into data mining practices is contingent on the acceptance and readiness of stakeholders within diverse industries. This section introduces a UAV Acceptance and Readiness Index, exploring the factors influencing the adoption of UAV technology for data mining and the preparedness of various sectors to embrace these advancements.

Table 1. Regulatory Environment

Regulatory Environment	
Acceptance	Readiness
The willingness of regulatory bodies to adapt and establish clear guidelines for UAV usage in data mining applications [9].	The development of regulatory frameworks that balance innovation with safety, fostering a conducive environment for UAV adoption [9].

Table 2. Technological Infrastructure

Technological Infrastructure	
Acceptance	Readiness
Embracing UAV technology requires a robust technological infrastructure capable of supporting data-intensive processes [10].	Investment in the necessary hardware, software, and communication systems to accommodate the seamless integration of UAVs into data mining workflows [11].

Table 3. Stakeholder Awareness and Education

Stakeholder Awareness and Education	
Acceptance	Readiness
Recognition and understanding of the benefits of UAVs in data mining among stakeholders, including industry professionals, policymakers, and the public [12].	Educational initiatives aimed at raising awareness about UAV capabilities, data security, and the potential impact on various sectors [13].

Table 4. Cost Considerations

Cost Considerations	
Acceptance	Readiness
Recognition of the cost-effectiveness of UAV technology compared to traditional data collection methods [14].	Financial preparedness to invest in UAV infrastructure, training, and maintenance, considering the long-term benefits and efficiency gains [15].

Table 5. Data Security and Privacy

Data Security and Privacy	
Acceptance	Readiness
Acknowledgment of the importance of data security and privacy in UAV-enabled data mining applications [16].	Implementation of robust security measures, encryption protocols, and adherence to privacy regulations to address concerns associated with data collection from aerial platforms [17].

Table 6. Industry Collaboration and Standards

Industry Collaboration and Standards	
Acceptance	Readiness
Willingness of industries to collaborate and share best practices for UAV integration into data mining workflows [18].	Development and adherence to industry standards that streamline UAV adoption, ensuring interoperability and consistency across applications [19].

Table 7. Training and Skill Development

Training and Skill Development	
Acceptance	Readiness
Recognition of the need for specialized training programs to equip professionals with the skills required for UAV-assisted data mining [20].	Implementation of training initiatives to build a workforce capable of operating, maintaining, and extracting valuable insights from UAV-collected data [21].

Table 8. Public Perception and Acceptance

Public Perception and Acceptance	
Acceptance	Readiness
Public perception and acceptance of UAV technology, considering factors such as noise, visual impact, and potential disruptions [22].	Public awareness campaigns to foster understanding and acceptance of UAVs, addressing concerns and showcasing the societal benefits of UAV-enabled data mining [23].

The UAV Acceptance and Readiness Index provides a comprehensive framework for assessing the feasibility of integrating UAV technology into data mining practices. As navigate through the subsequent sections, each factor in this index will be explored in detail, shedding light on the evolving landscape of acceptance and readiness for UAV-enabled data mining across various sectors.

4. Methods

A quantitative approach was employed in this research. A self-administered questionnaire with close-ended questions using a five-point Likert scale ranging. The final questionnaire has been through the preliminary tests which consist of validity and reliability analysis. Content validity is applied by inviting a group of five experts to be selected to pre-test and to review the set of questionnaires. They consist of two academicians working in universities as lecturers and researchers in the area of construction and town planning and three practitioners or project managers who have experience in managing development projects.

Meanwhile, the internal consistency of data reliability in this research was verified by using the Cronbach's Alpha coefficients through the pilot survey, where 30 usable responses from the pilot survey. It was found that the questionnaire coefficient of Cronbach's alpha was more than 0.7 which was deemed as reliable for quantitative data collection. In order to distribute the final survey, a simple random sampling was applied in this research as the population sample is known. In this research, the targeted population is 70 organizations from government agencies and development firms in Kuala Lumpur. Kuala Lumpur has a high number of developments of township and is considered as a positively data mining active location in Malaysia. A total of 56 required samples were successfully obtained and usable for analysis. Data collected were analysed using IBM SPSS Statistical Software which descriptive analysis has been employed in order to achieve the research objective.

5. Results and Discussion
5.1. *The level of the acceptance and readiness of Unmanned Aerial Vehicle implementation in data mining procedures*

Table 9 shows the means and standard deviation based on each variable which is for capabilities factor of Unmanned Aerial Vehicle technology. The variables consist of Regulatory Environment (mean=4.60, sd.=0.34), Technological Infrastructure (mean=4.47, sd.=0.44), Stakeholder Awareness and Education (mean=4.43, sd.=0.44), Cost Considerations (mean=4.41, sd.=0.47) and Data Security and Privacy (mean=4.38, sd.=0.52). Furthermore, for the readiness elements of implementation Unmanned Aerial Vehicle technology are Industry Collaboration and Standards (mean=3.56, sd.=0.44), Training and Skill Development (mean=3.24, sd.=1.38), Public Perception and Acceptance (mean=3.06, sd.=0.95). The mean value score interpretation of the variables used in this study were be interpreted.

Table 9. Descriptive Analysis

Variables	Mean	Std. Deviation	Mean Value Score Interpretation
Acceptance Factors			
Regulatory Environment	4.60	0.34	High
Technological Infrastructure	4.47	0.44	High

Stakeholder Awareness and Education	4.43	0.44	High
Cost Considerations	4.41	0.47	High
Data Security and Privacy	4.38	0.52	High
Readiness Elements			
Industry Collaboration and Standards	3.24	1.38	Moderate
Training and Skill Development	3.06	0.95	Moderate
Public Perception and Acceptance	3.12	1.37	Moderate

Note: The mean value categorized into three levels: low = 1.00 to 2.66; moderate = 2.67 to 3.33; and high = 3.34 to 5.00.

Based on the result illustrated in the Table 9, it can be concluded that in navigating the UAV Acceptance and Readiness Index for Data Mining, our exploration has unveiled critical dimensions shaping the trajectory of UAV technology adoption within the data mining landscape. The confluence of regulatory adaptability, technological infrastructure, stakeholder education, cost considerations, industry collaboration, and public perception forms a complex web influencing the acceptance and readiness for UAVs. The journey through this index demonstrates a forward momentum, wherein regulatory bodies are increasingly adapting to UAV applications, fostering an environment conducive to innovation while ensuring safety and compliance. Technological infrastructure investments have paved the way for seamless UAV integration, with robust hardware, software, and communication systems supporting data-intensive processes. Stakeholder awareness and education emerge as pillars of knowledge empowerment, fostering a deeper understanding of UAV benefits among professionals, policymakers, and the public. Educational initiatives play a crucial role in dispelling concerns, promoting transparency, and positioning UAVs as transformative tools in data mining.

6. Conclusions

As summarised, the UAV Acceptance and Readiness Index for Data Mining represents a holistic transition toward embracing UAVs as integral components of data mining workflows. As stakeholders align their strategies with the dimensions outlined in this index, the landscape of data mining is poised for transformative change [24].

The journey doesn't end here but propels us forward, navigating the future where UAVs play a central role in unlocking unprecedented insights and efficiencies in the world of data mining.

7. Acknowledgement

References

1. h. Yao, R. Qin, And X. Chen, Unmanned Aerial Vehicle for Remote Sensing Applications—A Review, Remote Sens. (Basel), Vol. 11(12), 2019, pp. 1443.
2. n. Bayomi and J. E. Fernandez, Eyes in The Sky: Drones Applications in The Built Environment Under Climate Change Challenges, Drones, Vol. 7(10), 2023, pp. 637.
3. d. Giordan Et Al., The Use of Unmanned Aerial Vehicles (UAVs) For Engineering Geology Applications, Bull. Eng. Geol. Environ., Vol. 79(7), 2020, pp. 3437–3481.
4. z. Zhang and L. Zhu, A Review on Unmanned Aerial Vehicle Remote Sensing: Platforms, Sensors, Data Processing Methods, And Applications, Drones, Vol. 7(6), 2023, pp. 398.
5. w. Alawad, N. B. Halima, And L. Aziz, An Unmanned Aerial Vehicle (UAV) System for Disaster and Crisis Management in Smart Cities, Electronics (Basel), Vol. 12(4), 2023, pp. 1051.
6. h.-W. Choi, H.-J. Kim, S.-K. Kim, And W. S. Na, An Overview of Drone Applications in The Construction Industry, Drones, Vol. 7(8), 2023, pp. 515.
7. l. Gao and Z. Liu, Unraveling the Multifaceted Nexus of Artificial Intelligence Sports and User Willingness: A Focus on Technology Readiness, Perceived Usefulness, And Green Consciousness, Sustainability, Vol. 15(18), 2023, pp. 13961.
8. S. A. H. Mohsan, N. Q. H. Othman, Y. Li, M. H. Alsharif, And M. A. Khan, 2023 Unmanned Aerial Vehicles (Uavs): Practical Aspects, Applications, Open Challenges, Security Issues, And Future Trends, Intell. Serv. Robot.
9. a. Raghunatha, P. Thollander, And S. Barthel, Addressing the Emergence of Drones – A Policy Development Framework for Regional Drone Transportation Systems, Transp. Res. Interdiscip. Perspect., Vol. 18(100795), 2023, pp. 100795.
10. m. A. Azizan Et Al., The Effectiveness of Highway Information Modeling in Kuala Perlis - Changlun Roadway, 2020 In 2nd International Conference on Materials Engineering & Science (IConMEAS 2019).
11. e. Fakhraian, I. Semanjski, S. Semanjski, And E.-H. Aghezzaf, Towards Safe And Efficient Unmanned Aircraft System Operations: Literature Review Of Digital Twins' Applications And European Union

Regulatory Compliance, Drones, Vol. 7(7), 2023, pp. 478.

12. k. Al-Dosari And N. Fetais, A New Shift In Implementing Unmanned Aerial Vehicles (Uavs) In The Safety And Security Of Smart Cities: A Systematic Literature Review, Safety (Basel), Vol. 9(3), 2023, pp. 64.

13. k. Al-Dosari, A. M. Deif, M. Kucukvar, N. Onat, And N. Fetais, Security Supply Chain Using Uavs: Validation and Development of a Uav-Based Model for Qatar's Mega Sporting Events, Drones, Vol. 7(9), 2023, pp. 555.

14. j. Matsimbe, W. Mdolo, C. Kapachika, I. Musonda, And M. Dinka, Comparative Utilization of Drone Technology Vs. Traditional Methods in Open Pit Stockpile Volumetric Computation: A Case of Njuli Quarry, Malawi, Front. Built Environ., Vol. 8, 2022.

15. k. Gunaratne, A. Thibbotuwawa, A. E. Vasegaard, P. Nielsen, And H. N. Perera, Unmanned Aerial Vehicle Adaptation to Facilitate Healthcare Supply Chains in Low-Income Countries, Drones, Vol. 6(11), 2022, pp. 321.

16. h. J. Hadi, Y. Cao, K. U. Nisa, A. M. Jamil, And Q. Ni, A Comprehensive Survey on Security, Privacy Issues and Emerging Defence Technologies for UAVs, J. Netw. Comput. Appl., Vol. 213(103607), 2023, pp. 103607.

17. u. Tariq, I. Ahmed, A. K. Bashir, And K. Shaukat, A Critical Cybersecurity Analysis and Future Research Directions for The Internet of Things: A Comprehensive Review, Sensors (Basel), Vol. 23(8), 2023, pp. 4117.

18. n.-A. Perifanis And F. Kitsios, Investigating the Influence of Artificial Intelligence on Business Value in The Digital Era of Strategy: A Literature Review, Information (Basel), Vol. 14(2), 2023, pp. 85.

19. S. A. H. Mohsan, N. Q. H. Othman, Y. Li, M. H. Alsharif, and M. A. Khan, Unmanned Aerial Vehicles (UAVs): Practical Aspects, Applications, Open Challenges, Security Issues, And Future Trends, Intell. Serv. Robot, 2023.

20. L. Li, Reskilling and Upskilling the Future-Ready Workforce for Industry 4.0 And Beyond, Inf. Syst. Front., 2022.

21. a. M. Muhmad Kamarulzaman, W. S. Wan Mohd Jaafar, M. N. Mohd Said, S. N. M. Saad, And M. Mohan, Uav Implementations in Urban Planning and Related Sectors of Rapidly Developing Nations: A Review and Future Perspectives for Malaysia, Remote Sens. (Basel), Vol. 15(11), 2023, pp. 2845.

22. b. Aydin, Public Acceptance of Drones: Knowledge, Attitudes, And Practice, Technol. Soc., Vol. 59(101180), 2019, pp. 101180.

23. f. Nex Et Al., Uav in The Advent of The Twenties: Where We Stand and What Is Next, Isprs J. Photogramm. Remote Sens., Vol. 184, 2022, pp. 215–242.

24. h. Desa, M. Azizi Bin Azizan, M. S. A. Khadir, M. S. Suhaimi, N. Z. Ramli, And Z. Hat, Feasibility Study of UAV Implementation in Route Surveying, J. Robot. Netw. Artif. Life, 2019.

Authors Introduction

Dr. Muhammad Azizi Bin Azizan

He received his PhD in Civil Engineering from the Universiti Malaysia Perlis. He is currently a Senior Lecturer in the same institution. He is Head of Project Integration & Management (PIM) at Centre of Excellence for Unmanned Aerial System (COE-UAS), Universiti Malaysia Perlis.

Dr. Nurfadzillah Binti Ishak

She received her PhD in Building Engineering from the Universiti Malaysia Perlis. She is currently a Senior Lecturer in the same institution. She is Head of Built Environment Intelligent (BELL) at Centre of Excellence for Unmanned Aerial System (COE-UAS), Universiti Malaysia Perlis.

Dr. Hazry Bin Desa

He obtained his PhD in Materials Science and Production Engineering (Robotics) from Oita University and currently holds the position as a Head at the Centre of Excellence for Unmanned Aerial Systems (COE-UAS) at Universiti Malaysia Perlis (UniMAP).

Development of Variable Arm to Control the Manoeuvrability of Quadrotor

Lim Yi Hong, Hazry Desa, Muhammad Azizi Azizan

Centre of Excellence for Unmanned Aerial Systems (COE-UAS), Universiti Malaysia Perlis, Block E, Pusat Perniagaan Pengkalan Jaya, Jalan Kangar – Alor Setar, 01000 Kangar, Perlis, Malaysia.

Muhammad Hassan Tanveer

Department of Robotics and Mechatronics Engineering, Kennesaw State University, Marietta, GA, 30067, USA.

Email: hazry@unimap.edu.my
www.unimap.edu.my

Abstract

This paper introduces the concept of a variable arm for a quadrotor which is able to perform the manoeuvrability of the quadrotor by changing the arm length's. The variation in arm's length affects the bending moment generated by the thrust force, resulting in the tilting and movement of the quadrotor. The primary goal of this project is to develop a quadrotor with an adjustable arm length to control its manoeuvrability effectively, thus minimizing the need for additional thrust force during flight control. The study focuses on designing a quadrotor with the capability to extend or retract its arms. The proposed concept relies on altering the bending moment through the variable arm to control the quadrotor's manoeuvrability. A quadrotor equipped with a variable arm was successfully designed and tested, with its performance evaluated in executing agile maneuvers. The experiment demonstrated that the variable arm induced body rotation in the quadrotor, effectively regulating its manoeuvrability and the study validates the potential of the variable arm approach for controlling quadrotor movement.

Keywords: N Retract, Control, Agile maneuvers, Body rotation, Performance evaluation, Validation.

1. Introduction

The quadrotor's working principle can be elucidated through Newton's Third Law of Motion, which states that every action has an equal and opposite reaction [1]. When the quadrotor's motor rotates the propeller, it generates a downward force on the air, a phenomenon explained by Bernoulli's Principle [2]. In a standard quadrotor, different motor speeds in the four motors allow for thrust generation, enabling precise control of the quadrotor's movements [3]. As the size and weight increase, greater thrust must be produced to lift the quadrotor effectively [4]. Typically, achieving this entails designing the quadrotor with higher-speed motors or larger propellers to generate the necessary thrust to counter the quadrotor's weight [5]. Consequently, this results in higher current consumption and increased power usage, leading to reduced flight endurance [6]. In this context, designing the quadrotor using conventional methods can be both costly and inefficient [7]. This project proposes a variable arm approach to control the quadrotor, utilizing different arm lengths to regulate its manoeuvrability [8].

The research on variable arm-controlled quadrotors draws on a foundation of related work in the fields of sensor networks, IoT applications, and intelligent systems. Kousik et al. leveraged a hybrid Convolution Recurrent Neural Network for improved salient object detection [9], showcasing advancements in intelligent systems. Singanamalla et al. addressed reliability and energy efficiency in emergency transmission within wireless sensor networks [10], contributing to the broader context of networked systems. Alzubi et al. conducted a survey on specific IoT applications, providing insights into the diverse applications of IoT across various domains [11]. Furthermore, Kallam et al. explored low-energy aware communication processes in IoT through a green computing approach [12], aligning with the focus on energy efficiency in the development of the variable arm-controlled quadrotor. Suresh et al. proposed an energy-efficient mechanism for leveraging IoT [13], setting the stage for considerations related to energy efficiency in unmanned aerial systems. Mekala et al. presented a computational intelligent sensor-rank consolidation approach for the industrial Internet of Things (IIoT) [14], providing valuable insights into intelligent systems applied to industrial contexts. Additionally, Reddy et al. focused on smart assistance for elderly individuals in emergency situations using IoT [15], showcasing the potential applications of intelligent systems in healthcare scenarios. Collectively, these works contribute to the foundation of knowledge necessary for understanding and advancing the development of innovative systems, such as the variable arm-controlled quadrotor presented in the current paper.

2. Materials and Methods

The quadrotor, depicted in Fig. 1, is designed in the form of a variable arm quadrotor, comprising 10 essential components. These components include the Arduino Mega, serving as the microcontroller, an ESC controller responsible for regulating the RPM of the brushless DC motor, the MDD3A driver for controlling the electric actuator's movement, and an MPU 6050 sensor located at the quadrotor's center to detect its acceleration and rotation. Additionally, there are two sets of variable arms and fixed arms, a LiPo Battery for power supply, an A2212 2200KV brushless DC motor, and a base frame. Specifically, the variable arm is engineered as an electric actuator with a linear guide and designed to securely hold the brushless DC motor.
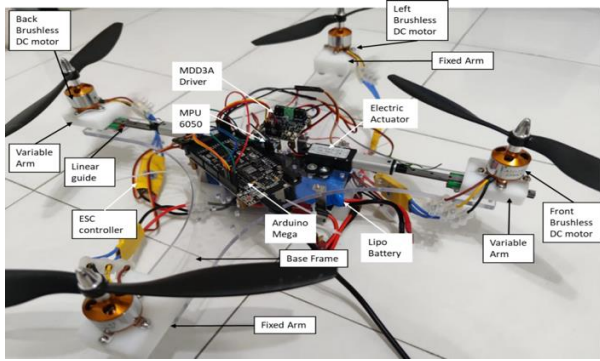
Fig. 1. Variable arm quadrotor.

3. Conceptual Modelling

The standard quadrotor achieves manoeuvrability by adjusting its speed to generate varying thrusts. The quadrotor's design follows a plus configuration, which offers simplicity in both calculation and analysis compared to the cross configuration. Hence, the calculations are conducted with two propellers while assuming the other two propellers remain constant. In Fig. 2, a free-body diagram illustrates both the normal quadrotor and the variable arm quadrotor, showcasing the relationship between angle, acceleration, and the difference in thrust production with various arm lengths. Fig. 2(a) displays the normal quadrotor, controlled by different thrust forces, while Fig. 2(b) demonstrates the variable arm quadrotor, which relies on adjusting the arm length for control.

$$W = L\cos(Angle) + R\cos(Angle) = 3L_y + R_y \quad (1)$$

$$F = ma \quad (2)$$

Based on Fig. 2, the quadrotor weighs 1.94 kg, and each arm requires a thrust force of 4.76 N to maintain equilibrium. By assuming that only one thrust is changed while the other three remain constant, Eq. (1) can be utilized to determine the force R required for the quadrotor to attain a specific angle. The acceleration generated by Eq. (2) is depicted in Fig. 3(a) while Fig.

3(b) illustrates the vertical force produced by both constant thrust (L_y) and the changed thrust (R_y).
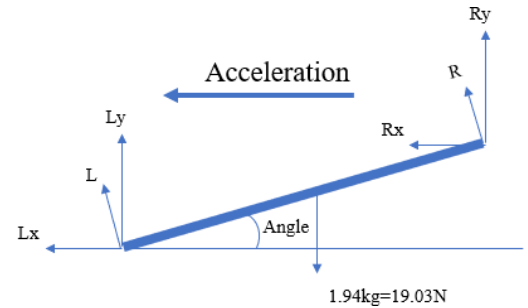
(a)
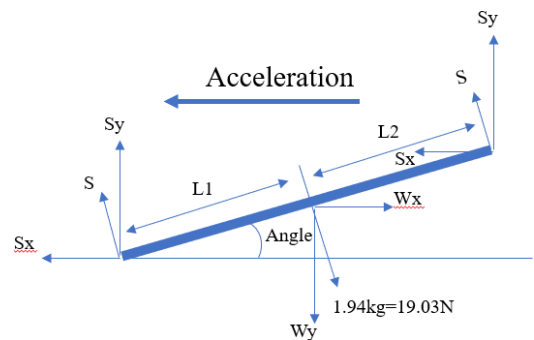
(b)

Fig. 2. Free body diagram for (a) normal quadrotor and (b) variable arm quadrotor.
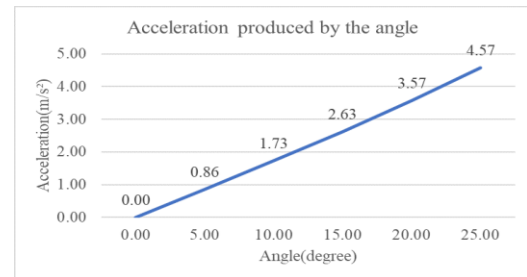
(a)
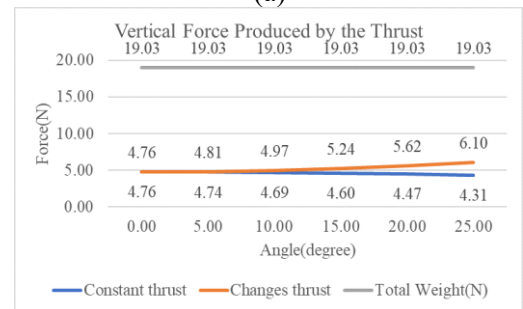
(b)

Fig. 3. (a) The acceleration and (b) changes of vertical force.

When the angle increases, the constant thrust generates less vertical force. Therefore, to achieve equilibrium for the quadrotor, the changed thrust should be increased to enhance the vertical force produced.

Additionally, this will result in a bending moment, as shown in Fig. 4(a), where the constant length of 200 mm is used in Eq. (3).

$$M = FL \qquad (3)$$

Based on the bending moment in Fig. 4(a), the concept of a variable arm is introduced, as demonstrated in Fig. 2(b). In this configuration, the quadrotor can be adjusted in angle by modifying the variable arm. When keeping the thrust constant, altering the length of the thrust from the quadrotor's center affects the bending moment, as indicated by Eq. (3).

By maintaining L1 and constant thrust S, the bending moment on the quadrotor can be made equivalent to that in Fig. 4(a) by adjusting the length of L2, as shown in Fig. 4(b). Consequently, increasing the length will lead to a higher bending moment.

(a)

(b)
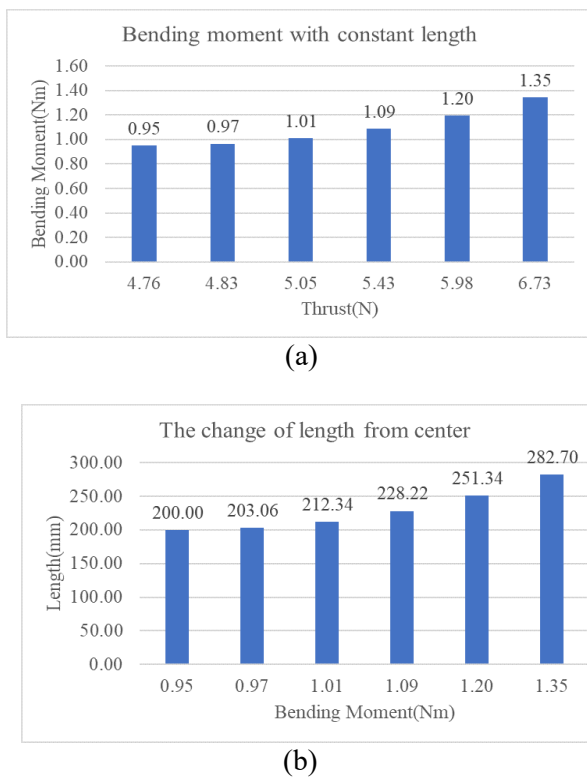
Fig. 4. Bending moment with (a) different thrust and (b) different length.

4. Result and Discussion

Based on the previous conceptual modeling, a prototype was developed to observe and gather data regarding the quadrotor's movement with varying arm lengths. The quadrotor is secured by a gripper and positioned in a way that allows testing for rotation and acceleration. To measure these parameters, an MPU6050 sensor is utilized, providing data along the X, Y, and Z axes.

The quadrotor's design consists of 2 variable arms and 2 fixed arms. Consequently, altering the length of the variable arm causes the quadrotor to rotate along one axis and move along another. Specifically, the quadrotor rotates on the Y axis to move along the X axis.

Additionally, data from other axes are used to assess the stability and functionality of the variable arm.
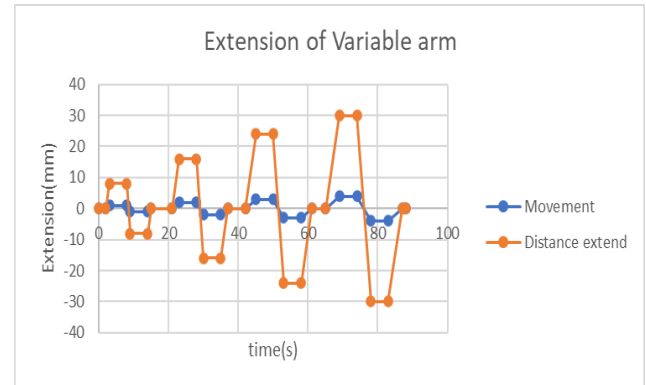
Fig. 5. Input signal and movement of the variable arm.

To control the extension or retraction of the variable arm, a signal is sent from the Arduino Mega microcontroller to the MDD3A driver. Fig. 5 illustrates the input signal for extending the variable arm. The positive values indicate forward movement, while negative values indicate backward movement. The electric actuator's extension speed is set at 8 mm/s with a 30 mm stroke length. By incorporating program delay time, the extension length of the variable arm can be precisely controlled.

For data collection, the variable arm is extended to different lengths (8 mm, 16 mm, 24 mm, and 30 mm) and held for a few seconds at each length to gather relevant information. The entire extension cycle takes approximately 88 seconds.

4.1. With Constant Thrust Force

Fig. 6 and Fig. 7 demonstrate that as the arm length increases, the rotation and acceleration of the quadrotor also increase. However, it is evident that the relationship between extension length and these variables is not linear. To address this, a simplified graph for Fig. 6(b) and Fig. 7(b) was created, selecting the second highest positive value or lowest negative value to minimize the influence of noise from Fig. 6(a) and Fig. 7(a).
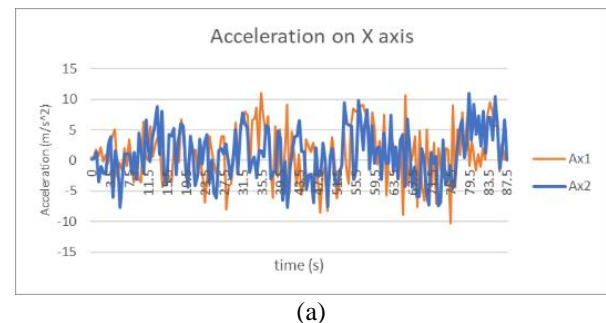
(a)
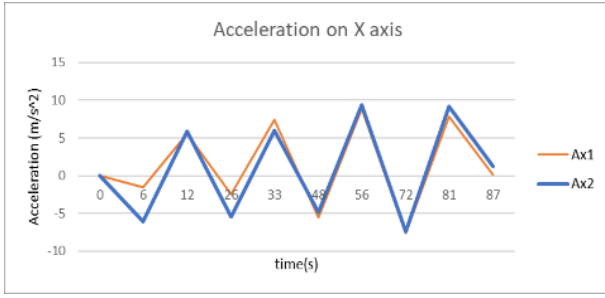
(b)

Fig. 6. (a) Acceleration on X axis for dataset 1 (Ax1) & dataset 2 (Ax2) and (b) simplify data for acceleration on X axis for dataset 1 (Ax1) & dataset 2 (Ax2).
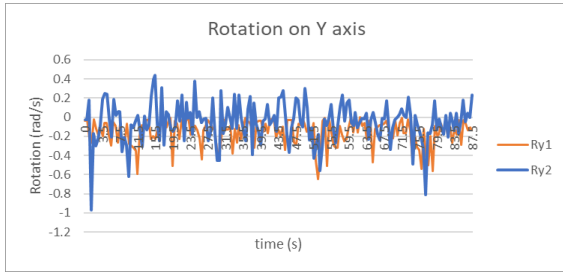
(a)
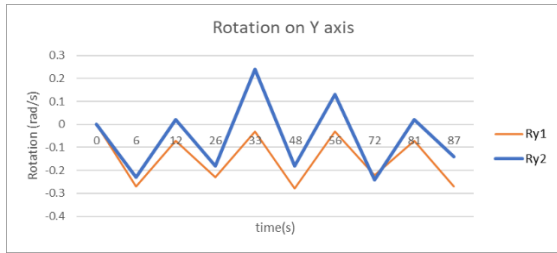
(b)

Fig. 7. (a) Rotation on Y axis for dataset 1 (Ry1) & dataset 2 (Ry2) and (b) simplify data for rotation on Y axis for dataset 1 (Ry1) & dataset 2 (Ry2).

(a)

(b)

Fig. 8. (a) Acceleration on Y axis for dataset 1 (Ay1) & dataset 2 (Ay2) and (b) rotation on X axis for dataset 1 (Rx1) & dataset 2 (Rx2).
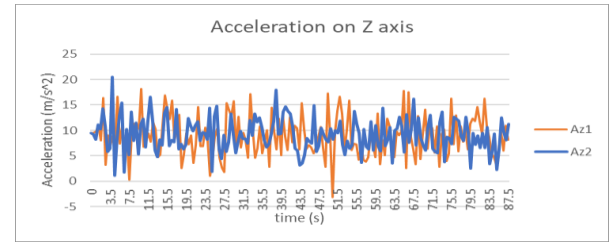
(a)
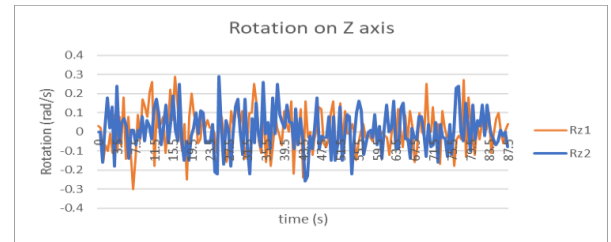
(b)

Fig. 9. (a) Acceleration on Z axis for dataset 1 (Az1) & dataset 2 (Az2) and (b) rotation for Z axis for dataset1 (Rz1) & dataset 2 (Rz2).

On the other hand, Fig. 8 and Fig. 9 display readings for Ax, Az, Rx, and Rz, but with considerable noise. Observing the graphs, most of the data for Ax, Az, Rx, and Rz falls within a range of values from 3 to -3, 5 to 15, 0.2 to -0.2, and 0.15 to -0.15, respectively. However, there is a significant amount of data that exceeds these ranges, indicating an unstable condition for the quadrotor. This instability could be attributed to various factors such as environmental conditions like surrounding air flows, unstable signals from the microcontroller, inconsistent power supply, and the accuracy and sensitivity of the sensor being used.

4.2. With Variable Thrust Force

The variable thrust force incorporates a straightforward program for regulating the RPM of brushless DC motors, utilizing the range established from earlier findings. During quadrotor operation, the craft is tilted at an angle to facilitate movement, and thus, the Rx, Rz, and Az ranges are utilized, with Ax being disregarded. Az comes into play to ensure the quadrotor remains in an equilibrium state and maintains its position.

Consequently, this approach has the potential to reduce noise and enhance the quadrotor's overall stability.
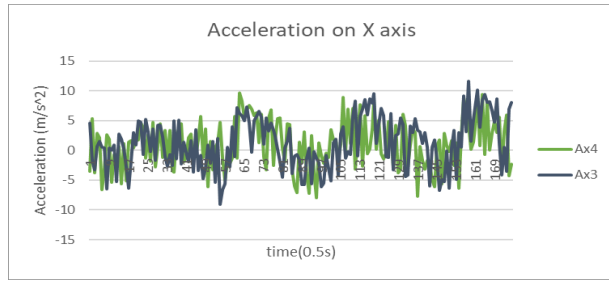
(a)
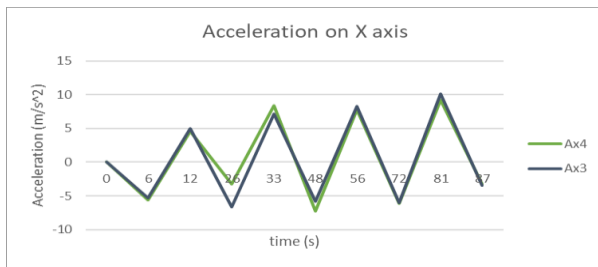
(b)

Fig. 10. (a) Acceleration on X axis for dataset 3 (Ax3) & dataset 4 (Ax4) and (b) simplify data for acceleration on X axis for dataset 3 (Ax3) & dataset 4 (Ax4).
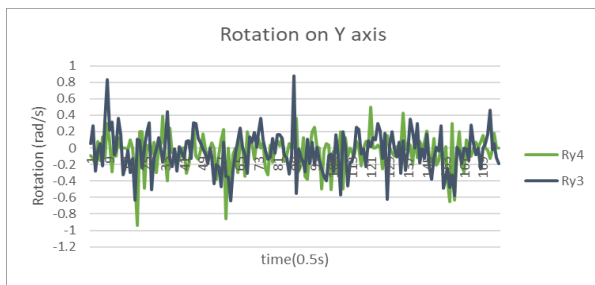
(a)
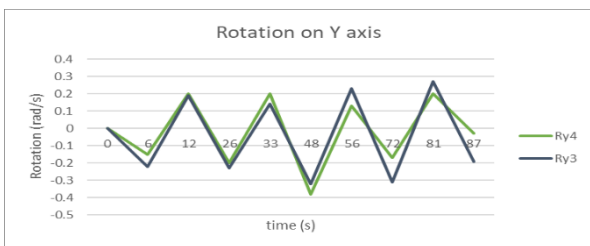
(b)

Fig. 11. (a) Rotation of Y axis for dataset 3 (Ry3) & dataset 4 (Ry4) and (b) simplify data for rotation on Y axis for dataset 3 (Ry3) & dataset 4 (Ry4).

In Fig. 10 and Fig. 11, as the variable arm length increases, both rotation and acceleration also experience a corresponding increase. To create a simplified representation, we referred to the movement depicted in Fig. 5 and extracted the second-highest positive value or lowest negative value for Fig 10(b) and Fig. 11(b). This approach helps mitigate any potential noise influence from Fig. 10(a) and Fig. 11(a) on the graphs.
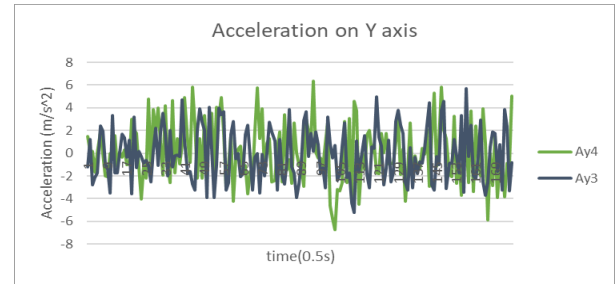
(a)
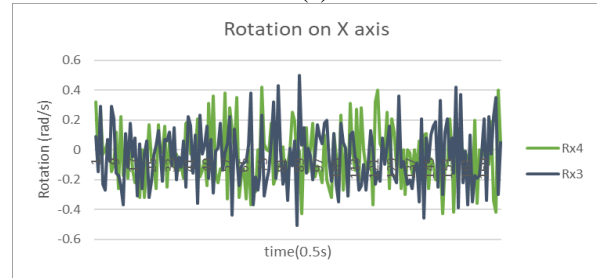
(b)

Fig. 12. (a) Acceleration on Y axis for dataset 3 (Ay3) & dataset 4 (Ay4) and (b) rotation on X axis for dataset 3 (Rx3) & dataset 4 (Rx4).
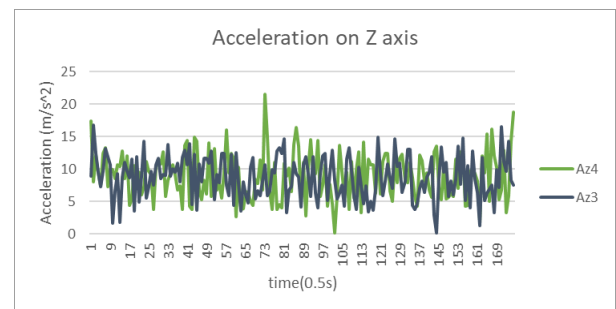
(a)
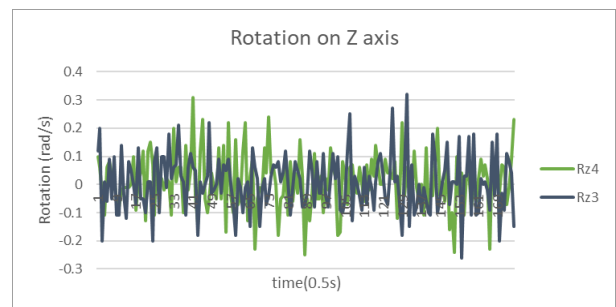
(b)

Fig. 13. Acceleration on Z axis for dataset 3 (Az3) & dataset 4 (Az4) and (b) rotation on Z axis for dataset 3 (Rz3) & dataset 4 (Rz4).

Lim Yi Hong, Hazry Desa, Muhammad Azizi Azizan, Muhammad Hassan Tanveer
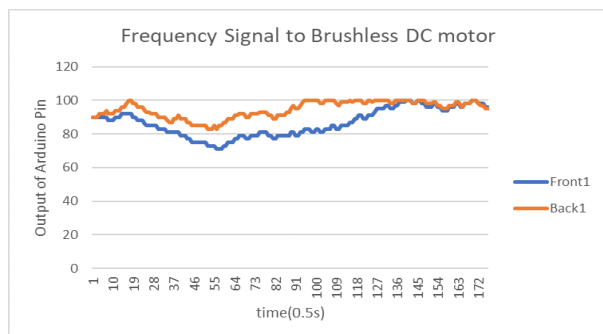
(a)
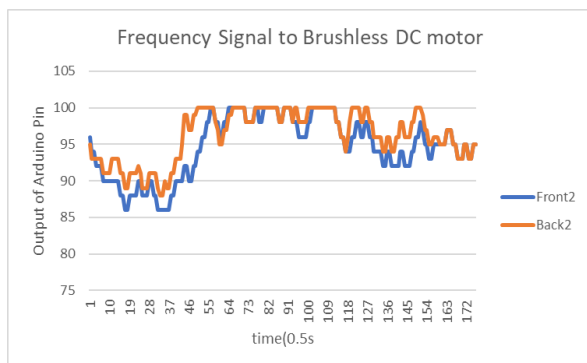
(b)

Fig. 14. Frequency signal to brushless dc motor for (a) dataset 3 and (b) dataset 4.

In Fig. 12 and Fig. 13, the presented results still exhibit some degree of noise. However, upon comparison with the outcomes from the previous section, it becomes evident that the quadrotor demonstrates enhanced stability and balance. Fig. 14 illustrates that the frequency signal sent to the motor undergoes alterations in a random pattern, without showing any discernible relationship to changes in the variable arm.

5. Conclusion

In this project, a quadrotor with a variable arm length was constructed. The variable arm is used to control the manoeuvrability of the quadrotor. Data were collected in four sets, consisting of two with constant thrust force and two with variable thrust force. The analysis of the constant thrust force data indicated that the quadrotor exhibited instability, likely resulting from factors such as surrounding airflow, an unstable signal from the microcontroller, and fluctuations in the power supply. To address these issues, a variable thrust force was implemented to minimize noise and stabilize the quadrotor. The results displayed a clear linear relationship between the variable arm length and the rotation and acceleration of the quadrotor. As the arm length increased, both the rotation and acceleration of the quadrotor also increased. Additionally, the findings revealed that extending the variable arm eliminated the need for extra thrust force, enabling the motor to fully utilize its capacity in carrying the payload. The graphs further indicated some noise variation, potentially caused by vibrations in the brushless DC motor. In conclusion, a

concept utilizing an electric actuator and linear guide for the variable arm was introduced, proving suitable for enhancing the manoeuvrability of the quadrotor by adjusting the arm length.

References

1. M. F. Ahmed and Y. S. Narayan, Fabrication and Testing of Quadcopter Prototype for Surveillance, International Journal of Mechanical and Production Engineering Research and Development (IJMPERD), June, 2018, pp. 99–105.
2. S. A. Khan, Z. Mehmood and Z. Afshan, Design, Analysis, and Topology Optimization of a Landing Gear Strut for a Quadcopter Upon Impact, International Conference on Applied and Engineering Mathematics (ICAEM) (IEEE), September 2021, pp. 37–42.
3. Y. R. Tang and Y. Li, Realization of the Flight Control for an Indoor UAV Quadrotor, IEEE International Conference on Information and Automation (ICIA) (IEEE), August, 2023, pp. 1278–1283.
4. N. Y. Kamil, D. Hazry, K. Wan and Z. M. Razlan, Trajectory Tracking Based on Arm's Length Variation, Journal of Theoretical & Applied Information Technology, Vol. 79(3), October, 2015, pp. 528-536.
5. P. E. I. Pounds, Design, Construction and Control of a Large Quadrotor Micro Air Vehicle, Thesis for the Degree of Doctor of Philosophy of the Australian National University, September, 2007.
6. N. Y. Kamil, D. Hazry, K. Wan and Z. Razlan, A Novel VAL: Quadrotor Control Technique for Trajectory Tracking Based on Varying the Arm's Length, ARPN J. Eng. Appl. Sci, Vol. 11, 2015, pp. 9195–9204.
7. R. Sekaran, R. Patan, A. Raveendran, F. Al-Turjman, M. Ramachandran and L. Mostarda, Survival Study on Blockchain Based 6G-Enabled Mobile Edge Computation for IoT Automation, IEEE access, Vol. 8, August, 2020, pp. 143453-143463.
8. M. Rizon, C. Ang, M. I. Solihin, M. Z. Razlan, D. Hazry, S. A. Bakar, W. Khairunizam and I. Zunaidi, Effects of Variable Arm Length on UAV Control Systems, Journal of Robotics Networking and Artificial Life, Vol. 7(2), September, 2020, pp. 91–97.
9. N. Kousik, Y. Natarajan, R. A. Raja, S. Kallam, R. Patan and A. H. Gandomi, Improved Salient Object Detection Using Hybrid Convolution Recurrent Neural Network, Expert Systems with Applications, Vol. 166, March, 2021, pp. 1-9.
10. V. Singanamalla, R. Patan, M. S. Khan, and S. Kallam, Reliable and Energy‐Efficient Emergency Transmission in Wireless Sensor Networks, Internet Technology Letters, Vol. 2(2), April, 2019.
11. J. A. Alzubi, R. Manikandan, O. A. Alzubi, O, N. Gayathri and R. Patan, A Survey of Specific IoT Applications, International Journal on Emerging Technologies, Vol. 10(1), 2019, pp. 47-53.
12. S. Kallam, R. B. Madda, C. Y. Chen, R. Patan and D. Cheelu, Low Energy Aware Communication Process in

IoT Using the Green Computing Approach, IET Networks, Vol. 7(4), July, 2018, pp. 258-264.

13. K. Suresh, M. RajasekharaBabu and R. Patan, EEIoT: Energy Efficient Mechanism to Leverage the Internet of Things (IoT), International Conference on Emerging Technological Trends (ICETT), October, 2016, pp. 1-4.

14. M. S. Mekala, P. Rizwan and M. S. Khan, Computational Intelligent Sensor-Rank Consolidation Approach for Industrial Internet of Things (IIOT). IEEE Internet of Things Journal, Vol. 10 (3), February, 2023, pp. 2121-2130.

15. A. R. Reddy, G. P. Ghantasala, R. Patan, R. Manikandan and S. Kallam, Smart Assistance of Elderly Individuals in Emergency Situations at Home, Internet of Medical Things: Remote Healthcare Systems and Applications, April, 2021, pp. 95-115.

Authors Introduction

Mr. Lim Yi Hong

He completed his bachelor's degree in Mechatronic Engineering from Universiti Malaysia Perlis and currently works as a Mechanical Design Engineer at Pentamaster Equipment Manufacturing Sdn Bhd in Penang, Malaysia.

Prof. Dr. Hazry Desa

He earned his Bachelor of Mechanical Engineering from Tokushima University, Japan, and proceeded to pursue his Ph.D. at the Artificial Life and Robotics Laboratory, Oita University, Japan, in 2003. Presently, he serves as the director at the Center of Excellence for Unmanned Aerial Systems (COE-UAS), Universiti Malaysia Perlis, Malaysia.

Dr. Muhammad Azizi Azizan

He received his PhD in Civil Engineering from the Universiti Malaysia Perlis. He is currently a senior lecturer in the same institution. He is head of Project Integration & Management (PIM) at Centre of Excellence for Unmanned Aerial System (COE-UAS), Universiti Malaysia Perlis.

Dr. Muhammad Hassan Tanveer

He is an Assistant Professor of Robotics and Mechatronics Engineering at Kennesaw State University. Prior to that, he was a Post Doctoral Research Associate at Virginia Tech, USA where he was working on development of Bio-sonar sensor for UAVs navigations. He received his Ph.D. in Robotics and Autonomous Sytems with specilization in 'Collaboration of Heterogenous Team of Robots' from University of Genova, Italy, in 2019.

Development of IoT-Enabled Smart Water Metering System

Saw Di Wen, Hazry Desa and Muhammad Azizi Azizan

Centre of Excellence for Unmanned Aerial Systems (COE-UAS), Universiti Malaysia Perlis, Block E, Pusat Perniagaan Pengkalan Jaya, Jalan Kangar – Alor Setar, 01000 Kangar, Perlis, Malaysia.

Abadal-Salam T. Hussain

Department of Medical Equipment Technology Engineering, College of Engineering Technology, Al-Kitab University, Kirkuk, Iraq.

Muhammad Hassan Tanveer

Department of Robotics and Mechatronics Engineering, Kennesaw State University, Marietta, GA, 30067, USA.

Rizwan Patan

Department of Software Engineering and Game Development, Kennesaw State University, Marietta, GA, 30067, USA.

Email: hazry@unimap.edu.my
www.unimap.edu.my

Abstract

This paper introduces a smart water meter that utilizes the capabilities of the Internet of Things (IoT) to automate the collection of meter readings. The primary goal of this project is to create an IoT-based device for reading water meters, while simultaneously developing a compatible mobile application. Instead of relying on manual meter reading, which requires human effort, this project proposes the use of an IoT-enabled water meter to collect the data automatically. The device employs a camera and Convolutional Neural Network (CNN) for image processing, making it easy to detect the meter reading accurately. The IoT system architecture involves the use of an ESP32 CAM for data collection, a laptop as a gateway, and the Message Queuing Telemetry Transport (MQTT) protocol for data transfer. The collected data is stored in Firebase's real-time database, and the mobile application is designed to monitor and analyze the data. A functional prototype of the device is constructed and tested in a housing area. The collected data is then monitored through the developed mobile application. Lastly, the data is analyzed to assess the suitability of the proposed method, and recommendations for future improvements are provided.

Keywords: IoT, Neural Network, Smart Water Management, Message Queuing system.

1. Introduction

Water consumption is currently in high demand and has been increasing due to population growth. Insufficient water management capacity, unstructured management, and the adverse effects of urbanization contribute to water shortages [1], [2], [3]. Particularly in residential areas, households consume a significant amount of water daily [4], [5]. This excessive consumption, exceeding the available water supply, raises concerns about future water shortages. Lack of public awareness, a significant percentage of non-revenue water, inefficient water demand management, and low water tariffs are generally recognized as the main causes of excessive water usage [6]. Excessive water consumption is also attributed to water losses. To address this issue, the water balance needs to be calculated, distinguishing between revenue water and non-revenue water. Unbilled permitted use and water losses are further categorized into apparent losses, which include customer meter errors and unauthorized use, and true losses, such as leaks in mains, service connections, and storage tanks [7].

Traditionally, water meters have been used for invoicing water consumption, and manual meter reading is the prevailing method. However, this approach requires a significant workforce and is prone to errors that can harm water companies. Moreover, manual meter reading often necessitates entering indoor premises, adding to the challenges faced by users and hindering daily operations [8]. Mechanical water meters, while reliable and affordable, do not allow for automatic or real-time monitoring of water usage. Consequently, water management companies have relied solely on them, despite the labor-intensive, time-consuming, and error-

prone nature of the process [8]. The market is gradually shifting towards smart meter reading methods, which eliminate the need for manual reading. This transition began in industrialized Western countries, with the United States leading the adoption of smart meter reading technology in the 1980s [9].

The automated collection and analysis of meter data are facilitated by smart water metering (SWM) systems, which incorporate various technologies. Customers can access their water usage data through online platforms such as websites or mobile applications, allowing them to manage their daily water usage more efficiently. Smart metering improves operations by reducing costs associated with operation and maintenance, per capita usage, waste, and leaks. Key benefits include monitoring water flow, distribution, and consumption, improving access to clean and safe water, enabling frequent or real-time access to water consumption information and billing, reducing the need for manual meter reading, enhancing leak and fraud detection, and improving data collection accuracy [10].

2. Methodology

2.1. Hardware Requirement

The primary microcontroller employed in this project is the ESP32 Cam, which runs the main program. The program code is uploaded to the microcontroller using a programming software platform. Additionally, an SD card is utilized to expand the memory capacity, enabling the installation of the program, and recording of collected data. The camera module, a crucial component, is connected to the microcontroller and serves the purpose of monitoring the water meter by capturing images of the meter reading as shown in Fig. 1.
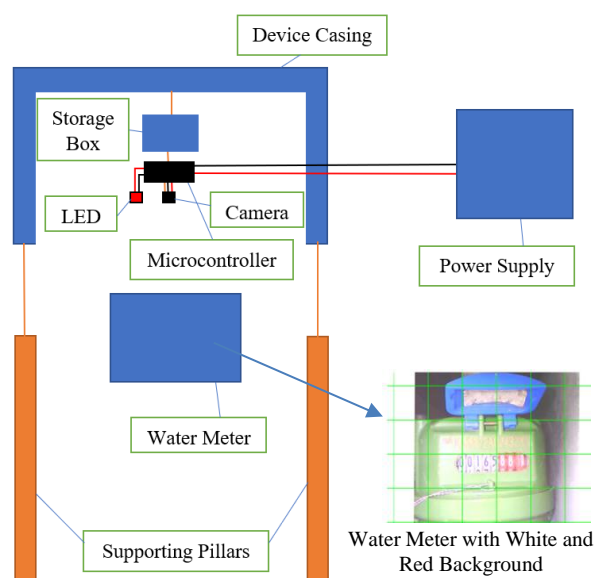
Fig. 1. The construction of a smart water meter featuring a display with white and red backgrounds.

2.2. Image Processing

The image processing algorithm used in this project has been modified from an open-source codebase originally created by Jomjol [11]. When an image is captured, it is sent to a web server for processing. On the server, the image alignment is adjusted to a specific orientation to facilitate smooth text recognition. Image segmentation can also be fine-tuned within a certain range to accurately detect numbers within the image. Additionally, artificial intelligence has been integrated into the image processing algorithm. By employing a Convolutional Neural Network (CNN) and training it with various types of water meter images, the accuracy of number recognition has significantly improved. The introduction of AI has enabled the microcontroller to perform data computation at the edge, rather than relying on cloud computing. This approach minimizes delays in obtaining data. Instead of waiting for data to be processed in the cloud, the microcontroller can analyze the information locally, resulting in real-time or near-real-time data processing.

2.3. Data Transfer

The image processing software developed by Jomjol utilizes MQTT to transfer data. The configuration for MQTT can be pre-set within the web server [11]. In this project, the microcontroller remains powered on for 24 hours, and data is collected at 30-minute intervals. To facilitate data transfer using the MQTT protocol, an MQTT broker must be installed on a device to host the server. In this case, the Aedes broker was chosen due to its easy installation process. Additionally, the Aedes broker is compatible with Node-Red, a programming tool used for sending data to Firebase. With some coding, the microcontroller can function as an MQTT publisher, responsible for publishing data on a specific topic. The Aedes broker then manages the published message and alerts subscribers who have subscribed to the same topic. In this scenario, the published message refers to the processed data, while the laptop acts as a subscriber, having Node-Red installed to receive the data. As a result, the microcontroller can establish communication with the laptop for data transfer. Leveraging the capabilities of Node-Red, the data can be further transmitted and stored in the real-time database of Firebase.

2.4. Cloud

For data transfer via MQTT, a connectivity platform is required, and in this case, a laptop is utilized as the gateway. The laptop acts as an intermediary, receiving data from the microcontroller and sending it to the cloud. Facilitating this communication is Node-Red, which has been installed on the laptop. All the data transmitted by the microcontroller is written into the Firebase real-time database. This includes the initial water meter reading as well as the most recent reading, along with timestamps indicating when the data was captured. The presence of data stored in the Firebase real-time database is considered a successful transfer of data from the microcontroller to the cloud at this stage. This mechanism

aligns with the cloud and network aspects in the architecture of IoT.

2.5. Application

The final component of the IoT architecture involves application and data analysis. In terms of application, users can access the water meter data through a dedicated app from anywhere, as long as they are connected to a Wi-Fi network. Achieving this functionality involves synchronizing the database with the developed mobile application. Additionally, the app has the capability to generate a bill based on the collected data.

2.6. Flowchart
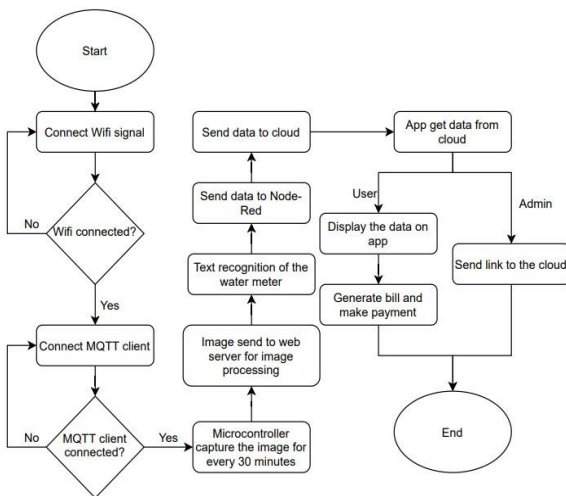
Fig. 2. Smart water meter system flowchart.

This paper proposes a method for obtaining water meter readings through a smart water meter system. To initiate the system, the microcontroller is powered on using a 5V DC power supply, enabling it to connect to a Wi-Fi network. Once a stable Wi-Fi connection is established, the program is designed to connect to the MQTT broker as a client. Once all connections are established, the microcontroller is ready to commence its operations.

The camera captures the first photo of the water meter immediately upon activation, with the triggering interval set to every 30 minutes. The captured image is then uploaded to a web server for further processing, aimed at extracting the data from the image. The text recognition process concludes by sending the detected value to the database via a tool called Node-Red.

The mobile application serves as a means for consumers to retrieve and monitor data from the cloud. It allows them to access the water meter readings and view them on their mobile devices. Additionally, administrators can utilize the app to monitor the readings and even send payment links to the cloud. Fig. 2 shows the flowcharts of the developed smart water meter system.

2.7. Circuit Connection
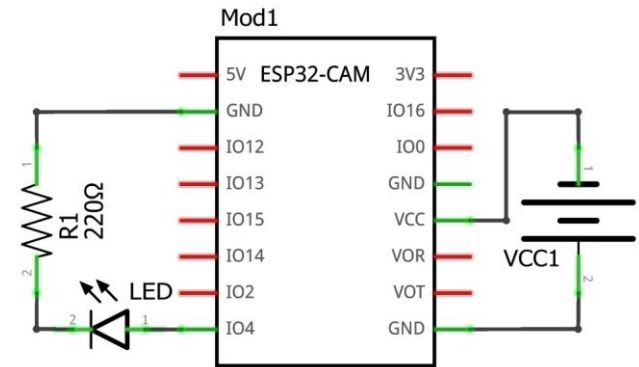
Fig. 3. Schematic diagram.

Fig. 3 illustrates the schematic diagram of the Smart Water Meter, showcasing the essential components of the electronic circuit setup for this project. The ESP32 Cam is a highly functional microcontroller board that integrates a powerful ESP32 chip, along with built-in Wi-Fi and Bluetooth capabilities. It serves as the primary component of the setup. Additionally, the hardware includes a camera module, enabling the device to stream video and capture high-resolution images. Furthermore, the microcontroller and an SD card are utilized to store the programmed code, offering versatile and expandable storage options. To enhance the lighting conditions for image capture, an additional LED is incorporated to illuminate the surroundings.

Regarding the power supply, this project utilizes a DC power source. To ensure compatibility with the hardware components, a 5V converter is employed. Acting as a crucial intermediary, the converter transforms the incoming AC power into a reliable 5V DC supply, which powers the main component of the project. This configuration guarantees a consistent and dependable power source, facilitating the smooth operation of the hardware and efficient execution of the programmed tasks.

3. Results and Discussions

3.1. Hardware Setup

A custom casing as shown in Fig. 4 has been designed specifically for the ESP32 Cam to serve as storage. This casing is positioned above the water meter and securely fixed in place using supporting pillars. This ensures that the captured images remain clear and avoid any blurring or distortion.
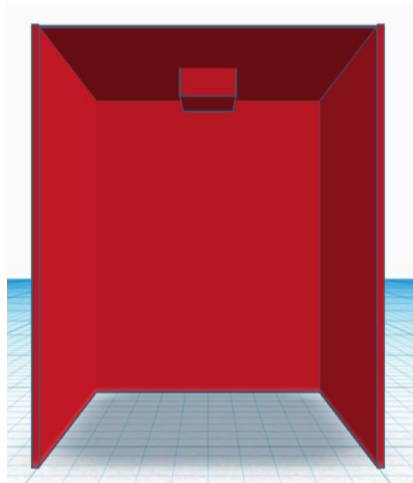
Fig. 4. Device casing - front view.

(a)

(b)

Fig. 5. (a) Water usage page and (b) water bill page.

3.2. Mobile Application

A mobile application has been developed with various functionalities, including the ability to scan the Smart Water Meter to establish a connection with the app. It enables users to monitor their water consumption and generates bills based on their usage. The usage page provides insights into the user's water consumption, while the billing page generates the corresponding water bill. The app retrieves data directly from the connected smart water meter, ensuring accurate and up-to-date information.

With this mobile application, users can easily track their latest water usage and view the associated bill to understand the amount they need to pay. This functionality promotes increased awareness of current water consumption, leading to potential cost savings. By being more mindful of their water usage, users can effectively manage their consumption and make informed decisions to optimize their expenses. Fig. 5 illustrates the display of information in the developed app, specifically showcasing water usage details and the current usage bill. This information can be accessed within the app and has the capability to be printed for reference or documentation purposes.

3.3. Data Analysis

Thorough analysis and discussion are conducted on the collected data, considering the output results. This data is obtained using the smart water meter device to retrieve the meter reading. All the collected data is extracted into an excel file to generate a graph for visualization as shown in Fig. 6. The blue line represents the value obtained after image processing, while the orange line represents the actual reading recorded from the captured image of the water meter.

Fig. 6 reveals the presence of spikes in the graph, indicating instances where the image processing detected false numbers from the water meter images. The primary cause of this issue is likely attributed to variations in the lighting conditions during image capture. The most common error occurs when detecting the last 3-digit number on the meter. Within the collected data, a cluster of errors is observed in the middle portion. This is attributed to a specific number repeatedly reading inaccurately, leading to the appearance of multiple spikes. However, as the dial turns to a different number, the errors diminish, resulting in fewer spikes toward the end of the data.

Another factor contributing to the occurrence of spikes is the presence of different background colors for the digits, specifically the red color shown in Fig. 1. The AI model used for the project was trained on images with a white background, which impacts the accuracy of the model and prevents it from reaching the desired level of performance.
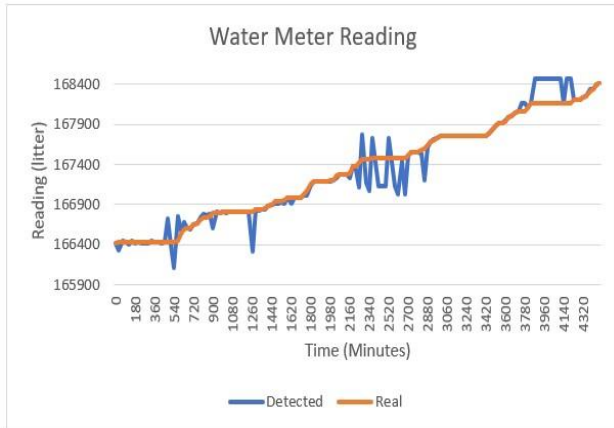
Fig. 6. Comparison of manual water meter actual reading between reading detected by the camera using image processing technique.

3.4. Data Accuracy

The spike in Fig. 6 may happen due to the different background color of the digits which is red color. The accuracy is only 3.33% for the device detecting all 8-digit numbers on the water meter. By reducing the last digit number with the red background, the accuracy has increased significantly to 53.33%. It is proved that the AI model can detect all the digits with white background as the accuracy for the first 5 digits is 100% accurate. The AI model is not trained with red background digit numbers, hence the accuracy for detecting the meter reading is not that high. Therefore, this explains the accuracy result in Fig. 7 that this AI model used has a low accuracy for detecting the meter digit number with red background.
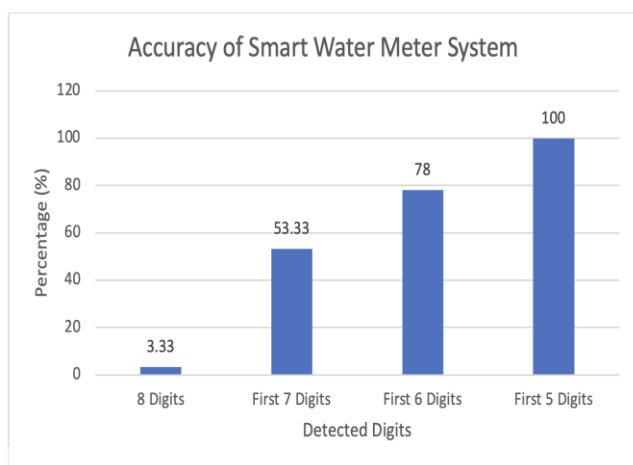
Fig. 7. White and red background effects on water scale meter accuracy.

4. Conclusion

Based on the aforementioned results, the system operates smoothly without encountering any critical errors, despite the detected data not being highly accurate. The accuracy of the data primarily depends on the AI model utilized for image processing. This discrepancy arises because the AI model was trained on a different type of water meter, distinct from the one employed in this project. Without proper training, the model's accuracy may be compromised. Nonetheless, the overall performance of the AI model is commendable, as it accurately identifies digits with a white background, exhibiting significantly higher accuracy than those with a red background. Thus, it can be concluded that a properly trained dataset greatly enhances accuracy. Utilizing the collected data enables us to ascertain the daily water usage for each household. By leveraging this data, predictions can be made to determine the water consumption in different areas, alerting the water supply company accordingly. This valuable information empowers the company to identify areas requiring additional water supply, ensuring residents do not face water shortages. Over time, as more data is gathered, the predictive capabilities of the system improve, resulting in more accurate forecasts. In summary, this IoT-based metering device resolves the issue of manpower required for water meter data collection while also offering long-term cost savings.

5. Future Recommendation

For the future implementation of this smart water meter system, there is a need to enhance the design of the device casing. The design should prioritize reliability and adhere to sound design principles. Additionally, the issue of sunlight penetration during the daytime must be addressed by reimagining the casing product. Furthermore, to address the common problem faced by all electronic devices, a comprehensive plan for sustainable energy usage in the power supply is necessary. Additionally, to improve the accuracy of the system, alternative algorithms can be explored and employed in conjunction with the AI model. Lastly, the mobile application can be enriched with additional features to enhance user experience. For example, incorporating a water leak detection feature would enable users to receive alerts regarding the condition of their pipes. Moreover, the billing system can be enhanced by expanding the range of payment methods, allowing for convenient online payments and improving user-friendliness.

Acknowledgment

Saw Di Wen, Hazry Desa, Muhammad Azizi Azizan, Abadal-Salam T. Hussain, Muhammad Hassan Tanveer,nd Rizwan Patan

References

1. M. Nadason, Joint Press Release by FOMCA, WECAM, FAM: Celebrating World Water Day 2021 – Valuing Water, March, 2022.
2. H. Abdul Rahman, A Review on Water Issues in Malaysia, International Journal of Academic Research in Business and Social Sciences, Vol. 11(8), August, 2021, pp. 860-875.
3. S. Boatwright, S. Mounce, M. Romano and J. Boxall, Integrated Sensor Placement and Leak Localization Using Geospatial Genetic Algorithms, Journal of Water Resources Planning and Management, Vol. 149 (9), June, 2023.
4. L. Canale, T. Cholewa, G. Ficco, A. Suita-Olcha, D. Di Petra, P. Kolodziej and M. Dell'Isola, The Role of Individual Metering in Reducing Domestic Hot Water Consumption in Residential Buildings: A Long-Term Evaluation, Journal of Building engineering, Vol. 73, August, 2023.
5. M. Kavya, A. Mathew, P. R. Shekar and P. Sarwesh, Short Term Water Demand Forecast Modelling Using Artificial Intelligence for Smart Water Management, Sustainable Cities and Society, Vol. 95, August, 2023.
6. N.S. Muhammad, J. Abdullah, N. A. Rahman and N. A. Razali, Water Usage Behaviour: Case Study in a Southern State in Peninsular Malaysia, IOP Conference Series: Earth and Environmental Science, Vol. 646, 2021.
7. J. Cassidy, B. Barbosa, M. Damião, P. Ramalho, A. Ganhão, A. Santos and J. Feliciano, Taking Water Efficiency to the Next Level: Digital Tools to Reduce Non-Revenue Water, Journal of Hydroinformatics, Vol. 23(3), May, 2021, pp. 453-465.
8. J. Marais, R. Malekian, N. Ye and R. Wang, A Review of the Topologies Used in Smart Water Meter Networks: A Wireless Sensor Network Application, Journal of Sensors, Vol. 2016, October, 2016.
9. X. Cao, Design of Remote Water Meter Reading System Based on GPRS Technology, Journal of Physics: Conference Series, Vol. 1601, 2020.
10. T. Csaba, Challenges of Smart Water Metering, WM Systems LLC - Innovation in Smart IoT Systems, Mar. 11, 2021.
11. GitHub. Retrieved January 7, 2023, from https://github.com/jomjol/AI-on-the-edge-device

Authors Introduction

Mr. Saw Di Wen

He completed his bachelor's degree in Mechatronic Engineering from Universiti Malaysia Perlis and currently works as a Mechanical Design Engineer at Pentamaster Equipment Manufacturing Sdn Bhd in Penang, Malaysia.

Prof. Dr. Hazry Desa

He earned his Bachelor of Mechanical Engineering from Tokushima University, Japan, and proceeded to pursue his Ph.D. at the Artificial Life and Robotics Laboratory, Oita University, Japan, in 2003. Presently, he serves as the director at the Center of Excellence for Unmanned Aerial Systems (COE-UAS), Universiti Malaysia Perlis, Malaysia.

Dr. Muhammad Azizi Azizan

He received his PhD in Civil Engineering from the Universiti Malaysia Perlis. He is currently a senior lecturer in the same institution. He is head of Project Integration & Management (PIM) at Centre of Excellence for Unmanned Aerial System (COE-UAS), Universiti Malaysia Perlis.

Dr. Abadal-Salam T. Hussain

He is currently a Ph.D. faculty assistant professor and head of the department for Medical Instrumentation and Technique Engineering at Alkitab University. Previously, he was a staff member in the Faculty of Electrical Engineering Technology at Universiti Malaysia Perlis (UniMAP), Malaysia.

Dr. Muhammad Hassan Tanveer

He is an Assistant Professor of Robotics and Mechatronics Engineering at Kennesaw State University. Prior to that, he was a Post Doctoral Research Associate at Virginia Tech, USA where he was working on development of Bio-sonar sensor for UAVs navigations. He received his Ph.D. in Robotics and Autonomous Sytems with specilization in 'Collaboration of Heterogenous Team of Robots' from University of Genova, Italy, in 2019.

Dr. Rizwan Patan

He is currently working as a Assistant Professor in the Department of Software Engineering and Game Development at Kennesaw State University, Marietta, USA. He receives his Ph.D. in 2017 from the School of Computing Science and Engineering at the Vellore Institute of Technology.

Object Detection and Instance Segmentation with YOLOV8: Progress and Limitations

Law Jack Lee, Hazry Desa, Muhammad Azizi Azizan
Centre of Excellence for Unmanned Aerial Systems (COE-UAS), Universiti Malaysia Perlis, Block E, Pusat Perniagaan Pengkalan Jaya, Jalan Kangar – Alor Setar, 01000 Kangar, Perlis.

Abadal-Salam T. Hussain
Department of Medical Equipment Technology Engineering, College of Engineering Technology, Al-Kitab University, Kirkuk, Iraq.

Muhammad Hassan Tanveer
Department of Robotics and Mechatronics Engineering, Kennesaw State University, Marietta, GA, 30067, USA.

Email: hazry@unimap.edu.my
www.unimap.edu.my

Abstract

This research employs object detection and instance segmentation algorithms to distinguish between objects and backgrounds and to interpret the detected objects. The YOLOV8 (You Only Look Once) framework and COCO dataset are utilized for detecting and interpreting the objects. Additionally, the accuracy of detection, segmentation, and interpretation is tested by placing objects at various distances from the camera. The algorithm's performance was evaluated, and the results were documented. In the experiments, a sample of 11 objects was tested, and 8 of them were successfully detected at distances of 45cm, 75cm, 105cm, and 135cm. For instance, segmentation, segmentation maps appeared clean when detecting a single object but faced challenges when multiple objects overlapped.

Keywords: Image Processing YoloV8, COCO, Segmentation, Object detection.

1. Introduction

Traditional methods in object detection rely on hand-crafted features and rules to identify and categorize objects, whereas machine learning-based approaches involve algorithms trained on extensive labeled image datasets. Notably, machine learning-based methods offer the advantage of handling complex and diverse environments by learning and adapting to new data [1]. However, they can be computationally demanding and require substantial labeled data for training. In the past few years, there have been notable strides in object detection algorithms, especially with the rise of deep learning-centered methods that have demonstrated cutting-edge capabilities on a variety of benchmarks. These methods utilize neural networks to learn features and classify objects, successfully finding applications in various domains [2], [3]. Nevertheless, prevailing monitoring systems might exhibit limited adaptability in computer vision, leading to difficulties in accurately detecting and classifying objects or features within visually intricate settings, including those characterized by low illumination, high contrast, or complex backgrounds [4]. Furthermore, current monitoring setups often heavily rely on human operators to analyze images and spot potential issues [5]. This manual procedure can be time-intensive and prone to errors, as human operators might encounter challenges in consistently and precisely interpreting images or detecting subtle anomalies [4]. To confront these challenges, this project introduces an object detection and self-interpreting application designed to augment monitoring systems.

2. Related Work

The research landscape in computer vision and deep learning has witnessed significant advancements in recent years, particularly in the domain of object detection and segmentation. Notably, Kousik et al. proposed a hybrid Convolution Recurrent Neural Network for improved salient object detection [9]. Additionally, Mekala et al. introduced a Deep Reinforcement Learning (DRL) based 4-r Computation Model for Object Detection on Roadside Units (RSU) using LiDAR on the Internet of Things (IoT) context [10]. Deep learning and image processing techniques for cancer diagnosis were effectively employed by Prassanna et al. [11], while Ghantasala et al. focused on texture recognition and image smoothing for microcalcification

Law Jack Lee, Hazry Desa, Muhammad Azizi Azizan, Abadal-Salam T. Hussain, Muhammad Hassan Tanveer

and mass detection in abnormal regions [12]. In the context of the Internet of Medical Things, Manimurugan et al. utilized a deep belief neural network for effective attack detection in smart environments [13]. Furthermore, Natarajan et al. applied fully convolutional deep neural architecture for the segmentation of nuclei in histopathology images [14]. Lastly, Selvaraj et al. proposed an optimal virtual machine selection approach using swarm intelligence for anomaly detection [15]. These diverse works collectively contribute to the broader understanding and application of advanced techniques in the realm of computer vision and deep learning.

3. Methodology

Overall Object Detection and Instance Segmentation Segmentation is a specific type of image analysis that focuses on identifying individual instances of objects and outlining their boundaries [6]. Two essential computer vision tasks, semantic segmentation, and object detection are closely related to instance segmentation [7]. Semantic segmentation seeks to correctly assign the appropriate object category to every individual pixel present within an image. However, it fails to differentiate among separate instances of identical object classes. For instance, if there are three dogs in an image, semantic segmentation will only identify pixels belonging to one of these three dogs.

Conversely, object detection involves predicting both the object's bounding box and its corresponding class for every instance of an object present in the image. However, it does not provide per-pixel identification for each object instance. Hence, object instance segmentation is notably more challenging than semantic segmentation and object detection because it seeks to both identify and label each object instance on a per-pixel basis [7].

To create a dataset, for example, one can collect images of cats and dogs, organizing them into two folders: "cats" and "dogs." Each folder should contain a similar number of images. In each folder, a total 80% of the images are allocated as training images, and the remaining 20% are used as test images. Annotations are required for the training images, involving manually drawing bounding boxes and segmenting the dogs or cats from the background. These annotations produce coordinates that are recorded in a blank text file, essential for the machine's ability to locate objects during image training and testing. Fig. 1 displays an example of object detection, with a bounding box around the Region of Interest (ROI), labelled with its name, and segmented with a coloured mask layer within the ROI, utilizing the COCO dataset [8].

Fig. 1. Object detection and instance segmentation.

4. Results and Discussion

4.1 Object Detection and Self-Interpreting Application Within Different Distance.

To test the ability of detection and interpreting application, the experiment was carried out in room. 11 objects were selected, and each object was placed in 45cm, 75cm, 105cm and 135cm apart from the webcam. The results were recorded in Table 1 below. A tick indicates that the detection was successful, and a cross indicates that the detection was unsuccessful.
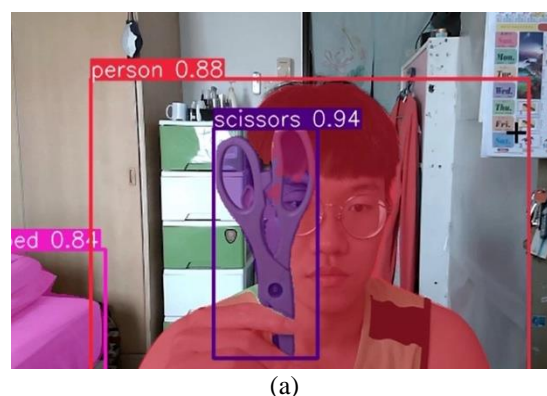
(a)

(b)

Fig. 2. (a) Scissor at 45 cm distance and (b) scissor at 75 cm distance from the camera.

(a)

(b)
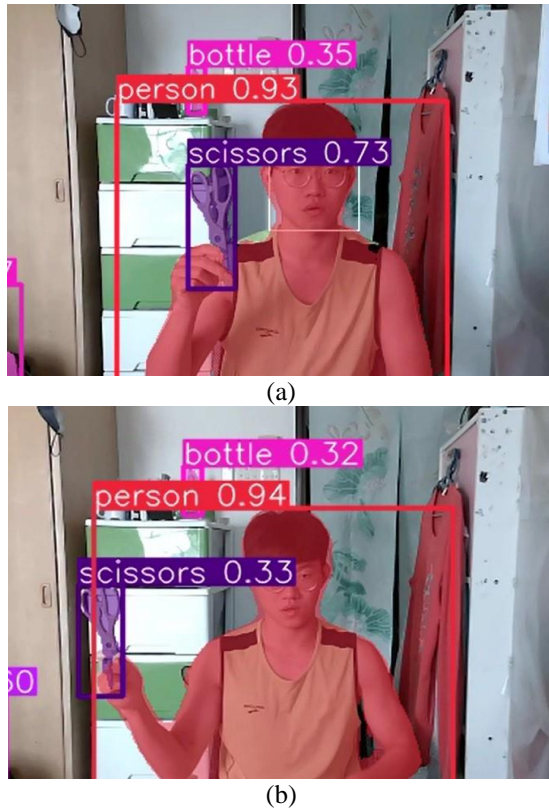
Fig. 3. (a) Scissor at 105 cm distance and (b) scissor at 135 cm distance from the camera.

failed to do so at 75cm, 195cm, and 135cm distances. Such limitations arise from the detection and segmentation networks, as overlapping objects pose challenges in separation and precise localization. The output image may look acceptable, but it merely represents the network's best estimation of object locations. When multiple objects overlap, the model may struggle, resulting in inaccurate bounding boxes or masks, which can sometimes lead to incorrect labels.

(a)

(b)
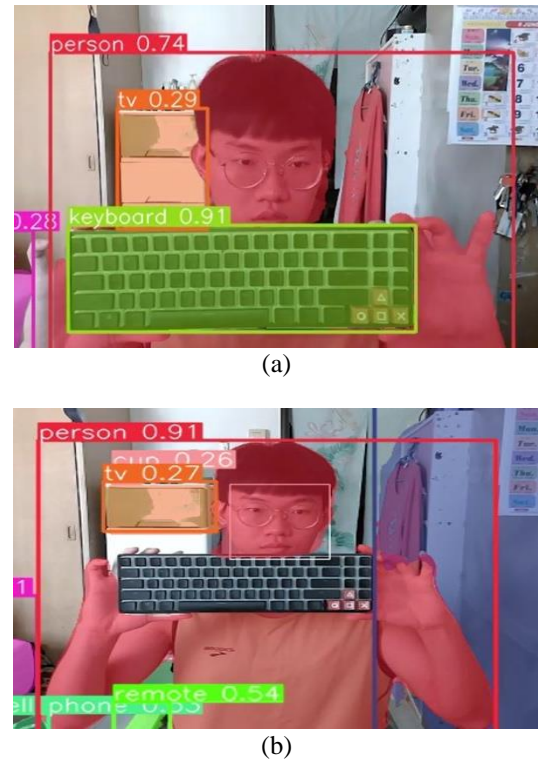
Fig. 4. (a) 45 cm distance (b) 75 cm distance away from the camera.

Fig. 2 and Fig. 3 present examples of various distances captured by the webcam. The figures show a pair of scissors, accurately detected, and identified by the system. Most large objects, such as a person, cell phone, keyboard, and bowl, were successfully detected and identified. However, at distances of 105 cm and 135 cm, the objects like the mouse, fork, and spoon were not recognized. This lack of recognition was primarily due to object occlusion. When the fork and spoon were held at 105 cm, the student's fingers obstructed part of the objects, making it challenging for the Yolo model to detect them. Furthermore, the webcam's resolution also impacted the Yolo model's detection rate. Although the webcam supported 1080p Full HD video quality, the output was compressed to 640px x 640px for smoother video operation. It is essential to note that this experiment was conducted on a CPU-only machine, resulting in significantly lower performance compared to using a GPU-supported machine.

Instance Segmentation Within Different Distances. The results obtained from the YOLOV8 model for instance segmentation were observed and collected. When a single object was detected at distances of 45cm, 75cm, 105cm, and 135cm, the segmentation maps of a person appeared to be clean, as evident from Fig. 2 and Fig. 3. However, in cases where multiple objects overlapped, the model encountered difficulties in accurate segmentation. Fig. 4 and Fig. 5 demonstrate the inconsistent segmentation of a keyboard at different distances. While the keyboard could be segmented successfully at 45cm when accompanied by a human, it

(a)

Law Jack Lee, Hazry Desa, Muhammad Azizi Azizan, Abadal-Salam T. Hussain, Muhammad Hassan Tanveer

(b)

Fig. 5. (a) 105 cm distance (b) 135 cm distance away from the camera.

Table 1. Instances of successful and unsuccessful attempts of being detected.

Objects	Distance away from camera (cm) on UAV				
	45	75	105	135	Percentage (%)
Scissors	✓	✓	✓	✓	100
Person	✓	✓	✓	✓	100
Bottle	✓	✓	✓	✓	100
Cup	✓	✓	✓	✓	100
Cellphone	✓	✓	✓	✓	100
Mouse	✓	✓	×	×	50
Keyboard	✓	✓	✓	✓	100
Bowl	✓	✓	✓	✓	100
Spoon	✓	✓	×	×	50
Fork	✓	✓	×	×	50
Toothbrush	✓	✓	✓	×	75

5. Conclusion

The experiment involving object detection and a self-interpreting system was carried out successfully, employing the YOLOV8 framework and the COCO dataset to collect the results. Out of the 11 objects that underwent testing, 8 of them were effectively identified at different distances: specifically, 45cm, 75cm, 105cm, and 135cm. This led to the algorithm used in the experiment achieving an accuracy rate of 72% in terms of object detection and interpretation. However, there were instances where 4 objects failed to be detected accurately at distances of 105cm and 135cm. This was attributed to the lower image quality at these distances. Notably, the segmentation maps displayed clean results when a single object was detected, but inconsistency in segmentation maps arose when multiple objects overlapped, mainly due to the limitations of the detection and segmentation network. In conclusion, while the experiment demonstrated promising results with a 72% accuracy in object detection and interpretation, challenges persisted in cases of poor image quality and overlapping objects, affecting the segmentation performance.

Acknowledgements

References

1. H. Ishibuchi, T. Nakashima and T. Kuroda, A Fuzzy Genetics-Based Machine Learning Method for Designing Linguistic Classification Systems with High Comprehensibility, ICONIP'99. ANZIIS'99 & ANNES'99 & ACNN'99, 6th International Conference on Neural Information Processing, August, 2002.
2. J. Han, D. Zhang, G. Cheng, N. Liu and D. Xu, Advanced Deep-Learning Techniques for Salient and Category-Specific Object Detection: A Survey, IEEE Signal Processing Magazine, Vol. 35(1), January, 2018, pp. 84-100.
3. Z. -Q. Zhao, P. Zheng, S. -T. Xu and X. Wu, Object Detection with Deep Learning: A Review, IEEE Transactions on Neural Networks and Learning Systems, Vol. 30(11), November, 2019, pp. 3212-3232.
4. S. Gundu, H. Syed and J. Harikiran, Human Detection in Aerial Images using Deep Learning Techniques, 2nd International Conference on Artificial Intelligence and Signal Processing (AISP), April, 2022, pp. 1-10.
5. R. A. Suárez Fernández, J. L. Sanchez-Lopez, C. Sampedro, H. Bavle, M. Molina and P. Campoy, Natural User Interfaces for Human-Drone Multi-Modal Interaction, International Conference on Unmanned Aircraft Systems (ICUAS), July, 2016, pp. 1013-1022.
6. A. M. Hafiz and G. M. Bhat, A Survey on Instance Segmentation: State of the Art, International Journal of Multimedia Information Retrieval, Vol. 9(3), July, 2020, pp. 171-189.
7. L. Ye, Z. Liu and Y. Wang, Depth-Aware Object Instance Segmentation, IEEE International Conference on Image Processing (ICIP), February, 2018, pp. 325-329.
8. T. Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Doll´ar and C. L. Zitnick, Microsoft COCO: Common Objects in Context, Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September, 2014.
9. N. Kousik, Y. Natarajan, R. A. Raja, S. Kallam, R. Patan and A. H. Gandomi, Improved Salient Object Detection Using Hybrid Convolution Recurrent Neural Network, Expert Systems with Applications, Vol. 166, March, 2021.
10. M. S. Mekala, R. Patan, A. H. Gandomi, J. H. Park and H. Y. Jung, A DRL based 4-r Computation Model for Object Detection on RSU using LiDAR in IIoT. IEEE Symposium Series on Computational Intelligence (SSCI), January, 2022, pp. 01-08.
11. J. Prassanna, R. Rahim, K. Bagyalakshmi, R. Manikandan and R. Patan, Effective Use of Deep Learning and Image Processing for Cancer Diagnosis, Deep learning for Cancer Diagnosis, September, 2020, pp. 147-168.
12. G. P. Ghantasala, S. Kallam, N. V. Kumari and R. Patan, Texture Recognization and Image Smoothing for Microcalcification and Mass Detection in Abnormal Region, International Conference on Computer Science, Engineering and Applications (ICCSEA), July, 2020.
13. S. Manimurugan, S. Al-Mutairi, M. M. Aborokbah, N. Chilamkurti, S. Ganesan and R. Patan, Effective Attack Detection in Internet of Medical Things Smart Environment Using a Deep Belief Neural Network, IEEE Access, Vol. 8, April, 2020, pp. 77396-77404.

14. N. A. Natarajan, M. S. Kumar, R. Patan, S. Kallam and M. Y. N. Mohamed, Segmentation of Nuclei in Histopathology Images Using Fully Convolutional Deep Neural Architecture, International Conference on Computing and Information Technology (ICCIT-1441), November, 2020, pp. 319-325.
15. A. Selvaraj, R. Patan, A. H. Gandomi, G. G. Deverajan and M. Pushparaj, Optimal Virtual Machine Selection for Anomaly Detection Using a Swarm Intelligence Approach, Applied Soft Computing Journal, Vol. 84, November, 2019.

Dr. Muhammad Hassan Tanveer

He is an Assistant Professor of Robotics and Mechatronics Engineering at Kennesaw State University. Prior to that, he was a Post Doctoral Research Associate at Virginia Tech, USA where he was working on development of Bio-sonar sensor for UAVs navigations. He received his Ph.D. in Robotics and Autonomous Sytems with specilization in 'Collaboration of Heterogenous Team of Robots' from University of Genova, Italy, in 2019.

Authors Introduction

Mr. Law Jack Lee

He completed his bachelor's degree in Mechatronic Engineering from Universiti Malaysia Perlis and currently works as a Mechanical Design Engineer at Pentamaster Equipment Manufacturing Sdn Bhd in Penang, Malaysia.

Prof. Dr. Hazry Desa

He earned his Bachelor of Mechanical Engineering from Tokushima University, Japan, and proceeded to pursue his Ph.D. at the Artificial Life and Robotics Laboratory, Oita University, Japan, in 2003. Presently, he serves as the director at the Center of Excellence for Unmanned Aerial Systems (COE-UAS), Universiti Malaysia Perlis, Malaysia.

Dr. Muhammad Azizi Azizan

He received his PhD in Civil Engineering from the Universiti Malaysia Perlis. He is currently a senior lecturer in the same institution. He is head of Project Integration & Management (PIM) at Centre of Excellence for Unmanned Aerial System (COE-UAS), Universiti Malaysia Perlis.

Dr. Abadal-Salam T. Hussain

He is currently a Ph.D. faculty assistant professor and head of the department for Medical Instrumentation and Technique Engineering at Alkitab University. Previously, he was a staff member in the Faculty of Electrical Engineering Technology at Universiti Malaysia Perlis (UniMAP), Malaysia.

A high-performance motion planning method based on asymptotically optimal RRT

Tianbin Meng
School of Mechanical and Electronic Engineering, Beijing Jiaotong University
Beijing, Haidian District, China

Jiwu Wang
School of Mechanical and Electronic Engineering, Beijing Jiaotong University
Beijing, Haidian District, China
E-mail: mengtianbin888@163.com, jwwang@bjtu.edu.cn
www.bjtu.edu.cn

Abstract

The motion planning algorithm of the robot arm plays an important role in the working process of the robot arm, especially in the complex environment, an efficient algorithm is more conducive to the robot arm to complete the corresponding planning task. Aiming at the problems such as low exploration efficiency and poor planning path in the current motion planning task of robotic arm, we propose a distance constraint mechanism. Based on RRT^* and $Informed-RRT^*$, the algorithm uses halton sequence to generate random points and introduces the current lowest cost path length. To avoid useless nodes extension. The simulation results show that the algorithm with distance constraint mechanism can improve the exploration efficiency and planning quality to some extent.

Key words: mechanical arm, Motion planning, Distance constraint, Improved RRT^* algorithm, halton sequence

1. Introduction

The robot arm has changed from the original rough industrial production equipment to today's multi-disciplinary industrial automation equipment, entering all corners of society, showing the powerful role of the robot arm. At the same time, higher requirements are also put forward for the motion planning of the robotic arm. Therefore, the study of the motion planning of the robotic arm has important engineering value [1], [2].

At present, there are four main types of path planning methods for robotic arms: graph search-based algorithm, artificial potential field method, random sampling method and deep reinforcement learning method [3]. The algorithm based on graph search needs to be discretized in the whole planning space, so it is not suitable for the space with higher latitude. The artificial potential field method is easy to make the manipulator fall into the local minimum under the influence of the influence field, and the efficiency is greatly reduced. Methods based on deep reinforcement learning also have problems of poor process interpretability and stability [4].

The method based on random sampling is one of the most common methods for path planning of manipulators. Its core idea is to replace the complex actual planning space by the simple planning space obtained by sampling, so as to reduce the space complexity. the method based on random sampling is efficient and suitable for high-dimensional space. Lavalle [5] proposed a classical RRT algorithm, which has strong exploration ability and

probability completeness, but cannot guarantee the optimal solution and a large number of unnecessary waypoint searches. Frazzoli [6] proposed RRT^* on the basis of RRT, which is an algorithm with asymptotic optimality. Compared with RRT, it adds two processes of re-selecting parent nodes and rewiring, Therefore, an optimal or suboptimal path can be found. Gammell [7] improved on the basis of RRT^* and proposed the $Informed-RRT^*$ algorithm. The principle is that after finding the first feasible solution, the subsequent sampling range is limited to a high-dimensional ellipsoid determined by the cost of the feasible solution. The convergence speed of RRT^* is improved.

To sum up, in order to solve the problem of low efficiency and too many redundant nodes in the RRT^* computing planning. In this paper, a distance constraint mechanism is introduced to improve RRT^* and $Informed-RRT^*$, which can greatly reduce the search and expansion of unnecessary waypoints and plan a high-quality path.

2. Methodology
2.1. Random tree growth distance constraint mechanism

RRT^* algorithm has asymptotic optimality, but the random number growth in the sampling process has a strong blindness, thus reducing the efficiency of the algorithm. $Informed-RRT^*$ Although the sampling points can be limited to an ellipse, the expansion of random trees within the ellipse is still blind. Therefore, this paper proposes a distance constraint mechanism, which introduces the current lowest cost path length. When the

random tree is extended, the minimum actual cost to reach the new node is calculated plus the Euclidean distance from the new node to the end point. If the estimated distance is greater than the current lowest cost path length, the new node will not be extended, which increases the search efficiency of the algorithm to a certain extent. A condition that determines whether a new node can be expanded:

$$g_T(x_{new}) + \hat{h}(x_{new}) < c_{best} \qquad (1)$$

Where $g_T(x)$ represents the minimum actual cost from the starting point to the new node, $\hat{h}(x)$ represents the Euclidean distance from the new node to the end point, and c_{best} is the current lowest cost path length.

Similarly, the current lowest cost path length is also introduced in the process of rewiring in random trees. When a new node is selected as the parent node at a certain point in the neighborhood, if the minimum actual cost to reach the point plus the Euclidean distance from the point to the end point is greater than the current lowest cost path length, the parent node of the node will not be replaced. A condition that determines whether a node in the neighborhood can replace the parent node:

$$g_T(x_{new}) + ||x_{new} - x_{near}|| + \hat{h}(x_{near}) < c_{best} \qquad (2)$$

Where $||x_{new} - x_{near}||$ indicates the Euclidean distance between x_{new} and x_{near}.

2.2. Random sampling point optimization strategy

RRT^* and $Informed - RRT^*$ algorithms have the problem of poor uniformity in the random sampling of points in space, so the space cannot be explored more fully. A random point sampling method based on halton sequence is proposed for reference to halton sequence. Random sampling method (Fig.1) and halton sequence-based sampling method (Fig.2) are shown in the figure below.
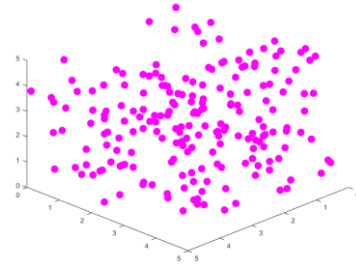
Fig.1 Random sampling method
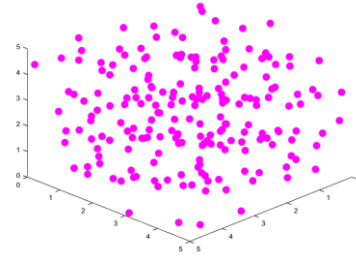
Fig.2 Sampling method based on halton sequence

It can be observed that the sampling points in the upper left part of Fig. 1 are sparse, while the middle part is dense. In Fig. 2, the distribution of sampling points is more uniform, thus improving the ability to explore the space.

3. Algorithm simulation analysis

In order to verify the validity of the distance constraint mechanism, the RRT^*, $Informed - RRT^*$ and the improved $DC - RRT^*$ and $DC - Informed - RRT^*$ are simulated and analyzed experimentally in a three-dimensional environment.

In MATLAB, simulation experiments in simple and complex obstacle avoidance environments are set respectively. The simulation environment was set to the three-dimensional space of $[0.2 \times 0.12 \times 0.2]$, the search starting point was $[-0.09\ 0.37\ 0.1]$, the target point was $[0.09\ 0.43\ 0.1]$, the exploration step was 0.01, and the number of iterations was set to 100 to reduce the randomness of the algorithm. The experimental results are shown in the figure and table below.

a. RRT^* algorithm tracing

b. $DC - RRT^*$ algorithm tracing

c. $Informed - RRT^*$ algorithm tracing

d. $DC - Informed - RRT^*$ algorithm tracing
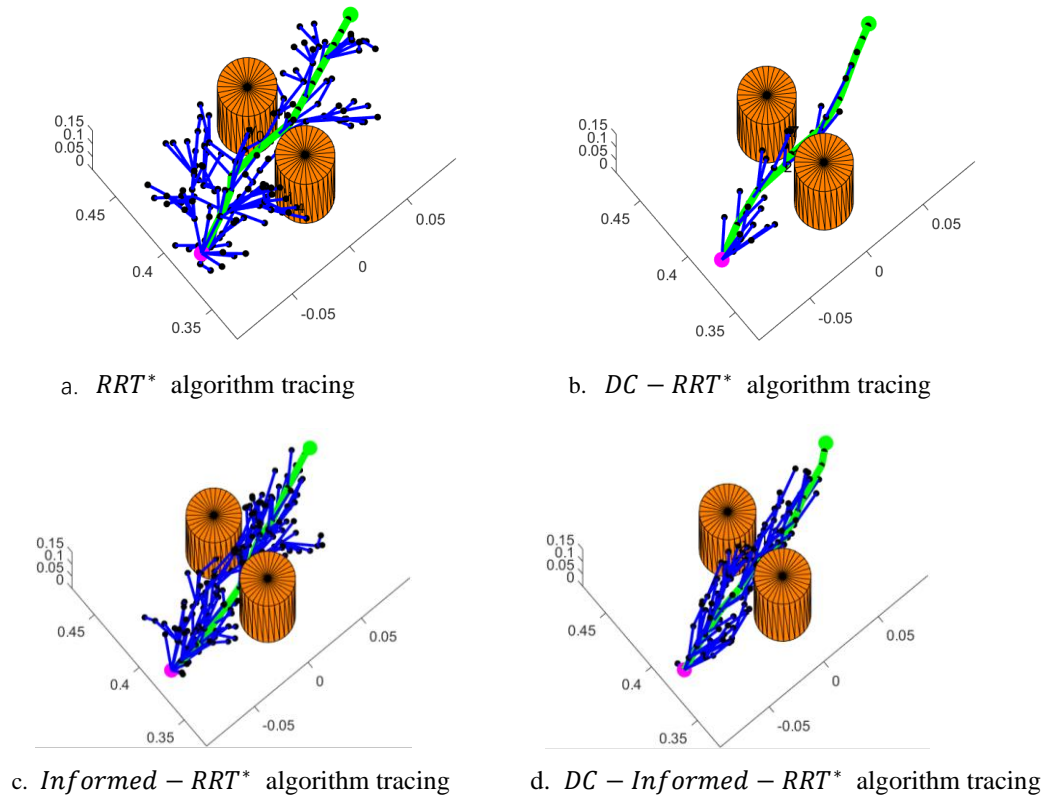
Fig.3 Algorithm tracing in 3D simple scenes

Table 1.   Performance comparison of algorithms in 3D simple scenes.

algorithm	average path length /m	average running time /s	Node usage /%
RRT^*	0.2042	6.67	6.62%
$DC - RRT^*$	0.1934	5.41	10.96%
$Informed - RRT^*$	0.1930	8.42	6.38%
$DC - Informed - RRT^*$	0.1906	5.72	12.31%

As shown in Fig. 3, The blue line represents the extended branch, and the green line represents the final path. RRT^* produces many useless branches. $Informed - RRT^*$ also has some useless branches, but fewer than RRT^*. The final path of the two is tortuous and the path length is long. After the introduction of distance constraint mechanism, the useless branches of $DC - RRT^*$ and $DC - Informed - RRT^*$ are significantly reduced, the final path is smoother and the path length is shorter. Combined

with the data in Table 1, the average path of $DC - RRT^*$ is reduced by 5.29% from 0.2042m to 0.1934m compared with RRT^*. The average running time decreased by 18.90% from 6.67s to 5.41s; The node utilization rate is increased from 6.62% to 10.96%, and the average path of $DC - Informed - RRT^*$ and $Informed - RRT^*$ is reduced from 0.1930m to 0.1906m by 1.24%. The average running time decreased by 32.07% from 8.42s to 5.72s; Node usage increased from 6.38% to 12.31%.

Tianbin Meng, Jiwu Wang

a. RRT^* algorithm tracing

b. $DC - RRT^*$ algorithm tracing

c. $Informed - RRT^*$ algorithm tracing

d. $DC - Informed - RRT^*$ algorithm tracing
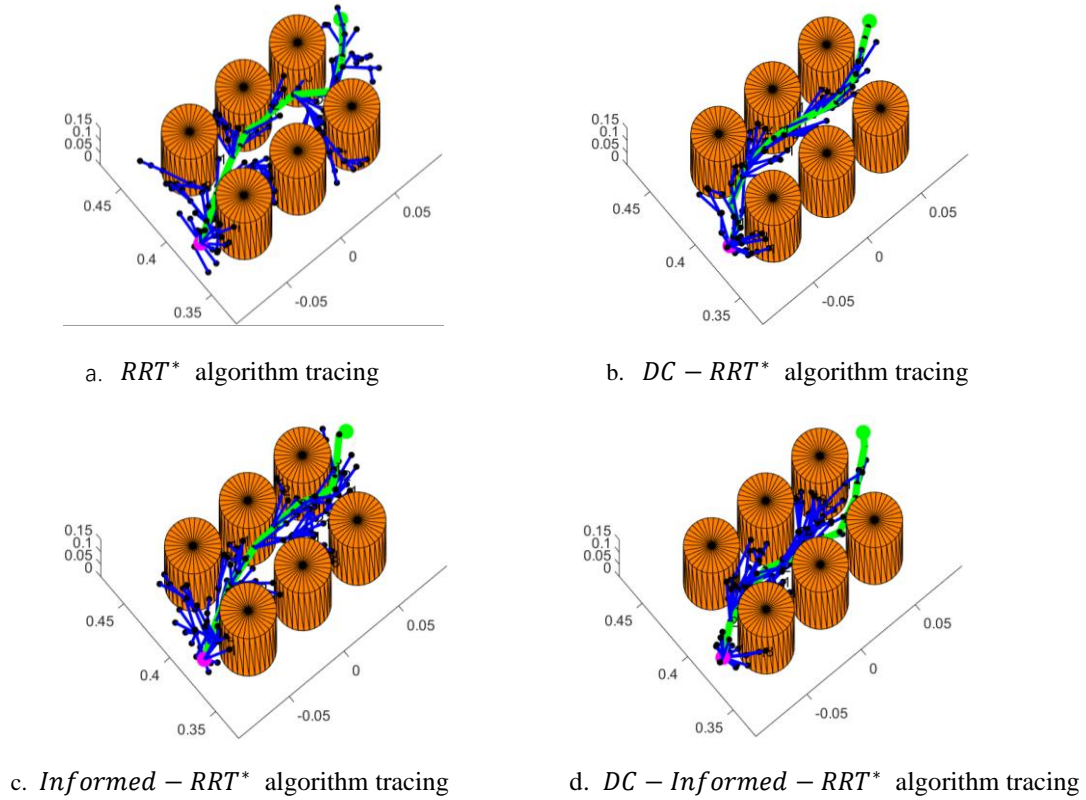
Fig.4 Algorithm tracing in 3D complex scenes

Table 2.   Performance comparison of algorithms in 3D complex scenes.

algorithm	average path length /m	average running time /s	Node usage /%
RRT^*	0.2191	17.16	7.63%
$DC - RRT^*$	0.2056	18.42	14.75%
$Informed - RRT^*$	0.2003	20.41	7.01%
$DC - Informed - RRT^*$	0.1965	21.95	16.66%

Fig. 4 shows the planning results of the algorithm in a complex environment. Combined with the data in Table 2, the average path of $DC - RRT^*$ decreases by 6.16% from 0.2191m to 0.2056m compared with RRT^*. The average running time is 17.16s and 18.42s respectively, the node utilization rate is increased from 7.63% to 14.75%, and the average path of $DC - Informed - RRT^*$ is reduced from 0.2003m to 0.1965m, a decrease of 1.89%. The average running time was 20.41s and 21.95s respectively, and the node utilization rate increased from 7.01% to 16.66%.

According to the analysis, since the introduction of distance constraint mechanism reduces the expansion of nodes that are useless for optimizing the path, the node utilization rate is higher, and the node distribution is more concentrated near the final path, then the node is more likely to optimize the cost of the current path. Secondly, although the introduction of distance constraint mechanism reduces the expansion of useless nodes, the nodes are more concentrated. In a simple obstacle avoidance environment, the time to reduce the expansion of useless nodes is greater than the time to increase the process of parent node re-selection and rewiring, so the average running time is reduced. In complex obstacle avoidance environments, parent node reselecting and rewiring processes increase the time more than the time to reduce the expansion of useless nodes, so the average running time increases. The distance constraint mechanism proposed in this paper enables the algorithm to select and optimize the path with higher efficiency during operation, which makes the node utilization rate increase and the final path length decrease. Especially in the simple obstacle avoidance environment, the search efficiency of the algorithm is improved to a large extent.

4. Conclusion

In this paper, a motion planning algorithm based on RRT^* and $Informed - RRT^*$ with distance constraint mechanism is proposed to improve the problems of low exploration efficiency and poor planning path in the current motion planning tasks of robotic arms under different complexity scenarios. It mainly reduces the expansion of useless nodes when the random tree grows, so that the algorithm can select and optimize the path with

higher efficiency when running. Combined with halton sequence sampling method, the space exploration ability of the algorithm is improved. The simulation results show that $DC - RRT^*$ and $DC - Informed - RRT^*$ can reduce the path length, improve the node utilization rate, and improve the search efficiency of the algorithm to a large extent, especially in simple environment.

References

1. Fang H, Ong S, Nee A. Robot path planning optimization for welding complex joints [J]. INTERNATIONAL JOURNAL OF ADVANC-ED MANUFACTURING TECHNOLOGY, 2017.
2. La Valle S M. Motion planning[J]. IEEE Robotics & Automation Magazine, 2011, 18 (2): 108-118.
3. Liang S. Research on off-line and on-line mo-tion planning methods for multi-degree of freedom manipulator (in Chinese) [D]. Zhejiang University,2023.
4. Ying K C, Pourhejazy P, Cheng C Y, et al. Deep learning-based optimization for motion planning of dual-arm assembly robots[J]. Computers & Industrial Engineering, 2021, 160: 107603.
5. Lavalle S M. Rapidly-Exploring Random Trees: A New Tool for Path Planning[J].Research Report, 1999.
6. Karaman S, Frazzoli E. Incremental Sampling-based Algorithms for Optimal Motion Planning[C]//Robotics: Science and Systems 2010.2010.
7. Gammell J D, Srinivasa S S, Barfoot T D.In-formed RRT*: Optimal Sampling-based Path Planning Focused via Direct Sampling of an Admissible Ellipsoidal Heuristic[J].IEEE, 2014.

Authors Introduction

Mr. Tianbin Meng

He is a postgraduate in Beijing Jiaotong University. His research interests are Robot arm motion planning.

Dr. Jiwu Wang

He is an associate professor, Beijing Jiaotong University. His research interests are Intelligent Robot, Machine Vision, and Image Processing.

CSM-RRT*: an improved RRT* algorithm based on constrained sampling mechanism

Hang Yang
School of Mechanical and Electronic Engineering, Beijing Jiaotong University

Beijing, Haidian District, China

Jiwu Wang
School of Mechanical and Electronic Engineering, Beijing Jiaotong University

Beijing, Haidian District, China

Xueqiang Shang
Aero Engine Corporation of China

Beijing, Changping District, China

E-mail: 21121252@bjtu.edu.cn,jwwang@bjtu.edu.cn,shangxueqiang@163.com

www.bjtu.edu.cn

Abstract

The Rapidly Exploring Random Tree Star (RRT∗) is recognized as a better path planning algorithm, but its path quality and path planning speed still have room for improvement. In this paper, an improved RRT* algorithm(CSM-RRT*) based on constrained sampling mechanism is proposed .The entire planning process is divided into two steps: fast exploration and optimization of the initial path.   Subsequently, a dynamic sampling region consists of removed redundant nodes and collision nodes is formed around initial path By prioritizing exploration within this dynamic region, computational resources can be saved and the asymptotic optimal path can be quickly converged from the initial path. Eventually, simulation results presented in various obstacle environments confirm the efficiency of CSM-RRT*.

Keywords: Path planning, RRT*, Constrained sampling mechanism, Dynamic sampling region

1. Introduction

Rapidly-exploring Random Tree(RRT) is a sampling- based algorithm. Many scholars have conducted research on it. To further improve the speed of the RRT, a double-tree algorithm RRT-Connect [1] makes the target tree and the start tree grow alternately. To improve the path quality, RRT* [2] conducts the process of parent node selection and Rewire for the samples. As the number of nodes increases, the path gradually optimizes, but this process will consume a lot of time.In order to plan the best possible path within the same time, Informed RRT* [3] adopts an elliptical heuristic domain sampling method and continuously shrinks the elliptical region.In order to reach an optimum or near optimum solution at a much faster rate, Smart-RRT* [4] accelerate the rate of convergence. In addition, P-RRT* [5] incorporates the artificial potential field in RRT* to provide transcendental information for the path.Quick-RRT* [6] enlarges the set of possible parent vertices, which generates a better initial solution and converges to the optimal faster than RRT*.PQ-RRT* [7] combines the P-RRT* with the Quick-RRT* to generate better initial solutions.

Further research on path quality and planning speed of the RRT* has specific significance for solving the problem in complex constraints. Based on constrained sampling mechanism ,this paper proposes an improved RRT*algorithm, CSM-RRT*.By comparing with other algorithms,   effectiveness of the proposed algorithm is verified.

2. CSM-RRT*

This section introduces the strategies of the CSM-RRT*, including the following aspects: ChooseParent, Rewire and constrained sampling mechanism.

2.1. *Constrained sampling mechanism*

2.1.1. *Node saved strategy*

The regional exclusion mechanism(Fig.1) was proposed, in which random sampling points were eliminated in a circular area of the node to ensure the sparsity of the entire tree and improves the exploration efficiency.
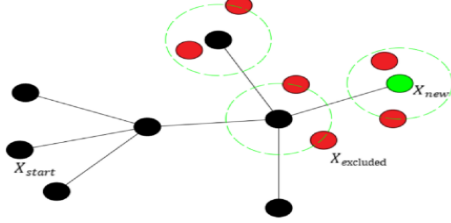
Fig.1 regional exclusion mechanism.

X_{start} represents the initial node, X_{goal} represents the target, $X_{excluded}$ represents the nodes abandoned by the regional exclusion mechanism, and the circular dashed line represents the size of the exclusion region.

However, the x_{rand} was eliminated rather than x_{new}. Additionally, this method only focuses on the sparsity of the entire tree, without considering the sparsity of individual node. To address this issue, regional exclusion mechanism and node saved strategy are combined and applied to the path planning of the manipulator, ensuring the sparsity of both the entire tree and individual sub-node. The results proved to have good effect.

As shown in Fig.2, after adopting the region exclusion mechanism, the parent node O can extend 6 nodes if the exclusion radius r is equal to step size, and the distance between each node is equal to the step size. However, in general, the distance of each node is greater than the step size, when a parent node extends 4 nodes. For a node, we can record the number of expansions and collisions. And the constraint condition is set to 5 times. When the sum of the number of extensions and collisions reaches 5 times, it means that the node has been expanded, or there are a lot of obstacles around it, and it is difficult to continue to expand. In this case, we can save itself to ensure that the next sampling node will not appear in the vicinity again. At the same time, the redundant nodes of the regional exclusion mechanism will also be saved separately for path optimization.

(a)Node extension range   (b) collision detection failed node

Fig.2 Node saved strategy

2.1.2 Dynamic sampling region

As shown in Fig.3, after the initial path planning is completed, each node in the path has an exclusion region, which contains a large number of redundant nodes and collision points. For path optimization, the main goal of path optimization is to obtain an optimal or asymptotic optimal path.In general, an optimal path can only be obtained when the path is close enough to the obstacle.Therefore, redundant and collision nodes close to the initial path should be employed to concentrate on meaningful areas, while discarding nodes with higher path costs.Starting from the node closest to the endpoint, the nodes with the lowest path cost in the dynamic sampling domain are selected to continuously improve path quality. The initial path quickly converges to the optimal path through cyclic iteration.
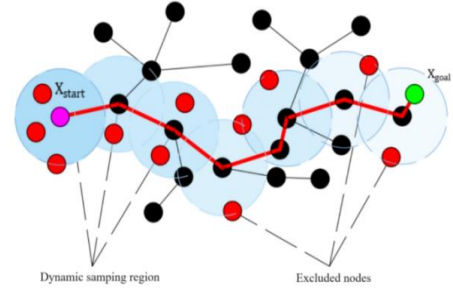
Fig.3 Dynamic sampling region

With further optimization of the path, dynamic region will change and contain many nodes. Nodes in dynamic regions can be included multiple times, resulting in increased runtime. To accelerate exploration, hash table is used. In other words, even if a node is included in the exclusion region many times, its distance to the starting point only needs to be detected once, and the result is saved and used directly later to avoid repeated distance detection.

2.2. *ChooseParent and Rewire*

RRT* sets X_{near} within a certain radius range of the new node and as candidate parent nodes and candidate child nodes during ChooseParent and Rewire.The process is shown in Fig.4.
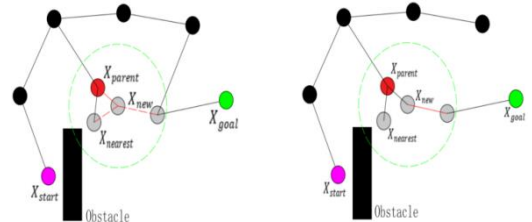
(a)chooseParent   (b)Rewire

Fig.4 Process of ChooseParent and Rewire.

3. Simulation results and Analysis

In order to ensure the reasonability of the statistical results, 50 sets of valid data are gotten for each algorithm when comparing each indicator. The experimental environment is shown in Fig.5.

(a)   (b)
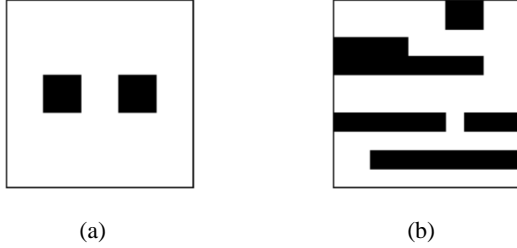
Fig.5 Simulation environment (a)Simple environment (b)Cluttered environment. The entities are obstacles

The results are compared in four directions: initial path planning time T_{init}, the suboptimal path planning time $T_{5\%}$, the initial path cost C_{init} and the suboptimal path cost C_{cost}.

3.1. *Compare of algorithm convergence speeds*

The convergence speed of the algorithm is compared by T_{init} and $T_{5\%}$.

3.1.1 Simple environment

Fig.6 shows the path planning results of the three algorithms in the simple environment, and Table 1 shows the average value of T_{init} and $T_{5\%}$.

(a)   (b)   (c)
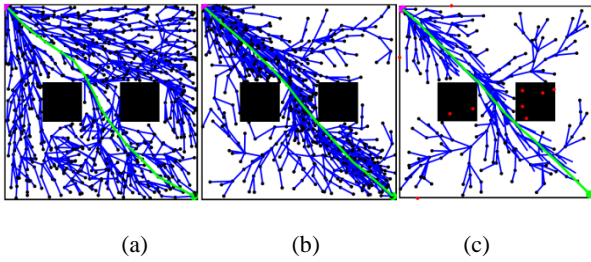
Fig.6 Path planning results in the simple environment (a) RRT*; (b)Informed-RRT*;(c) CMS-RRT*

As shown in Table 1, in terms of $T_{5\%}$,the CMS-RRT * algorithm respectively saved 2.3 seconds and 1 seconds compared to RRT * and Informed-RRT*. The initial path planning speed is 2.5 times faster than the RRT* algorithm. Considering the impact of time saved by the initial path, the CSM-RRT * algorithm did not show significant performance in path optimization in simple environments.

Table1.Average of T_{init} and $T_{5\%}$ of the two algorithms in the simple environment

Algorithm	RRT*	Informed-RRT*	CSM-RRT*
Average(T_{init}/s)	2.93	1.72	0.84
Average($T_{5\%}$/s)	5.46	3.23	1.56

3.1.2 Cluttered environment

Fig.7 shows the path planning results of the three algorithms in the cluttered environment,and Table 2 shows the average value of T_{init} and $T_{5\%}$.

(a)   (b)   (c)
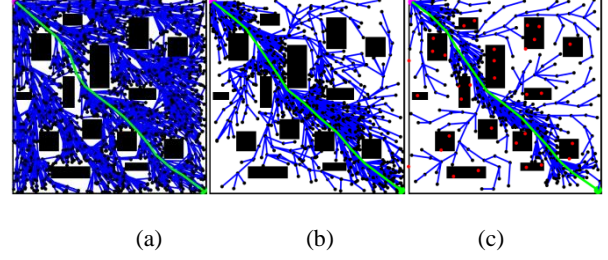
Fig.7 Path planning results in the cluttered environment (a) RRT*; (b)Informed-RRT*; (c) CSM-RRT*

As shown in Fig.7 and Table 2, it can be seen that as the map becomes more complex, the initial path search efficiency of the CSM-RRT * algorithm is not as significant as in a simple environment.However, its path optimization performance is still much better than other algorithms, with an optimization speed increased by 2~5 times.The CSM-RRT* algorithm only samples around the initial path, rather than the entire environment, avoiding path optimization in invalid areas and ensuring fast convergence to the asymptotic optimal path.

Table2.Average of T_{init} and $T_{5\%}$ of the two algorithms in the cluttered environment

Algorithm	RRT*	Informed-RRT*	CSM-RRT*
Average(T_{init}/s)	4.64	2.76	1.58
Average($T_{5\%}$/s)	12.93	5.02	2.58

3.2. *Compare of path quality*

The path quality of the algorithm is compared by C_{init}, C_{cost} .At the same time, in order to avoid similar simulation results, the number of algorithm iterations is set to 2000 for the convenience of viewing algorithm differences.

3.2.1 Simple environment

It can be seen from Table 3 that the path quality of CSM-RRT * is not the beset among all initial paths.This is mainly because the region exclusion mechanism and node preservation strategy in the CSM-RRT * algorithm focus more on fast exploration of the initial path.

However, the final path cost of CSM-RRT * algorithm remains minimal, proving that CSM-RRT * is superior to other algorithms.

Table3.Average of C_{init} and C_{cost} of the two algorithms in the simple environment

Algorithm	RRT*	Informed-RRT*	CSM-RRT*
Average(C_{init})	712.06	707.58	716.84
Average(C_{cost})	703.85	697.59	692.21

3.2.2 Cluttered environment

In order to verify the stability of the algorithm, 50 repeated experiments were conducted on algorithms, and the final path cost simulation results(Fig.8) of the three algorithms in cluttered environments were obtained.
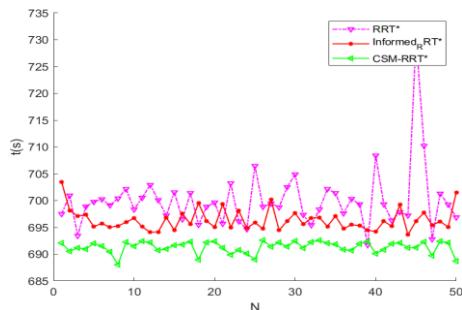
Fig.8 Path cost results in the cluttered environment

As shown in Fig. 8, the final path cost of the CSM-RRT * algorithm is much lower than those of the other algorithms, and the fluctuation between each path result is very small, indicating stable optimization ability. The performance indicates that with a limited iterations, the CSM-RRT* algorithm can quickly converge to the optimal path from initial path every time.

4. Conclusion

This paper proposes an improved RRT* algorithm, CSM-RRT*, which has some advantages in planning speed of initial path, convergence speed and path quality. The main idea of the CSM-RRT* algorithm is to remove and record the redundant nodes and collision points, control the direction of path growth and optimize paths in dynamic sampling region. Compare RRT* and Informed-RRT*,it is verified that the CSM-RRT* algorithm has good applicability to both simple and cluttered environments.

References

1. Kuffner JJ, LaValle SM. RRT-connect: an efficient approach to single-query path planning. Proceedings of IEEE International Conference on Robotics and Automation. 2000 Apr 24–28; San Francisco, USA. pp. 1–7.

2. Perez AT, Karaman S, Shkolnik AC, et al. Asymptotically optimal path planning for manipulation using incremental sampling-based algorithms. IEEE/RSJ International Conference on Intelligent Robots & Systems. 2011 Sep 25-30; San Francisco, USA; pp. 4307–4313.

3. GAMMELL JD, SRINIVASA SS, BARFOOT TD, et al. Informed RRT*: Informed RRT*: Optimal Sampling-based Path Planning Focused via Direct Sampling of an Admissible Ellipsoidal Heuristic[C]//2014 IEEE/RSJ International Conference on Intelligent Robots and Systems. Chicago, USA: IEEE,2014, pp.2997-3004

4. Islam, F., Nasir, J., Malik, U., Ayaz, Y., & Hasan, O. . Rrt*-smart: Rapid convergence implementation of rrt* towards optimal solution. In 2012 IEEE International Conference on Mechatronics and Automation ,2012, pp. 1651–1656.

5. Qureshi AH, Ayaz Y. Potential functions based sampling heuristic for optimal path planning. Auton Robot. 2015, vol.40(6), pp.1079–1093.

6. Jeong IB, Lee SJ, Kim JH. Quick-RRT*: triangular inequality-based implementation of RRT*with improved initial solution and convergence rate. Expert Syst Appl. 2019, vol.123, pp.82–90.

7. Li Y, Wei W, Gao Y, et al. PQ-RRT*: an improved path planning algorithm for mobile robots. Expert Syst Appl.2020,vol.152,pp.113425–113436.

Authors Introduction

Mr. Hang Yang

He is currently studying for a master's degree in mechanical engineering at Beijing Jiaotong University. He is mainly engaged in research on robot path planning.

Dr. Jiwu Wang
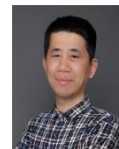
He is an associate professor, at Beijing Jiaotong University. His research interests are Intelligent Robots, Machine Vision, and Image Processing.

Mr. Xueqiang Shang

He is a senior engineer at Aero Engine Corporation of China, primarily engaged in hydraulic component design and verification.

Small Sample Object Detection Based on Improved YOLOv5

Yuxuan Gao
School of Mechanical and Electronic Engineering, Beijing Jiaotong University
Beijing, Haidian District, China
Jiwu Wang
School of Mechanical and Electronic Engineering, Beijing Jiaotong University
Beijing, Haidian District, China
Zixin Li
Aero Engine Corporation of China,Beijing, Changping District, China
Email: 18815511536@163.com, jwwang@bjtu.edu.cn,15311426793@163.com
www.bjtu.edu.cn

Abstract

Object detection is widely used in various production and life, such as mask detection and recognition during the epidemic, face recognition with masks. Object detection algorithm based on deep learning has always been an important research content and implementation method in the field of object detection. Due to the large number of lead seals and fuses, their locations are not fixed, the lead seals and fuses have difficulties such as few sample datasets, complex target background and easy to be blocked, and strong reflective interference, and the conventional image processing methods are difficult to solve the problem of effective object recognition. In this study, by expanding the datasets, using different data enhancement methods, and training in the improved algorithm, the detection accuracy, detection speed, and adaptability were effectively improved.

Keywords: YOLOv5; Deep learning; Object detection; Few-shot learning

1. Introduction

Visual inspection, as a key application in related fields, provides fundamental support for achieving automation and is a key link in promoting the development of industrial automation [1]. Feature based visual detection methods have been widely used in industrial inspection, but this traditional detection method relies too much on manually designed feature extractors based on experience, and when faced with changes in target morphology and interference, the detection accuracy will significantly decrease. With the development of machine vision and artificial intelligence technologies, deep learning has become a new paradigm for object detection tasks in the industrial field [2].

However, deep learning relies on a large amount of labeled sample data during the training phase, which has drawbacks in real application scenarios. In factories, it is often difficult to obtain sufficient sample data due to issues such as a single model [3]. In such cases, the limited number of labeled samples can lead to severe overfitting of deep learning models and a decrease in recognition accuracy. Moreover, with sufficient datasets, labeling data can be time-consuming and labor-intensive, leading to issues of omission and mislabeling during labeling.

In this study, we used different image processing methods to further expand the original dataset and perform other data augmentations. At the same time, the automatic annotation of the dataset is completed through code to ensure the accuracy of the annotation, and it is trained in the improved YOLOv5 network to improve detection accuracy.

2. Methodology
2.1. Yolov5 algorithm

The core idea of YOLOv5 algorithm is to treat the object detection task as a regression problem and segment the feature map into L×L sized cells, M bounding boxes are set for each cell to surround the target. Each cell is responsible for detecting targets that fall within its area. In one operation, the bounding boxes, localization confidence, and probability of each category of targets contained in all cells are predicted. Then, the detected bounding boxes are filtered using NMS (Non Maximum Suppression) method to obtain the position, category, and confidence of the target to be tested.

The implementation of YOLOv5 algorithm can be summarized as the following steps:

(1) Preprocessing. The input image is first preprocessed, including scaling, cropping, and normalization operations, to convert it into a format suitable for network input.

(2) Feature extraction. The input image is fed into the backbone network and the feature is extracted through the backbone network.

(3) Feature pyramid network. YOLOv5 utilizes Feature Pyramid Networks (FPN) [4] to fuse feature maps of different scales, capture targets of different sizes, and improve detection accuracy.

(4) Predictive header. YOLOv5 uses a prediction head to predict each cell, obtaining bounding boxes (including the position and size of the target), confidence (probability of target existence), and class probabilities.

(5) Decoding prediction. Decode and convert the output of the prediction head into actual image coordinates, including adjusting the predicted bounding box to the original image size, and calculating the category and confidence score of each target.

(6) Non maximum suppression. Use NMS algorithm to remove redundant detection boxes with high overlap, and finally retain the most likely object detection result.

(7) Post processing. Perform post-processing on the detection results processed by NMS, including visualizing and saving the results.

The YOLOv5-7.0 network model structure is shown in Fig. 1, which mainly includes Conv, CSPDarkNet, and SPPF modules.
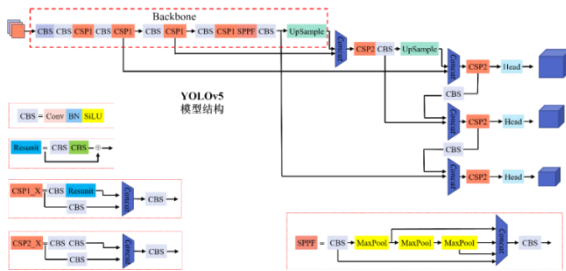
Fig.1 YOLOv5-7.0 model structure

2.2. SENetAttention

In order to obtain attention in the channel dimension, the input characteristic graph is convoluted to obtain the dimension H × W × C, and then compress the input characteristic graph through global average pooling. This step converts the characteristic graph of each channel into the scalar of $1 \times 1 \times C$, thus forming a channel description vector, which captures the global spatial information of each channel and corresponds to the initial weight of each channel. This step is to compress (squeeze),as shown in equations (1) [5]:

$$Z_c = F_{sq}(x_c) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} x_c(i,j) \qquad (1)$$

Where, *H* and *W* represent the height and width of the feature map, and x_c represents the input at each channel.

Then the second step is the excitation operation, which uses the small network of the Fully Connected (FC) layer to learn the relationship between channels. This network usually contains two full connected layers. The first full connected layer is responsible for mapping the channel description vector to a low dimensional space, with a

ReLU activation function in the middle, and the second full connected layer maps it back to the original number of channels, This process can be regarded as a bottleneck structure, which aims to extract the dependencies between channels. The mathematical expression of this process is shown in Formula (2):

$$s = F_{ex}(z,W) = \sigma(g(z,W)) = \sigma(W_2\sigma(W_1z)) \qquad (2)$$

Where, σ Represents the ReLU function, W_1 represents the parameters of the dimension reduction layer, W_2 represents the parameters of the dimension increase layer, and *z* is the output result from the previous compression step.

The last step is Re-scale, which multiplies the previously learned channel weight vector with the original feature map to achieve the function of adaptively adjusting the feature response of each channel and improve the performance of the network model. The network architecture of its algorithm module is shown in Fig. 2.
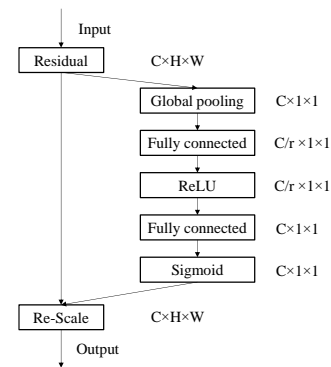
Fig.2 The diagram of the squeeze-and-excitation block

Adding the SENet attention module to the last layer of the Backbone network can effectively reduce the false detection rate and missing detection rate of lead seals and fuses, and improve the detection accuracy.

3. Data processing

3.1. image processing

For the image acquisition work of the product object, considering the complex surface structure of the product and the obvious reflective effect of the metal material and paint surface, the shooting effect is easily affected by the lighting conditions. According to this feature, the product is placed in a variety of lighting conditions, and different types of camera equipment are used to collect images. At the same time, local images and unmarked background images are taken in each lighting environment, in order to maximize the diversity of target images.Sample is shown in Fig. 3.

a) Low brightness environment   b) Normal lighting environment   c) High brightness environment
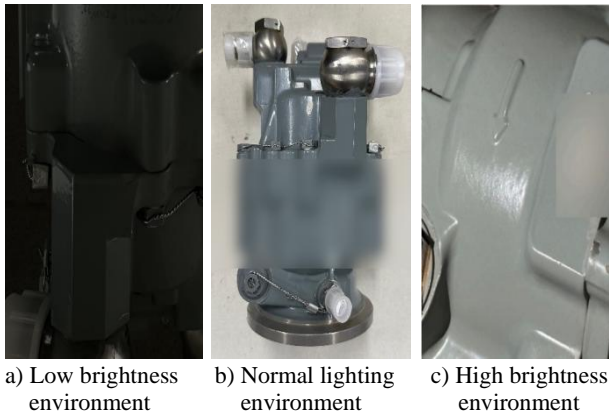
Fig.3 Shooting in various lighting conditions

ImgAug is a powerful and flexible image enhancement library based on Python language, which is mainly used in the field of deep learning and computer vision. It supports a variety of image processing libraries such as OpenCV. By calling ImgAug's API, the image can be transformed in various aspects such as geometry, color, contrast, etc., so as to expand the original image data set and improve the generalization ability and robustness of the model. With the help of its official API document, the corresponding Python code is written, and the image is transformed as follows. The process can be shown in Fig. 4.
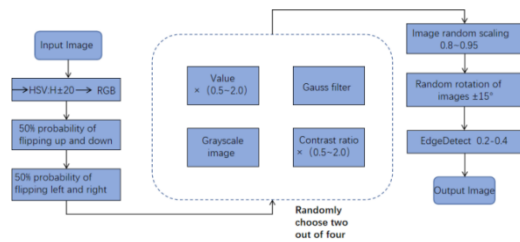
Fig.4 Image enhancement process

Write python code to execute the above image enhancement operation, and realize the automatic annotation function of new images while realizing various transformations of images. By reading the XML annotation file corresponding to the original image, the position and coordinate information of all the annotation boxes are extracted, and according to the random enhancement process of each new image, the code automatically calculates and adjusts the information of the annotation box, thus generating a new XML file that is completely aligned with the target position in the new image. This method avoids the tedious work of manual marking and saves a lot of human resources. By automatically updating the annotation information, the enhanced image is consistent with the original annotation, thus improving the accuracy and reliability of the model. The new sample image of the original image after image enhancement is shown in Fig. 5.

a) Original image   b) New samples 1

c) New samples 2   d) New samples 3
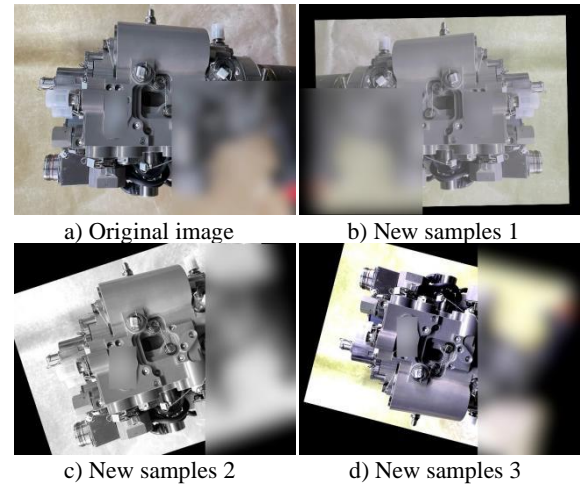
Fig.5 New samples after image augmentation

Through the above image enhancement method, each initial image is expanded to 6 new images, and finally 441 images are obtained as the final data set.

4. Experiment and Comparison

After completing the dataset production and modifying the YOLOv5 model, conduct training and compare the training results of the unimproved model. At the same time, verify the accuracy of the same image.

Baseline is the yolov5 model without adding SE attention mechanism, and the training is carried out using the non-expanded dataset. The dataset is as shown in the figure. SENet is the yolov5 model with adding SE attention mechanism, and the training is carried out using the non-expanded dataset. SENet+dataset is the yolov5 model with adding SE attention mechanism, and the training is carried out using the expanded dataset. The unexpanded dataset has 63 pictures, and the expanded dataset has 441 pictures, as shown in the following Fig. 6.

a) Baseline   b) Expand the dataset
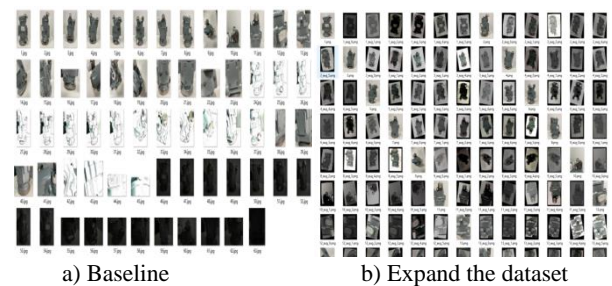
Fig.6 Image expansion

The key data of the training results are shown in Table 1. Among them, mAP (mean Average Precision) refers to the average precision, which comprehensively considers precision and recall to measure the performance of the model on detecting targets.

Yuxuan Gao, Jiwu Wang, Zixin Li

Table1. Comparison of data under different training conditions

	mAP (%)	precision (%)	time (ms)
Baseline	81.4	81.4	9.9
SENet	87.5	89.5	9.3
SENet+ dataset	96.3	96.9	12.3

From the table, it can be seen that compared to the baseline reference benchmark, the use of SENet attention mechanism in mAP has improved by 6.1%, and the combination of SENet and expanded dataset has improved by 8.2%; In terms of accuracy, compared to Baseline, the use of SENet attention mechanism improved by 7.9%, and the combination of SENet and expanded dataset improved by 7.4%.

The above data results demonstrate the advantages of SENet+augmented dataset. The best weight files trained using Baseline, SENet, and SENet+augmented dataset are tested on the images in the test set. The image detection results are shown in Fig. 7.

a) Baseline   b) SENet

c) NSENet+ Expand the dataset
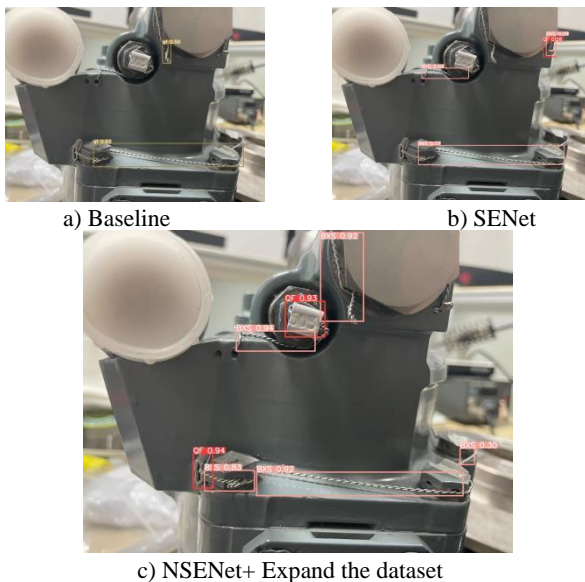
Fig.7 Comparison of detection under different training conditions

From the figure, it can be seen that the effect is significantly improved after adding the SENet module network. However, due to the limited amount of data, there are still issues of false positives and missed alarms. After expanding the dataset, it can be clearly observed that not only is the accuracy further improved, but also false positives are reduced. In summary, the addition of SENet attention mechanism and the scalability of the dataset are helpful for model training and improving detection accuracy.

5. Conclusion

In this study, by improving the YOLOv5 model and expanding the dataset, the training accuracy is improved. The experimental results show that this method can meet the measurement accuracy requirements. Compared with the unimproved method, the method proposed in this paper improves detection accuracy and accuracy. This method can provide good support for deep learning based object detection research. However, due to the increased computational complexity of the algorithm, the running time is longer. In order to meet the real-time requirements of object detection, the speed of the algorithm needs to be further improved.

References

1. Fu Binbin Application and development trend of industrial machine vision [J] China Industry and Informatization, 2021 (11): 18-24
2. Guo Jing, Luo Hua, Zhang Tao Machine Vision and Application [J] Electronic Science and Technology, 2014, 27 (07): 185-188
3. Xing Z, Chen X. Lightweight algorithm of insulator identification applicable to electric power engineering[J]. Energy Reports, 2022, 8: 353-362.
4. Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 2117-2125.
5. Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 7132-7141.

Authors Introduction

Dr. Jiwu Wang
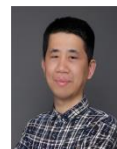
He is an associate professor, Beijing Jiaotong University. His research interests are Intelligent Robot, Machine Vision, and Image Processing.

Mr. Yuxuan Gao

He is a postgraduate in Beijing Jiaotong University. His research interests are deep learning and yolo.

Ms. Zixin Li

She is an engineer at Aero Engine Corporation of China. Her research direction is 3D reconstruction

Research on Gas Pore Prediction Method Based on Sand Core Characteristic Time

Xiaolong Wang
School of Mechanical and Electronic Engineering, Beijing Jiaotong University
Beijing, Haidian District, China

Qihua Wu
Weichai Power Co., Ltd.
Weifang, China

Jiwu Wang
School of Mechanical and Electronic Engineering, Beijing Jiaotong University
Beijing, Haidian District, China

Jinwu Kang
School of Materials Science and Engineering, Tsinghua University
Beijing, Haidian District, China

Na Li
Weichai Power Co., Ltd.
Weifang, China

Yucheng Sun
Weichai Power Co., Ltd.
Weifang, China
E-mail: 21121249@bjtu.edu.cn, wuqh@weichai.com, jwwang@bjtu.edu.cn, kangjw@tsinghua.edu.cn,
lin@weichai.com, sunyuc@weichai.com

Abstract

In the production of castings, gas pores are a prevalent defect, particularly in components where air-tightness is crucial, such as cylinder blocks and heads. These defects primarily arise from the intrusion of gases into the casting during the combustion of resin in the sand core. Due to the complexity of the sand core's gas evolution and venting process, predicting gas pores using numerical simulation poses significant computational and time challenges. This paper introduces a rapid prediction method for gas pores based on the characteristic time of the sand core. By setting the heat transfer boundaries, initial conditions, and termination criteria for computation, the thermal conductivity of the sand core is adjusted. The termination computation time is used as the characteristic time of the sand core. This time is then compared with a critical characteristic time to predict the distribution of gas pores. As the analysis of more sand cores is incorporated, the precision of the critical characteristic time improves, leading to more accurate predictions of gas pores in castings.

Keywords: Gas pore; Numerical simulation; Iron casting; Sand core; Resin

1. Introduction

Gas pores are common and typical defects in cast iron and aluminum alloy castings, with entrapped gas pores being a major type of pore defect. In complex castings such as engine blocks and cylinder heads, numerous intricately shaped sand cores are employed. During the filling and solidification of molten metal, the sand cores are subjected to heating, causing the organic binder within them to decompose or burn at high temperatures, generating gas [1], [2]. If the gas cannot be expelled from the core and instead infiltrates the molten metal, it results in entrapped gas pores [3]. The factors influencing the occurrence of entrapped gas pores are numerous, making

Xiaolong Wang, Qihua Wu, Jiwu Wang, Jinwu Kang, Na Li, Yucheng Sun

it challenging to eliminate. This issue significantly impacts product quality, sometimes leading to the rejection of castings or the discovery of pore defects after machining. The complex formation mechanism of entrapped gas pores makes it difficult to predict using numerical simulation methods accurately [4], [5], [6].

This paper adopts a strategy for identifying hazardous sand cores and their critical points. Based on the principles of similarity in heat transfer within the sand core and the transport of gas within it, a virtual heat transfer method for sand cores is employed. The paper proposes a characteristic time indicator for predicting the occurrence of casting porosity induced by sand cores. This indicator is used to assess the critical points and severity of risks associated with sand cores, enabling the determination of whether casting porosity will be induced and predicting the location of gas pore formation. The characteristic time values are stored in a database for validation against actual test results. Furthermore, these values serve as a reference for predicting porosity in new sand cores.

2. Methods

2.1. *Heat transfer analysis and modified thermal property parameters*

Gases produced from the high-temperature decomposition of binders are expelled through exhaust channels. However, due to the intricate geometries of sand cores and variations in the distances from different points to the core head, some points, especially those farthest from the core head, may pose challenges in gas evacuation. This difficulty increases the probability of porosity formation. The term "distance" here is not spatial but pertains to the movement of gas within the sand core, exhibiting a certain similarity principle with heat transfer within the sand core. To identify the point farthest from the sand core, i.e., the most challenging gas evacuation point, a virtual heat transfer method is employed. The numerical simulation for virtual heat transfer focuses solely on the specific sand core in question, with other sand cores, moldings, and castings excluded from the simulation. These excluded components are, however, crucial for automatically determining the core head. This determination is achieved by identifying areas where the sand core does not contact the casting, signifying the presence of the core head.

Heat transfer analysis is the foundation of the resin burning, gas release and flow, the melt solidification, and the formation of gas pores. The heat transfer inside the casting, mold, and sand core during the casting process is expressed by the Fourier equation,

$$\rho c \frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(k \frac{\partial T}{\partial x}\right) + \frac{\partial}{\partial y}\left(k \frac{\partial T}{\partial y}\right) + \frac{\partial}{\partial z}\left(k \frac{\partial T}{\partial z}\right) + q \quad (1)$$

where ρ is the density, c is the specific heat capacity, T is the temperature, t is the time, k is the thermal conductivity, and q is the latent heat term as the melt solidifies.

The nominal thermal conductivity of the sand core is adjusted based on the permeability of the sand core, and the adjustment formula is as follows:

$$\lambda_1 = \frac{K}{100}\lambda_0 \quad (2)$$

where λ_0 is the real thermal conductivity of the sand core (W/(m·K)), λ_1 is the nominal thermal conductivity of the sand core (W/(m·K)), K is the permeability of the sand core.

The nominal specific heat of the sand core is adjusted based on the gas generation of the sand core, and the adjustment formula is as follows:

$$c_1 = \frac{G}{20}c_0 \quad (3)$$

where c_0 is the real specific heat of the sand core (J/(kg·K)), c_1 is the nominal specific heat of the sand core (J/(kg·K)), G is the gas generation of the sand core.

2.2. *Determination of dangerous gas penetration points in sand cores*

During the simulation process, six distinct boundary interfaces are evident, as depicted in Fig. 1: metal-sand core, metal-air, metal-sand mold, sand core-sand mold, sand mold-air, and sand core-air. The aforementioned coating is applied to both the surface of the sand core and within the sand mold cavity. Consequently, the interfaces involving metal-sand core, metal-sand mold, and sand mold-sand core are influenced by the presence of this coating. Its inclusion results in heightened thermal and penetration resistance at these boundaries. Moreover, owing to the coating's exceptionally low gas permeability, certain segments exhibit zero gas permeability at room temperature, signifying an incapacity to permit gas passage.
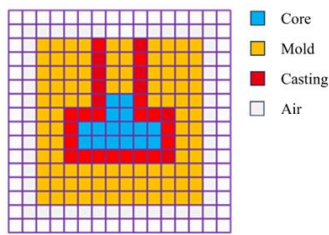
Fig. 1 Boundary setting conditions.

Establishing the boundary conditions for the sand core involves setting the surface temperature of the core head to a constant 1500℃. The remaining exterior surfaces of the sand core are considered adiabatic. The initial temperature of the sand core's elements is set at 25℃. The virtual heat transfer computation is programmed to automatically terminate when the volume of the unheated portion of the sand core reaches 100 mm³ , as illustrated in Fig. 2. This moment is recorded as the characteristic time of the sand core.

(a) Before heat transfer

(b) After heat transfer
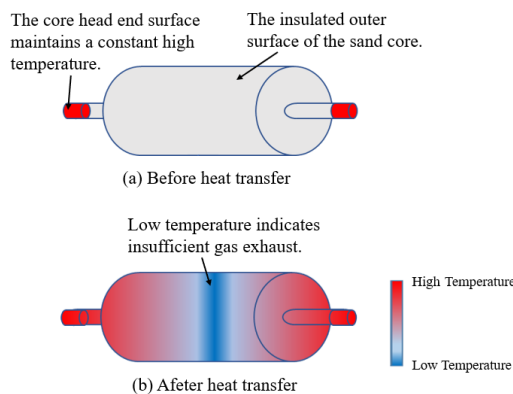
Fig. 2 Virtual heat transfer of sand core identifies the most challenging gas exhaust point.

When the characteristic time of the sand core exceeds the critical time for gas pore initiation, it is determined that the sand core induces gas pores in the casting. Conversely, when the characteristic time of the sand core is less than the critical time for gas pore initiation, it is determined that the sand core does not induce gas pores in the casting. The topmost units in the unheated portion of the sand core are selected, and these units are identified as the critical points for gas pore initiation in the casting. These critical points represent the locations in the sand core most susceptible to inducing gas pores in the casting.

Save the characteristic time of the sand core to the sand core characteristic time database, and adjust the critical characteristic time based on the actual occurrence of gas pores for each sand core.

3. Pore prediction results

3.1. *Parameters*

Predictions of pores were performed on a single cylinder head. The casting consists of the intake manifold, exhaust manifold, upper water jacket, lower water jacket, and chassis core.

Due to the relatively simple geometry of the intake and exhaust manifold cores, only the cores themselves were considered in the determination of dangerous areas during the assessment of the intake and exhaust cores. For virtual heat transfer computations, the thermal property parameters of the cores were set with a density of 1.6×10^3 kg/m³, a nominal specific heat of 1×10^3 J/(kg·K), and a nominal thermal conductivity of 0.5 W/(m·K).

3.2. *Gas pore*

The virtual heat transfer simulation was conducted for the casting and sand cores, and the computation was terminated upon meeting the specified conditions. The characteristic time for the sand cores was determined to be 2.1 s. The critical units of each sand core, as illustrated in Fig. 3, align with the experimental results, where the critical characteristic time for gas pore generation in the sand cores was identified as 1 second. Consequently, the casting exhibits gas pores, and the locations of these pores are depicted in Fig. 4.

(a) Intake and exhaust manifold cores

(b) Upper and lower water jacket cores

(c) Chassis core
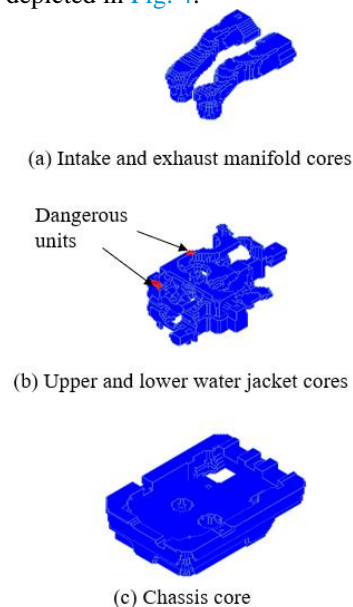
Fig. 3 Dangerous units in each sand Core

Xiaolong Wang, Qihua Wu, Jiwu Wang, Jinwu Kang, Na Li, Yucheng Sun
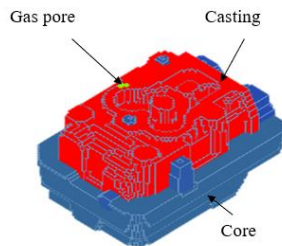
Fig. 4 Casting pore results

Comparison with castings produced in the actual factory workshop reveals a consistent distribution of gas pores, validating the effectiveness of the proposed method.

4. Conclusion

This paper proposes a gas pore prediction method based on the characteristic time of sand cores, enabling the forecasting of pores in castings. It addresses the challenges associated with the complex gas generation and exhaust processes in sand cores, difficulties in predicting them through numerical simulation techniques, and the significant costs and time associated with such simulations. This method provides guidance for practical casting production. As the analysis of sand cores increases, the critical characteristic time becomes more precise, leading to more accurate gas pore predictions.

References

1. Daronde, S. et al. The Effect of Vacuum on the Mechanical Properties of Sand Cast AA6061 Alloy. J. Mater. Eng. Perform. 31, 262–271 (2022).

2. Wang, J., Jiang, H. & Jiang, N. Study on the pyrolysis of phenol-formaldehyde (PF) resin and modified PF resin. Thermochimica Acta 496, 136–142 (2009).

3. Wang, X. et al. Study on the Gas Release of 3D-Printed Furan Resin Sand Core during the Casting Process. Materials 16, 4152 (2023).

4. Mocek, J., Kaźnica, N. & Zych, J. Kinetics of Gases Emission from Surface Layers of Sand Moulds. Archives of Foundry Engineering; 2018; vol.18; No 1 (2018).

5. Carlson, K. D., Lin, Z. & Beckermann, C. Modeling the effect of finite-rate hydrogen diffusion on porosity formation in aluminum alloys. Metall. Mater. Trans. B-Proc. Metall. Mater. Proc. Sci. 38, 541–555 (2007).

6. Qian, X. et al. Gas evolution characteristics of three kinds of no-bake resin-bonded sands for foundry in production. China Foundry 19, 140–148 (2022).

Authors Introduction

Mr. Xiaolong Wang

He is currently studying for a master's degree in mechanical engineering at Beijing Jiaotong University. He is mainly engaged in research on intelligent casting and porosity prediction.

Mr. Qihua Wu

He is a process engineer at Weichai Power Co., Ltd., primarily engaged in research on intelligent manufacturing.

Dr. Jiwu Wang

He is an associate professor, at Beijing Jiaotong University. His research interests are Intelligent Robots, Machine Vision, and Image Processing.

Dr. Jinwu Kang

He is an associate researcher and doctoral supervisor at Tsinghua University, specializing in the fields of intelligent casting and simulation modeling.

Mrs. Na Li

She currently works at Weichai Power Co., Ltd., primarily responsible for the intelligent manufacturing sector.

Mr. Yucheng Sun

He currently works at Weichai Power Co., Ltd., primarily responsible for the intelligent manufacturing sector.

Optimization Analysis of a Deep Learning-Based Model for Predicting Temperature Fields in the Solidification Process of Castings

Yahui Yang

School of Mechanical and Electronic Engineering, Beijing Jiaotong University

Beijing, Haidian District, China

Jiwu Wang

School of Mechanical and Electronic Engineering, Beijing Jiaotong University

Beijing, Haidian District, China

Jinwu Kang

School of Materials Science and Engineering, Tsinghua University

Beijing, Haidian District, China
E-mail: 22126065@bjtu.edu.cn,jwwang@bjtu.edu.cn,kangjw@tsinghua.edu.cn

Abstract

In the foundry industry, accurate prediction of the temperature field of castings during solidification is crucial to ensure high product quality and improve productivity. This study provides an in-depth investigation of the Unet deep learning model used for temperature field prediction during solidification, focusing on the impact of key training parameters such as optimiser selection, batch size, number of iterations, and loss function selection on the prediction performance of the Unet model. The results of the study demonstrate the analysis of selecting and tuning these parameters to improve the accuracy and reliability of model predictions. The results of this study not only provide new ideas for the practice of deep learning applications in the foundry industry, but also help to improve the accuracy and efficiency of the production proces.

Keywords: casting, deep learning, temperature field prediction, parameter optimization.

1. Introduction

In recent years, deep learning has expanded in field prediction, simplifying traditional numerical simulations [1], [2] while maintaining accuracy. Specifically, in casting solidification, data-based methods predict temperature fields [3] effectively. However, incorrect network structures or training parameters can cause underfitting or overfitting, leading to errors. Thus, optimizing these parameters is vital for enhancing accuracy, reducing costs, and speeding up design iterations for precise temperature field prediction in casting processes.

2. Model Introduction

Selecting an appropriate network model structure is crucial for high-precision prediction of the temperature field during the solidification process of castings. The U-Net [4], [5] network demonstrates exceptional performance in the field of image segmentation, particularly in achieving precise segmentation with smaller datasets.
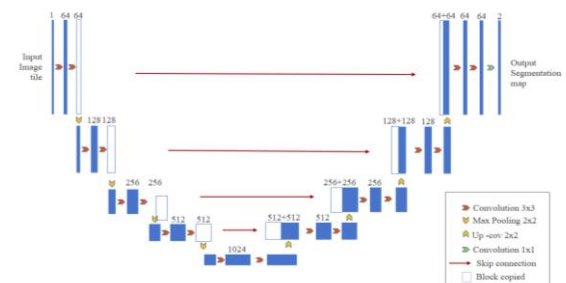
Fig.1. Unet network model [6]

In the prediction of temperature field during casting solidification, due to the time-consuming computation of

the finite difference method, it is difficult to generate a large amount of training data, and the encoder-decoder network architecture is able to efficiently extract the semantic features of the image and retain the positional information, which is suitable for prediction tasks with small datasets. The Unet network model is shown in Fig. 1.

In this study, we employ a network model (Fig. 1) with the classic U-Net structure to predict the temperature field during the solidification process of castings. The model takes the mold images of sand castings as input, and the output is the temperature field image after a certain time step. Both input and output are single-channel images.

3. Optimization Methods

In the realm of deep learning network architectures, the addition or alteration of various modules can significantly impact the final predictive performance. To

achieve optimal neural network prediction capabilities, this study delves into several key factors affecting model training, including the choice of optimizer, batch size, number of iterations, and the selection of loss functions. The following sections will present a detailed comparative analysis of these crucial elements, aiming to elucidate how to optimize deep learning network structures for higher accuracy in predictions. This translation is intended for application in academic papers.

3.1. *The Impact of Different Optimizers on the Prediction Accuracy of the Temperature Field*

In this set of comparative experiments, the batch size was set to 32, and the number of epochs was fixed at 200. Mean Absolute Error(MAE) was chosen as the loss function. To analyze the impact of different optimizers, three distinct optimizers were selected for comparison: Adam, Nadam and SGD. The learning rate for each was adjusted using a cosine annealing learning rate strategy. The experimental results are presented in Table 1.

Table 1. Comparison of results of different optimizers

Serial Number	Optimizer	Accuracy	Pearson Correlation Coefficient	Loss	Val-Loss	Training Time
1	Adam	0.9156	0.9989	0.0096	0.0230	66727.5657
2	SGD	0.8198	0.7685	0.1977	0.1769	68961.8629
3	Nadam	0.9118	0.9989	0.0099	0.0236	65335.4521
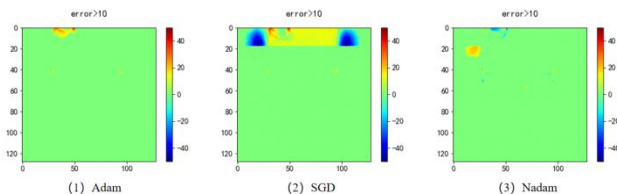
Fig.2. Error Plots of Predicted versus Actual Values Using Three Optimizers

According to the data presented in Table 1, the model demonstrates higher accuracy and correlation when employing the Adam optimizer, with the smallest loss value observed. To further analyze and compare these results, samples at a specific moment (t=10 minutes) were

randomly selected for visual analysis. Fig.2 demonstrates the distribution of errors between the temperature fields predicted by the model and the actual temperature fields, under three different optimizers: Adam, Nadam, and SGD.

3.2. *The Impact of Different Batch Sizes on Prediction Accuracy*

In this set of comparative experiments, building upon the results of the previous set, Adam was selected as the optimizer, with the number of epochs set to 200 and MAE employed as the loss function. To investigate the impact of batch size on model performance, three different batch sizes (16, 32, and 64) were established for comparative analysis. The experimental results are shown in Table 2.

Table 2. Comparison of results for different batch sizes (batch size)

Serial Number	Optimizer	Accuracy	Pearson Correlation Coefficient	Loss	Val-Loss	Training Time
1	16	0.9014	0.9889	0.0181	0.0217	68537.3182
2	32	0.9118	0.9989	0.0096	0.0230	66727.5657
3	64	0.8089	0.9979	0.0214	0.0286	66357.1713

Based on the data analysis from Table 2, it is observed that setting the batch size to 32 results in higher model accuracy and correlation, with the smallest loss value and shorter training time. Samples at the same moment (t=10 minutes) were selected for visual analysis.
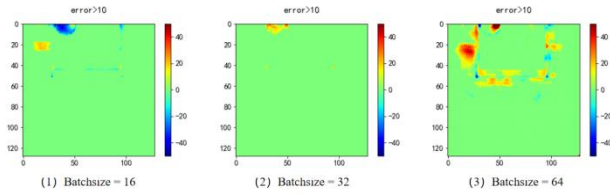
Fig.3. Error Plots of Predicted versus Actual Values for Three Batch Sizes

Fig. 3 illustrates the distribution of errors between the predictedand actual values of the model under three different batch sizes (16, 32, and 64). The model exhibits relatively smaller prediction errors when the batch size is set to 32.

3.3. *The Impact of Different Iteration Counts on Prediction Accuracy*

In this set of comparative experiments, building upon the results of the first two sets, Adam was chosen as the optimizer, with the batch size set to 32 and Mean Absolute

Error (MAE) used as the loss function. To explore the impact of different numbers of epochs on model performance, three distinct numbers of epochs (50, 200, and 500) were established for comparative analysis. The experimental results are shown in Table 3.

According to the experimental data in Table 3, the model exhibits the highest accuracy and correlation, with the smallest loss value and relatively short training time, when the number of epochs is set to 200. With an increase in the number of epochs, the computational cost correspondingly rises. Therefore, when selecting the optimal parameters for the model, it is crucial to consider both prediction accuracy and computational cost as key factors in determining the most suitable parameter settings. Fig. 4 displays error distribution of the temperature field at different numbers of epochs (50, 200, 500), providing a more intuitive assessment of model performance.
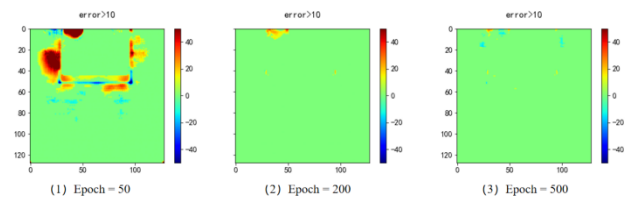
Fig.4. Error Plots of Predicted versus Actual Values for Three Different Epoch Counts

Table 3. Comparison of results for different number of iterations (epoch)

Serial Number	Epoch	Accuracy	Pearson Correlation Coefficient	Loss	Val-Loss	Training Time
1	50	0.7264	0.8963	0.0302	0.0361	17488.4049
2	200	0.9118	0.9989	0.0096	0.0230	66727.5657
3	500	0.9083	0.9988	5.3330e-04	0.0040	157621.8659

3.4. *The Impact of Different Loss Functions on Prediction Accuracy*

In this set of comparative experiments, based on the results of the previous three sets, we selected the Adam optimizer with the number of epochs set to 200 (epoch = 200). Three different loss functions were chosen for comparative analysis: Mean Squared Error (MSE), Mean Absolute Error (MAE), and their combination (MSE + MAE). The experimental results are presented in Table 4.
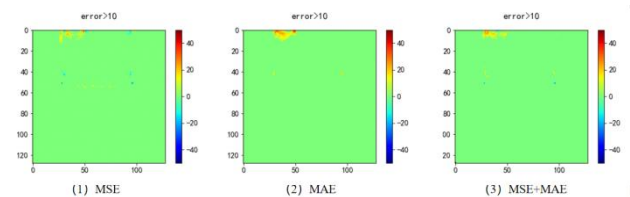
Fig.5. Error Plots of Predicted versus Actual Values for Three Different Loss Functions

Based on the experimental data presented in Table 4, the highest accuracy and correlation, along with the smallest

Table 4. Comparison of results of different loss functions

Serial Number	Loss Function	Accuracy	Pearson Correlation Coefficient	Loss	Val-Loss	Training Time
1	MSE	0.9293	0.9990	4.9873e-04	0.0038	64514.8804
2	MAE	0.9118	0.9989	0.0096	0.0230	66727.5657
3	MAE+MSE	0.9083	0.9988	5.3330e-04	0.0040	67621.8659

Yahui Yang, Jiwu Wang, Jinwu Kang

loss value and relativly short training time, are observed when Mean Squared Error (MSE) is used as the loss function. Thus, choosing MSE as the loss function is the optimal parameter choice for the model. Fig. 5 illustrates the visualization of the error distribution of the temperature field predictions when different loss functions (MSE, MAE, and MSE+MAE) are employed.

3.5. *Experimental Results*

In the four comparative experiments conducted, we determined that the Adam optimizer is the optimal choice, surpassing Nadam and SGD in accuracy, correlation, and loss minimization. The model exhibits its best performance with a batch size of 32, effectively balancing efficiency and accuracy. Setting the number of epochs to 200 achieves a balance between computational cost and precision. Among the loss functions compared, Mean Squared Error (MSE) significantly excelled in predictive accuracy and minimizing loss.

4. Conclusion

In this study, we employed a data-driven Unet neural network model for temperature field prediction, optimizing its structure and training parameters (optimizer, batch size, epochs, loss function). The model, trained on 128x128 pixel image data, showed optimal performance with the Adam optimizer, a batch size of 32, 200 epochs, and Mean Squared Error (MSE) as the loss function, achieving an accuracy of 0.9293 and a correlation of 0.9990. These findings offer new insights and valuable references for the application of deep learning in the casting industry.

Reference

1. Lee S M, Lee W J. Finite-element analysis on thermornechanical behavior of a marine propeller casting in the sand-casting process[J]. JOURNAL OF MATERIALS ENGINEERING AND PERFORMANCE, 2005, 14(3): 388-394.
2. Liu S, Bai L, Wang B, el. Numerical simulations of solidification and hot tearing for continuous casting of duplex stainless steel[J]. JOURNAL OF IRON AND STEEL RESEARCH INTERNATIONAL, 2020, 27(6): 643-655.
3. Cai S, Wang Z, Wang S, et al. Physics-Informed Neural Networks for Heat Transfer Problems[J]. Journal of Heat Transfer, 2021, 143(6): 060801.
4. Gao W, Zhao J, Peng S. UNet–Based Temperature Simulation of Hot Dry Rock in the Gonghe Basin[J]. Energies, 2022, 15(17): 6162.
5. Raissi M, Perdikaris P, Karniadakis G E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations[J]. Journal of Computational Physics, 2019, 378: 686-707.
6. Han xiao. Research on Temperature Field Prediction Method of Sand Mold Casting Solidification Process Based on U-NET [D]. Beijing Jiaotong University,2022

Authors Introduction

Ms. Yahui Yang

She is currently pursuing her master's degree in Mechanical Engineering at Beijing Jiaotong University. She is mainly engaged in the research on deep learning analogue simulation.

Dr. Jiwu Wang

He is an associate professor, at Beijing Jiaotong University. His research interests are Intelligent Robots, Machine Vision, and Image Processing.

Dr. Jinwu Kang

He is an associate researcher and doctoral supervisor at Tsinghua University, specializing in the fields of intelligent casting and simulation modeling.

Portable Green Energy Mobile Laptop Charging Station

Ammar A.M. Al-Talib
Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment,
UCSI University, 56000 Kuala Lumpur, Malaysia

Rodney Tan
Department of Electrical and Electronic, Faculty of Engineering, Technology and Built Environment,
UCSI University, 56000 Kuala Lumpur, Malaysia

Ang Aun Jie
Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment,
UCSI University, 56000 Kuala Lumpur, Malaysia

Noor Idayu Mohd Tahir
Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment,
UCSI University, 56000 Kuala Lumpur, Malaysia

Sarah 'Atifah Saruchi
Faculty of Manufacturing and Mechatronic Engineering Technology,
Universiti Malaysia Pahang Al-Sultan Abdullah, 26600 Pekan, Pahang, Malaysia

Cik Suhana Hassan
Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment,
UCSI University, 56000 Kuala Lumpur, Malaysia

Amar Ridzuan Abd Hamid
Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment,
UCSI University, 56000 Kuala Lumpur, Malaysia
E-mail: ammart@ucsiuniversity.edu.my, rodneytan@ucsiuniversity.edu.my, 1001747031@ucsiuniversity.edu.my
nooridayu@ucsiuniversity.edu.my, sarahatifah@umpsa.edu.my, suhana@ucsiuniversity.edu.my,
amar@ucsiuniversity.edu.my

Abstract

Mobile phones and laptop computers require electrical power to recharge when the battery is down. As a result, it would be very useful if a portable charging station derived from renewable energy harvesting could be built, so that individuals could recharge their phones and laptops whenever needed. The objective of this project is to design and develop a green energy mobile and laptop charging station that uses wind and solar energy and evaluate the performance of the designed station under different working conditions. The efficiency of the power generated to charge the station is achieving 95.6% for solar charging, which is considered a high efficiency for a renewable energy charging station. Based on the analysis of the charging station results, it has been proven that it can provide sufficient power and is safe for use as a portable mobile laptop charging station.
Keywords: Energy conversion, Portable power supply, Universal mobile laptop charging

1. Introduction

Electricity is a vital part of modern life and for a country's economic growth and development. Over the next two decades, average power consumption is predicted to nearly double, with overall electricity demand expected to rise by 2.3 percent per year on average from 13,290 billion kWh in 2001 to 23,072 billion kWh in 2025 [1]. In Malaysia, home electric power consumption has grown substantially over the past years, resulting in high demand for energy to satisfy increasing social and economic activities.

Fossil fuels are the most famous resources used. However, these fuels are creating huge carbon emissions to the atmosphere which has negatively impacted the environment. As the demand for electric power rises, additional power generation is required to meet the demand. Due to the depletion of conventional resources, renewable energy resources are the most suitable resources for generating electric power with little environmental effect. Renewable or non-conventional energy are tidal energy, wind energy, hydropower, biomass, geothermal energy, and solar energy [2]. It also

Ammar A.M. Al-Talib, Ang Aun Jie, Idayu M.Tahir, Sarah Atifah Saruchi, Cik Suhana Bt. Hasan, Amar Rizwan, Rodney Tan

contributes to the achievement of Sustainable Development Goal 7 (SDG7), which calls for universal access to sustainable and renewable energy by 2030.

This project's goal is to create a portable dual-source charging station that can charge mobile phones and laptops using both solar and wind energy. A hybrid power system is a system that combines two or more renewable energy sources to provide continuous energy generation to fulfil load needs. Researchers have discovered that a combined solar and wind hybrid system is the most reliable and efficient, as well as being ecologically friendly and affordable in cost [3]. Several recent research [4], [5], [6] have combined two non-conventional sources to create a dual-mode portable mobile charger. Wind and solar hybrid systems for mobile chargers with minimal operating costs have been designed in this research. Additional components are necessary for laptop charging, which necessitates the use of an AC power source. When constructing a portable charging station for mobile and laptop devices, certain specifications and data are critical.

2. System overview

Major components of the system include power inverter, battery bank, system and battery controller, photovoltaic modules, wind turbines, and, in some cases, the prescribed electrical demands. Fig. 1 from MEDA [7] illustrates the system's basic working principle. The acquired energy from the wind turbine and PV array will be transformed into electrical DC power and stored in the battery bank. For the system under research, a monocrystalline solar panel and a vertical axis wind turbine are recommended. As compared to polycrystalline and thin-film solar panels, monocrystalline solar panels offer the highest efficiency of 15–20 percent [8]. According to the research [9], vertical axis wind turbines are more efficient to build and use in Malaysia's tropical and windy weather. A charge controller for wind and solar power is required to maintain your batteries properly supplied and avoid overcharging. For DC loads, the electric power stored in the battery bank may be utilized directly, however for AC loads, an inverter is required to connect to the battery bank. A power inverter is an electrical device that converts direct current (DC) from a battery bank to alternating current (AC). Finally, the system can be used to charge mobile phones and laptops.

2.1. *Fabricated prototype*

The prototype has been successfully built using
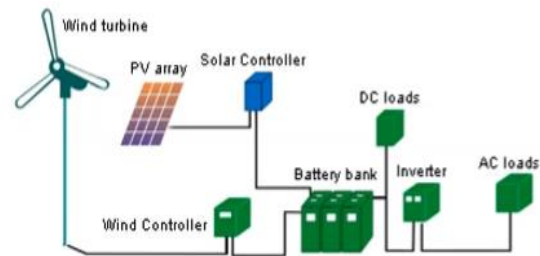
Fig. 1 A block diagram of Hybrid Solar-wind Energy system [7]

Fig. 2 Full components of prototype

Table 1 List of prototype components

No.	Component	Specifications
1	Power Inverter	200W
2	Dokio Solar charge controller	30A
3	PWM Solar Charge controller	10A
4	Battery Indicator Voltmeter	N/A
5	Dokio Solar Panel connector	N/A
6	SLA Battery	12V, 7.2Ah
7	Plastic Ammo Box	N/A
8	4-in-1 output panel	Voltmeter, switch, USB ports, 12V cigarette port

polystyrene boxes, a plastic ammo box, and cardboard. The components of the prototype are labelled in Fig. 2 and listed in Table 1. Fig. 3 shows the complete fabrication of the prototype.

The portable charging station consists of a Dokio solar charge controller, a PWM solar charge controller, a battery indicator voltmeter, a 12V, 7.2Ah GP SLA Battery, a 200W power inverter, and a 4-in-1 output panel as shown in Fig. 4. All the components in the circuit are linked in parallel with the 12V 7.2AH SLA Battery,

Fig. 3 The completed prototype

Fig. 4 The charging mobile phone and laptop with using the charging station

ensuring that the station can function even if one of the components fails. A switch, voltmeter, USB ports, and a 12V cigarette port are all included in the 4-in-1 output panel. The charging station can be turned on and the mobile phones can start charging through the USB port, while a power inverter must be connected to the 12V cigarette port to convert power into AC power for laptop charging.

3. Result and Discussion

The system has been tested to evaluate the performance of the designed station. The equipment used in the experiments was an 80 W Mono-solar panel, a 100 W wind turbine, and a 1000 W halogen lamp to generate electrical power. Several experiments have been conducted and the data were recorded and analyzed to evaluate the performance of the designed station under different working conditions. A multi-meter has been used to measure the current and voltage output from the solar panel and wind turbine, and then the energy output over the state of the battery of the experiments have been measured by the charge controller. The experiments have

Table 2. Data collected for four different experiments conducted

Experiment	Trial	Average Voltage Generated (V)	Average Current Generated (A)	Charging Duration (Hour:minute)	Solar Irradiances (W/m²)	Light Intensity (W/m²)	Speed of Wind (m/s)
1	1	12.75	0.935	2:42	734.95	-	-
	2	12.67	0.873	3:01	679.46	-	-
	3	13.50	1.450	1:08	839.69	-	-
	4	13.24	1.140	1:36	837.85	-	-
	5	12.85	0.982	2:14	760.80	-	-
	6	12.72	0.955	2:38	745.23	-	-
2	1	12.41	0.340	18:18 (Not completed)	-	238.17	-
	2	12.62	0.450	24:25	-	451.99	-
	3	12.65	0.480	19:09	-	493.19	-
3	1	11.30	0.250	23:00 (Not completed)	-	-	3.0
	2	11.32	0.275	23:12 (Not completed)	-	-	3.3
	3	11.38	0.300	22:30 (Not Completed)	-	-	3.6
4	1	12.64	0.750	14:51	-	474.61	3.6

Ammar A.M. Al-Talib, Ang Aun Jie, Idayu M.Tahir, Sarah Atifah Saruchi, Cik Suhana Bt. Hasan, Amar Rizwan, Rodney Tan

been conducted to fully charge the 12V 7.2Ah Sealed, Lead Acid Rechargeable Battery for each trial of experiments.

Experiments 1 and 2 are solar power charging for the designed station, with experiment 1 using sunlight to generate electrical power outdoor under the sun, whether experiment 2 is by utilizing a halogen lamp as the source of light to simulate the solar energy. The wind power charging for the station is evaluated with a power fan in experiment 3 to simulate the wind energy because of the low wind speed available outdoors. Experiment 4 is a hybrid solar and wind power charging system for the station, with a halogen lamp and a power fan used indoors to get more consistent and precise results. Table 2 shows the data collected for the four different experiments conducted.
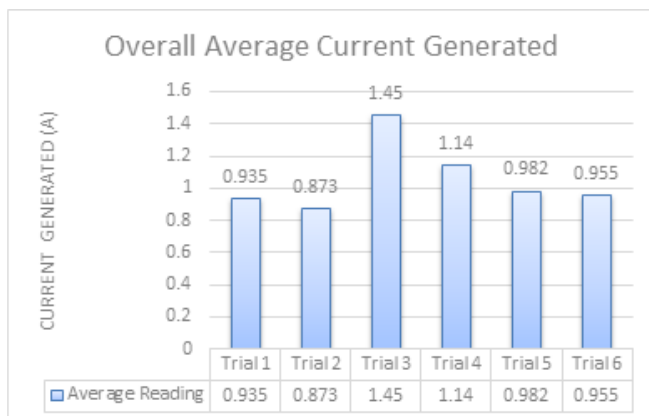
Fig. 5 Overall average current generated and solar irradiance for experiment 1
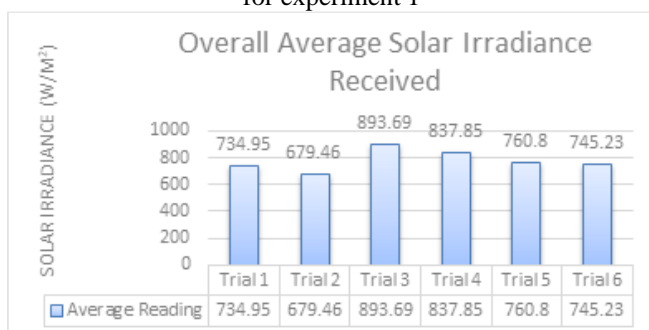
Fig. 6 Overall average solar irradiance received for experiment 1.
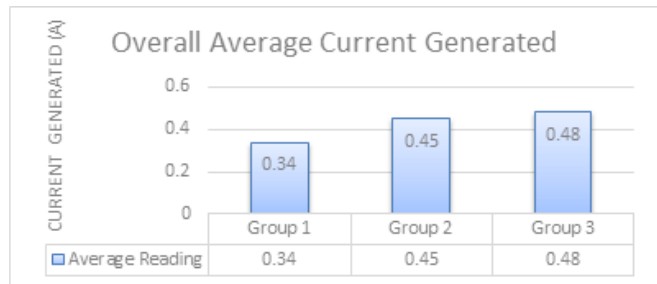
Fig. 7 Overall average current generated for experiment 2.
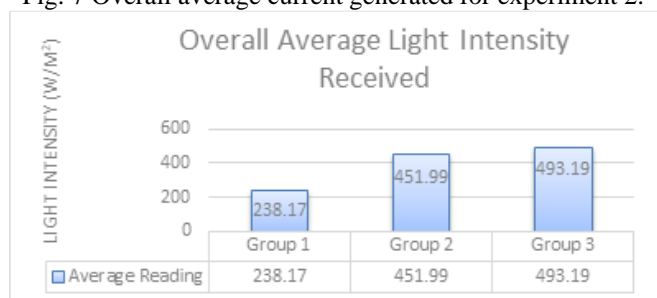
Fig. 8 Overall average light intensity received for experiment 2.
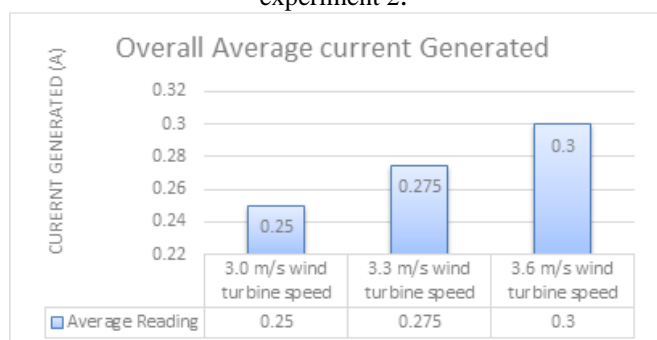
Fig. 9 Overall average current generated with three different speeds for experiment 3.

The average overall electrical current generated and overall average solar irradiances received for experiment 1 are plotted and illustrated in Fig. 5 and Fig. 6 based on the data obtained. In Fig. 7 and Fig. 8, the overall average current generated and overall average light intensity received for experiment 2 are plotted and shown. Fig. 9 illustrates the relationship between current generated and wind speed.

According to Fig. 5 and Fig. 6, it is shown that the higher the solar irradiances come the higher the current generation output result. The results reveal that even the lowest solar irradiation of 679.46 W/m2 can generate an average of 12.67 volt and 0.873 amperes.

Fig. 7 and Fig. 8 are showing that as the light intensity increase, the current generated and output power increase.

As the solar panel receives more light intensity from the halogen lamp, the current generated is increasing as well as the power output. Although the light source to the solar panel is stable, the current and power generation is respectively low which is not utilizing the sunlight that contains higher solar irradiances. As a result, it also corresponds to the research done by [10], whereby the current rises steadily with an increase in solar irradiances or light intensity.

 Fig. 9 shows that when the wind speed increase, the current generated will also increase. However, it can still be observed that the current is increasing with the speed of wind increase which gets a maximum of 0.3A with a 3.6 m/s wind speed. As a result, it can be concluded that the greater the wind speed, the greater the magnitude of the current.

From Table 2, it is shown that an average of 0.75A is generated with an average light intensity of 474.61 W/m^2 and a constant speed wind of 3.6m/s. The results have been taken when both solar and wind harvesting systems are working together to charge the station. The duration to fully charge the station is around 14.85 hours which is faster than experiments 2 and 3.

The duration to fully charge the battery is also dependent on the renewable power supply. The higher the power supply to charge the station, the shorter the duration of the battery to be fully charged. The fastest battery of 7.2 Ah capacity could be fully charged in around 68 minutes with a charging current of 1.45 A. To protect the battery, the battery is recharging at 12 V which is around 50 % to avoid irreversible damage occurring. The fully charged battery can fully charge an iPhone 11 with a battery capacity of 11.9 Wh for 5.5 cycles and an Asus gaming laptop with a battery capacity of 48 Wh for 1.8 cycles.

3.1. *Result comparison*

The average experimental results from experiment 1 through experiment 4 are plotted and illustrated in Figure 10. The average result is calculated by averaging the measurements based on the state of charge of the battery, which acts as a comparison point. This is because the charge controller will supply power following the battery's status, which means that the power delivered will gradually grow following the status of the battery to prevent the battery from being damaged.

According to Fig. 10, experiment 1 is the most efficient

	Experiment 1	Experiment 2	Experiment 3	Experiment 4
Average Voltage (V)	12.95	12.56	11.33	12.64
Average Current (A)	1.06	0.42	0.275	0.75
Average Duration to Fully Charge (hours)	2.2	20.62	22.8	14.85
Theoretical Duration to Fully Charge (hours)	9.51	24	36.65	13.44
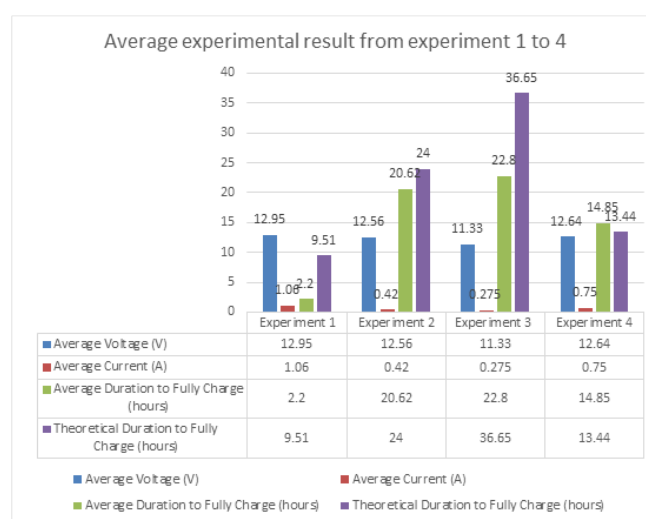
Fig. 10 Average experimental result from experiment 1 to experiment 4.

way, taking just 2.2 hours to fully charge the station with an average voltage of 12.95V and a current of 1.06A. It took around 7.31 hours less time than the expected result based on calculation. Even though the readings fluctuated during the experiment, it was still the most efficient method for solar charging when compared to experiment 2 which utilized a halogen lamp as the light source.

From Fig. 10, experiment 2 took an average of 20.62 hours to completely charge the station battery, which is around 18.4 hours longer than experiment 1. With an average of 12.56 V and 0.42A, the power generated was likewise significantly lower than experiment 1. It took around 3.38 hours less time than the expected result based on calculation. As a result, the most effective approach for successfully charging the station batteries for the solar harvesting system is charging with a solar panel directly under the sun, which will maximize the power generated among other methods.

Ammar A.M. Al-Talib, Ang Aun Jie, Idayu M.Tahir, Sarah Atifah Saruchi, Cik Suhana Bt. Hasan, Amar Rizwan, Rodney Tan

Due to the low wind speed available in Malaysia, experiment 3 aims to see how much power can be generated at various wind speeds. It only can generate an average of 11.33V and 0.275A, which only charged station batteries to 60% in an average of 22.8 hours. This is because the charge controller utilized for the wind harvesting system is a solar charge controller, resulting in the charge controller power output being unstable and low. Furthermore, the solar charge controller is only designed for solar power systems with electrical characteristics that differ from those of wind power, resulting in low power received to charge the station. As a result, by adopting an appropriate wind charge controller, the wind power charging system may be more effective.

Experiment 4 feeds the station using both solar and wind harvesting technologies. It aims to see if operating the combined harvesting system at the same time as charging the station is more efficient than working independently. According to Figure. 10, the result showed an average of 12.64V and 0.75A, and it took an average of 14.8 hours to fully charge the station. However, the required duration to completely charge took 1.41 hours longer than the theoretical calculation. This outcome might be impacted by the fact that an inappropriate charge controller is used for a combined solar and wind harvesting system that prolongs the duration to be fully charged. As a result, it can be proven that the combined solar and wind harvesting system has a better performance than a system that only harvested wind or solar.

4. Conclusion

In summary, the design of a mobile dual source charging station, by using wind energy and solar energy, has been fabricated with several materials such as polystyrene boxes, cardboard, and plastic ammo toolbox. The performance of the designed station under different working conditions was investigated and evaluated using handphone and laptop as load. Based on the comparison, the designed station has the best performance when both solar and wind harvesting systems are working at the same time. Therefore, this study is promising to develop a clean energy portable charging station that is environmentally friendly and able to act as a charging station when there is no power supply available.

References

1. N.A. Mahusin and R. Baharun, Utility Consumption among Malaysian Electricity Users in Government Buildings. International Symposium on Technology Management and Emerging Technologies, 2014.
2. P. Pranjela and K.H Deepali, Non Conventional Sources of Energy-Applications of Solar Energy in Architectural Buildings, International Journal of Computer Applications 2013
3. K.V. Anila, Design of a Hybrid Power Generation System Using Solar-Wind Energy, M.S. Thesis, Deakin University, Geelong, Australia, 2018.
4. R. Chakma, T. Chawaphan, K. A. A. Mamum, A. Chakma, S. Harun, Portable Smart Phone Charger Using Human Mechanical Energy by Gear Train with Hand Crank, Journal of Electrical and Electronics Engineering, Vol 12, Issue 3, pp. 20-25, June 2017.
5. Md. A. Rahaman, N. Hoque, N. K. Das, F. N. Maysha, MD. M. Alam, Portable Dual Mode Mobile Charger with Hand Crank Generator and Solar Panel, Indonesian Journal of Electrical Engineering and Computer Science, Vol. 1, No. 2, February 2016, pp. 282 ~ 287.
6. P. Vijay, T. Manglani, P. Kumar, R. Meena, A. Khedia, Wind and Solar Mobile Charger, International Journal of Recent Research and Review, Vol. 7, Issue 4, Dec 2014.
7. Mahrashtra Energy Development Agency (MEDA), Small Wind Energy and Hybrid Systems Programme. https://www.mahaurja.com/meda/en/off_grid_power/small_wind_solar_hybrid
8. A. Saleem, K. Mehmood, F. Rashid, The Efficiency of Solar PV System, Proceedings of 2nd International Multi-Disciplinary Conference 19-20 December 2016.
9. M. K. Johari, M. A. A. Jalil, M. F. M. Shariff, Comparison of horizontal axis wind turbine (HAWT) and vertical axis wind turbine (VAWT), International Journal of Engineering & Technology, Vol 7, no 4, (2018), pp 74-80.
10. Z. Li, J. Yang, P. A. N. Dezfuli, Study on the Influence of Light Intensity on the Performance of Solar Cell, International Journal of Photoenergy, vol. 2021, 2021.

Authors Introduction

Ammar Abdulaziz Al T

He has finished his B.Sc and M.Sc degrees in Mechanical Engineering from the University of Mosul/Iraq.
He has finished his Ph.D degree from UPM University / Malaysia.
He is also a Chartered Engineer and Member of the Institute of Mechanical Engineers / UK. (CEng. MIMechE). He has developed all the Postgraduate Programs at the Faculty of Engineering at UCSI University / Malaysia, and worked as the Head of Postgraduate and Research department at the same faculty for the years 2010-2018.
alib

Rodney Tan

He currently is an Associate Professor at UCSI University, Malaysia. He received his PhD degree in Electrical Engineering in 2013 from UNITEN Malaysia. His research interests include Renewable Energy, Energy Storage System and Power Electronic Converter.

Ang Aun Jie

He graduated from the Mechanical Engineering Department at UCSI University in 2021, and currently is working as an Engineer.

Noor Idayu Mohd Tahir

She graduated from University Teknologi MARA, Malaysia in Mechanical Engineering. She was design engineer in oil and gas industries for 7 years. Currently, she works as a lecturer under the Department of Mechanical and Mechatronics, Faculty of Technology and Built Environment, UCSI Kuala Lumpur. Her research interests include design and autonomous robotics.

Sarah 'Atifah Saruchi

She graduated from Nagoya University, Japan in Mechanical and Aerospace Engineering. She received her Master and Doctoral degrees from Malaysia-Japan International Institute of Technology (MJIIT), Universiti Teknologi Malaysia. Currently, she is working as senior lecturer at Universiti Malaysia Pahang Al-Sultan Abdullah, Pahang, Malaysia. Her research interests include mechatronics and artificial intelligence

Cik Suhana Hassan

She is an Assistant Professor at the Department of Mechanical and Mechatronics Engineering of UCSI University. She received her bachelor's and master's degrees in 2009 and 2011, respectively, from Universiti Teknologi PETRONAS, and her PhD in 2019 from Universiti Putra Malaysia. Her research interests include the investigation of bio-composites for use in automotive applications. She is also an active member of the materials community, having been recognized as a Professional Member of the Institutes of Materials Malaysia and a Professional Technologist of the Malaysian Board of Technologists in the Material Science Technology field

Amar Ridzuan Abd.Hamid

He is a Professional Technologist, Head of Programmes and lecturer of Mechanical and Mechatronic programmes from Department of Mechanical Engineering, UCSI University, Malaysia. He received his Bachelor's degree in automotive engineering (Hons) from UTeM, Melaka in 2011 and his Master's degree in Mechanical Engineering from UPM, Selangor in 2015. His research interests are in the design and analysis of automotive compartments.

Auto Indoor Hydroponics Plant Growth Chamber

Ammar A. M. Al-Talib

Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment,
UCSI University, 56000 Kuala Lumpur, Malaysia

Tew Hwa Hui

Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment,
UCSI University, 56000 Kuala Lumpur, Malaysia

Sarah 'Atifah Saruchi

Faculty of Manufacturing and Mechatronic Engineering Technology,
Universiti Malaysia Pahang Al-Sultan Abdullah, 26600 Pekan, Pahang, Malaysia

Noor Idayu Mohd Tahir

Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment,
UCSI University, 56000 Kuala Lumpur, Malaysia

Nor Fazilah Binti Abdullah

Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment,
UCSI University, 56000 Kuala Lumpur, Malaysia
E-mail: ammart@ucsiuniversity.edu.my, 1001955085@ucsiuniversity.edu.my, sarahatifah@umpsa.edu.my,
nooridayu@ucsiuniversity.edu.my, norfa@ucsiuniversity.edu.my

Abstract

The objective of this project is to build an auto indoor hydroponics plant growing chamber that has an auto monitoring and controlling system. A ESP32 based hydroponics electrical system is built with the attachment of hardware components such as temperature and humidity sensor, light intensity sensor, water level sensor, and water flow rate sensor. The software development of the system is through Arduino IoT Cloud platform, which has an overall suitability in terms of features, cost, and user intuitiveness for starters. Results have shown that ESP32 can ensure stable power supply. After testing and validation, all of the electrical components are stored in a power enclosure box to prevent contact with liquid. In short, the developed auto indoor hydroponics plant growth chamber has effectively demonstrated the ability in easing the plant cultivation procedure for agricultural community.

Keywords: Auto Indoor Hydroponics, ESP32, Controlling and Monitoring Hydroponics

1. Introduction

The global population is expected to reach 9.8 billion in 2050 [1]. Meanwhile, scientists anticipate that a growing population will result in a decline in the amount of available land. The scarcity of land in the future is predicted to become a bottleneck in all industries, but especially in agriculture [2]. Not to mention, a lack of available land will make traditional farming difficult and increase the likelihood of a food crisis due to rising food demand but low yield rates [3].

All the aforementioned issues have made it difficult to provide adequate and nutritious food, particularly in urban areas. Therefore, new methods of producing enough food must be developed in order to sustainably feed the world's expanding population. New and modern farming techniques marked a significant technological advance for humanity. The substitution of liquid as the new growing medium for traditional growing medium can be the different approach for consistent crop production, preservation of rapidly depleting land and food availability. On that being in case, the rapid growth of hydroponic system is getting attention and increasing

Ammar A. M. Al-Talib, Tew Hwa Hui, Sarah 'Atifah Saruchi, Noor Idayu Mohd Tahir, Nor Fazilah Binti Abdullah

widespread all over the world. It is accounted to have a world growth of 18.8% from 2017 to 2023 [4]. Hydroponics is derived from Greek word means Hydro "water" and Pons "labour". It is a method to grow various plants in water containing fertilizer without any soil by using different inert mediums such as sand, gravel, rockwool that act as a mechanical support [5]. The growth of hydroponics plants must be in a controlled environment and often affected by different parameters such as electrical conductivity in nutrients, nutrients circulation rate, PH, light intensity air and water temperature. Hence, the plant can have growth in an optimal condition with frequent monitoring for better yield rate of the plants [2].

Therefore, this project aims to develop an auto indoor hydroponic system prototype with "auto" monitoring and controlling features with IoT capabilities.

2. Literature Review

In this section, understanding and evaluating the monitoring and controlling systems from past research studies can be well understood.

2.1. Water and Nutrient Flow Path Monitoring and Controlling System

Water flow path and control takes a critical role in hydroponics system as the growth condition of plant was highly determined by the nutrients absorbed.

Fig.1. refers to the system used by M. E. H. Chowdhury et al., 2021 [6]. It contains two circulation paths: nutrient solution circulation and clean water supply.

The pump on the clean water source will only operate when the nutrient container does not meet the height of the water level with the aid of the water level sensor. The placement of the water level sensor was determined with the water height formula used; the water level sensor of the system will be placed 22.4cm vertically on the inner container to ensure enough 37L water in the tank.

An additional water flow sensor has been provided on the nutrient container. The water flow sensor was capable of measuring volume circulation per day and water. The system will notify when there is no water flow sensing.
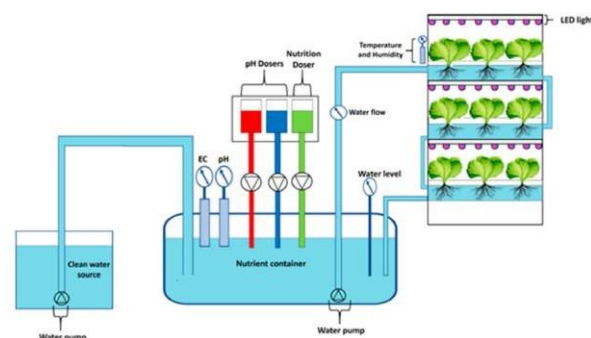
Fig. 1. Water, Nutrient and pH Monitor and Control System

2.2. Temperature and Humidity Monitoring and Controlling System

Temperature and Humidity around the plants is crucial as it may affect the growth of plants in different temperature environments.

Fig. 2. refers to the system used by S. Seni et al, 2020. [7]. It is an indoor hydroponics project in Malaysia that grow Pak Choy. There were two environmental sensors and two exhaust fans on each planting level on the system to ensure proper air circulation and detection. It's also observed that the reading of sensors values on different levels will be slightly different.

The temperature of the project is set within the allowable operating range of Pak Choy by implementing environment sensors and exhaust fans. The exhaust fans will be operating when exceed the allowable operating range. The project also identified that the temperatures will periodically rise in the afternoon that will cause continuous operation of exhaust fans in the afternoon.

Fig. 2. Temperature and Humidity Controlling System

2.3. *Light Monitoring and Controlling System*

Referring to another study about effects of Artificial Night LED Lighting by L. S. Wei et al, 2019, [8]. The proposed hydroponics system in Malaysia cultured the plant with a 18/6 (light/dark) photo period. Based on Table 1 below, it shows that the plant was cultured with 8hours of artificial lighting and 16 hours with natural lighting that continuously operate until plant maturity day.

Table 1. Light Duration Used in the Study

Natural Light without LEDs Support	Night-time without LEDs Support	Night-time with LEDs support
8am to 6pm	6pm to 12am	12am to 8am

2.4. *IoT Monitoring System*

Fig. 3. shows the overall system setup for the monitoring process. The proposed system is an Arduino Mega based hydroponic system with MQTT protocol to communicate between Arduino Mega and Node MCU for data processing. This system proposed precise monitoring of the nutrient solution's properties, such as electrical conductivity (EC), pH, and temperature, to ensure that the plants receive optimal levels of nutrients.The sensors result will then convert to JSON format and sent to cloud platform for monitoring. Overall, the proposed system is an important step forward in the field of hydroponics farming. It allows for precise and automated monitoring of nutrient solution levels reducing the need for manual labor of the farming process.
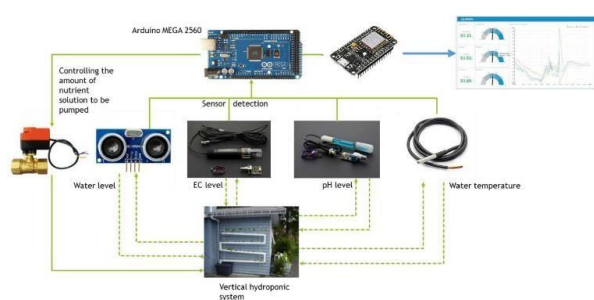
Fig. 3. IoT Monitoring Hydroponics System

3. Methodology and Setup

In this section, the monitoring and controlling systems from past research will be presented.

3.1. *Mechanical Design and Setup*

Fig. 4. shows the proposed hydroponics mechanical design with multiple main components. The hydroponic structure is arranged in a horizontal shape which has an overall dimension of around 1.25m x 0.79m x 1.2m (W x L x H) with 2 layers of planting platforms that can cover a total of 72 plants with 38 plants each layer. The system has additionally equipped with two ball valves that can help in controlling the water circulation and water level to void overflow piping [9]. The system has been installed in a chamber for easily controlled environments. The hydroponics system has mainly been using UPVC pipe as the structural material of the design for easily accessible, fabrication and assembly friendly as well as cost effective. Eventually, the chamber is also smaller in size as compared to a 160cm human being.
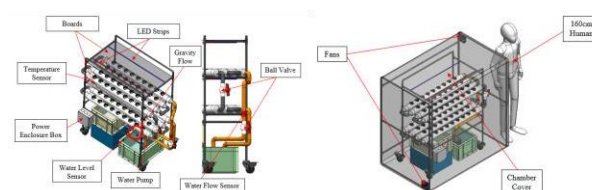
Fig. 4. Hydroponics Design with Main Components

3.2. *Electrical System Infrastructure*

Fig. 5. shows the proposed electrical schematic designed with Proteus. The proposed design has utilized ESP32 as the mainboard and different sensors such as the DHT11 for temperature and humidity sensing, the YF-S201 for water flow rate sensing, a water level sensor for nutrient level sensing, as well as light sensor module for light intensity sensing. Relays and actuating devices have been used to control the system automatically. In general, the total cost of components required to set up the system is about RM185.60

Ammar A. M. Al-Talib, Tew Hwa Hui, Sarah 'Atifah Saruchi, Noor Idayu Mohd Tahir, Nor Fazilah Binti Abdullah
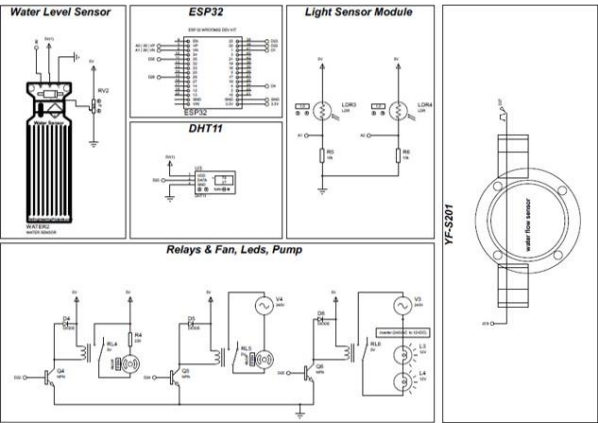
Fig. 5. IoT Monitoring Hydroponics System

3.3. Software System Flow

Ventilation Flow Chart

Fig. 6. shows the hydroponics system monitoring and control flow chart. The main controller, an ESP32, will retrieve data from the sensor, DHT11. The ventilation fan will turn on when the temperature have reached the threshold set, vice versa. Meanwhile, the main microcontroller will then publish that data to the Arduino cloud to allow real-time monitoring for the user.
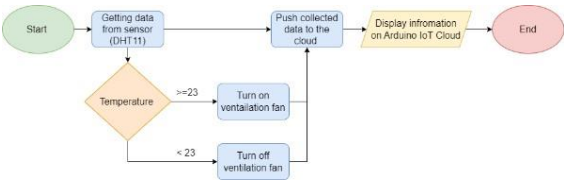
Fig. 6. Auto Ventilation Flow Chart

Grow Light Flow Chart

Fig. 7. shows the hydroponics system monitoring and control flow chart. The main controller, an ESP32, will retrieve data from the NTP servers and LDRs sensors. The LED strips will turn on when the time has reached the threshold set, vice versa. Meanwhile, the main microcontroller will then publish that data to the Arduino cloud to allow real-time monitoring for the user.
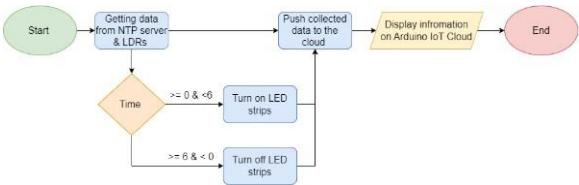
Fig. 7. Auto Grow Light Flow Chart

Nutrient Level and Water Flow Rate Flow Chart

Fig. 8. shows the hydroponics system monitoring and controlling flow chart. The main controller, an ESP32, will retrieve data from the water flow and water level sensors. The monitoring system will notify the user when the water level and water flow has reached the threshold set. Meanwhile, the main microcontroller will then publish that data to the Arduino cloud to allow real-time monitoring for the user.
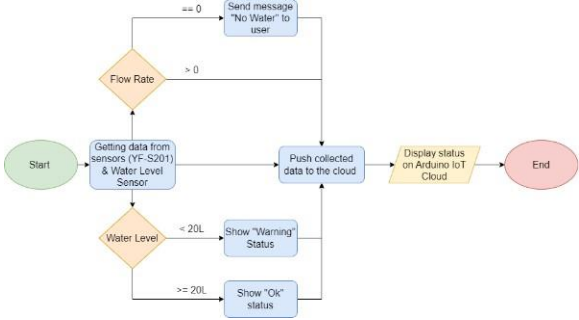
Fig. 8. Auto Grow Light Flow Chart

4. Results and Discussion

4.1. Testing, Analysis and Validation

Table 2 below shows the comparison of measured voltage and actual voltage of power supply pin on the microcontroller. The percentage error of most pins is within an acceptable range. However, the percentage error for the 3V pins is 10.33%, indicating that there is an issue with the expansion board. This might be due to the voltage drops across the pcb trace of the used expansion board. Nonetheless, since the measured voltages are 3.31V, it is possible that there is wrong labelling on the

board. The 3V pins will be used as a 3.3V power supply to the system.

Table 2. Measured Voltage and Actual Voltage

Input	Pin Symbol	Measured Voltage (V)	Actual Voltage (V)	Percentage Error (%)
5V 2A	VIN	4.98	5.00	0.40
	5V (1)	4.92	5.00	1.60
	5V (2)	4.91	5.00	1.80
	3.3V	3.31	3.33	0.60
	3V (1)	3.31	3.00	10.33
	3V (2)	3.31	3.00	10.33

Temperature And Humidity Testing and Analysis

Table 3 and Table 4, Are showing the comparison of temperature and humidity of used sensor and measurement tools. The results are retrieved every 5 minutes until 30 minutes. Based on the results, it is observed that the temperature measurements have an average percentage difference of 1%. Nevertheless, humidity measurements have an average percentage of 3.53%. Both measurements are relatively acceptable.

Table 3. Measured and Actual Temperature

Input	No	Temperature (°C)		% Difference
		DHT 11 Sensor	Measuring Tool	
23°C	Initial Data	23.50	23.20	1.29
	Data at 5 Minutes	23.40	23.20	0.86
	Data at 10 Minutes	23.40	23.20	0.86
	Data at 15 Minutes	23.30	23.10	0.87
	Data at 20 Minutes	23.30	23.00	1.30
	Data at 30 Minutes	23.20	23.00	0.87
Average Error Percentage Difference (%)				1.00

Table 4. Measured and Actual Humidity

Input	No	Relative Humidity (%)		% Difference
		DHT 11 Sensor	Measuring Tool	
23°C	Initial Data	60	58	3.45
	Data at 5 Minutes	59	57	3.51
	Data at 10 Minutes	59	57	3.51
	Data at 15 Minutes	58	56	3.57
	Data at 20 Minutes	58	56	3.57
	Data at 30 Minutes	58	56	3.57
Average Error Percentage Difference (%)				3.53

Table 5 and Fig. 9 are for the Water Level Testing and Analysis. It shows the ADC reading of the water level sensor corresponding with the height of the measuring tools. The ADC input channels of ESP32 have a 12-bit resolution, this means that analog readings of 0 to 4095 shall be obtained. However, the measured value is only ranging from 0 to 2019 only, this might be due to the corrosion of the water level sensor. The testing results helped in benchmarking the operating range of the water level sensor. Therefore, the water level sensor can be used in this system for water level measurements.

Table 5. Comparison of Level Sensor and Ruler

No	Water Level Sensor (12-bit ADC Reading)	Measuring Tools (cm)
1	0	0
2	976	1
3	1454	2
4	1703	3
5	2019	4

Ammar A. M. Al-Talib, Tew Hwa Hui, Sarah 'Atifah Saruchi, Noor Idayu Mohd Tahir, Nor Fazilah Binti Abdullah
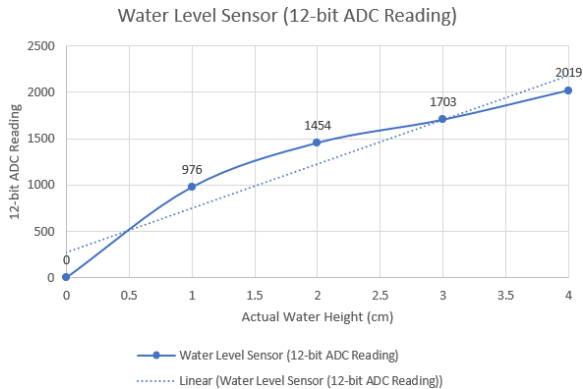
Fig. 9. Actual Water Height Correspond to 12-bit ADC

For the water flow rate sensor testing and analysis, a mathematical equation is used to calculate the flow rate of the system is as follows, [6].

$$Flow\ Rate\ (L/min) = \frac{Sensor\ Frequency\ (Hz)}{7.5}$$

Based on Fig. 10, the obtained data suggests a linear relationship between the valve angle of rotation and the water flow rate, this proves that the water flow rate increases with increasing the valve angle. The R^2 value of 0.9635 indicates that the data points are well-fitted to the linear regression line, indicating that the valve angle and water flow rate have a significant correlation.
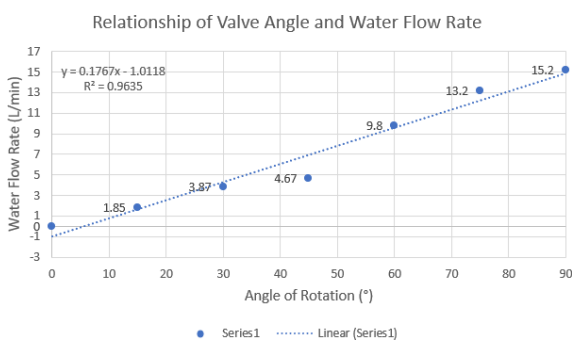
Fig. 10. Relationship of Valve Angle and Water Flow Rate

For the Light Intensity Testing, and based on the measurement conducted by comparing the light intensity sensor and lux meter in Table 5, it's identifying that the room lights itself is not enough to supply enough light intensity on the hydroponics system. The percentage difference of sensors and lux meter has less than 18% at different room conditions. It indicates that the light intensity sensor can be implemented in this prototype. Nevertheless, it's important to place the hydroponics system near a room that has exposure to sunlight for enough brightness for plants to conduct photosynthesis.

Table 5. Light Intensity Sensor and Lux Meter

Room Condition	Light Intensity Sensor (lux)		Lux Meter (lux)		% Difference	
	Top	Bottom	Top	Bottom	Top	Bottom
Sunlight	128.0	94.0	151.30	105.40	15.40%	10.82%
Room Lights	41.0	17.0	50.0	19.30	18.00%	11.92%
Without Lights	0	0	0	0	0%	0%

For the Relay Testing, Table 6 shows the results of controlling ventilation fan with relay, it is noticed that the ventilation fan will be turned on every afternoon. The fan will be turned off during mornings and nights. This is due to the reason that the ventilation fan will be turned on when reaching the state point.

Table 6. DHT11 and State of Fan

Threshold	Time (a.m. / p.m.)	Temperature of DHT 11 Sensor (℃)	State of Ventilation Fan (ON / OFF)
29.50℃	Morning 9 a.m.	25.40	OFF
	Afternoon 3 p.m.	30.30	ON
	Night 11 p.m.	27.80	OFF

Table 7 shows that the threshold was set, where the LED strips will be only turned on during 12 a.m. to 8 a.m. The LED strips will only turn on when it's midnight.

Table 7. Results of Time and State of LED

Threshold	Time (a.m. / p.m.)	State of LED Strips (ON / OFF)
12a.m. to 8a.m.	Morning 6 a.m.	ON
	Afternoon 12 p.m.	OFF
	Mid-night 1 a.m.	ON

4.2. Hardware Assembly and Implementation

Fig. 11. shows the detail of the placement of the electrical components. All the electronics components are placed in an enclosure to prevent direct contact with the water and easy for maintenance. The temperature sensor is located on the top layer of the hydroponics system. Meanwhile, there is also a light intensity sensor that is placed nearby. The second light intensity sensor is placed on the second layer. Next, the water flow rate and water level sensor are placed along the flow of circulation and inside the nutrient container. The ventilation fan is placed on the front of the chamber due to the reason that it can provide the lowest wire connection distance.

Fig. 11. Placement of Electrical Components

Final Implementation

Fig. 12. shows the final assembly of the hydroponics system. The assembly of the hydroponics system started with the structure, continued with circulation channels, chamber and finally electrical and electronics system.

Fig. 12. Mechanical Final Assembly

4.3. Software and Interface Implementation

Fig. 13. and Fig. 14. shows the Arduino IoT cloud dashboard and mobile app that measures the temperature, humidity, water flow rate, water level and light intensity on top and bottom layer. There is only one page of the dashboard and app interface of the system for user friendliness. The top part of the interface shows the measurement values and status of sensors in real time. Meanwhile, the bottom part of the interface shows the measurement values retrieved over time.
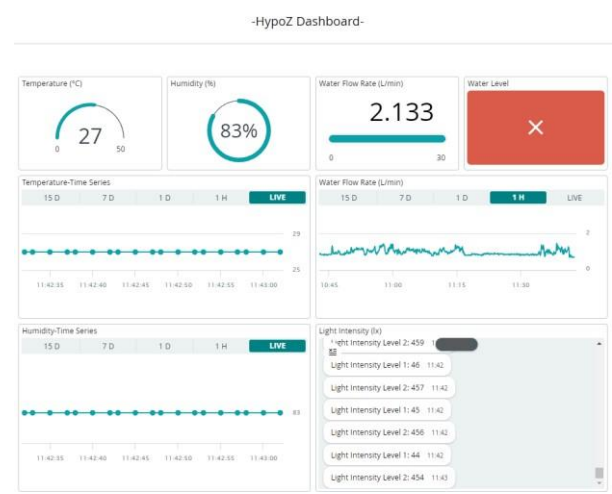
Fig. 13. Web Interface of Arduino IoT Cloud System

Fig. 14. App Interface of Arduino IoT Cloud System

4.4. Hydroponic System Comparison

Table 8 shows two different similar commercial solutions as compared to the proposed system which are all solutions to indoor agriculture. It can be seen obviously that the proposed system offers more features with a monitoring and controlling system. Yet, remaining significantly cheaper and affordable to anybody to venture into agriculture.

Besides that, the proposed system has a medium power consumption as compared to the Grow IT premium kit. This indicates that the system has consumed less energy and power.

In general, the proposed system provides a cheaper and more features alternative for indoor farming, which is also suitable for growers that just started to venture into agriculture.

Table 8. Comparison of Hydroponics System

Similar Work or Commercial Solution	HypoZ	City Vertical Farm L	GrowIT Premium Kit
Plant Capacity	72	48	70
Volume	1.19 m^3	1.48 m^3	1.31 m^3
IoT Solution	Yes	None	None
Light Monitoring	Yes	None	None
Temperature Monitoring (M) and Control (C)	M and C	None	None
pH Monitoring (M) and Control (C)	None	None	None
Power Consumption	58.665 W	18 W	120 W
Price Information	RM 700.07	RM 2899.00	RM 2999.00
Vendor	Self Developed	CityFarm	GrowEatWell

5. Conclusion

The designed and tested prototype has shown successful results in achieving the objectives of this project. There is still room for improvements. It is recommended that the hydroponics structure can be attached with additional features such as air- conditioning and humidity controlling devices to effectively control the growing environment of the plants. Besides, features like pH and EC monitoring and controlling devices can be equipped, this may help to adjust accurately nutrient solutions up taken by the plants.

Acknowledgment

References

1. Nations, U. World Population Projected to Reach 9.8 Billion in 2050, and 11.2 Billion in 2100 Available online: https://www.un.org/en/desa/world-population-projected-reach-98-billion-2050-and-112-billion-2100 (accessed on 18 July 2022).
2. Shubhashree, B.S.; Hiremath, H.M. Design of Hydroponics System for Remote Automation. *Int. J. Eng. Res.* 2020, 8, 3.
3. Okemwa, E. EFFECTIVENESS OF AQUAPONIC AND HYDROPONIC GARDENING TO TRADITIONAL GARDENING. Available online: https://www.semanticscholar.org/paper/EFFECTIVENESS-OF-AQUAPONIC-AND-HYDROPONIC-GARDENINGOkemwa/4b3978feb2b103a04026bb4f7d46213b9a6a928 5 (accessed on 29 July 2022).
4. Aires, A. Hydroponic Production Systems: Impact on Nutritional Status and Bioactive Compounds of Fresh Vegetables. In *Vegetables - Importance of Quality Vegetables to Human Health*; Asao, T., Ed.; InTech, 2018 ISBN 978-1-78923-506-7.
5. Sharma, N.; Acharya, S.; Kumar, K.; Singh, N.; Chaurasia, O. Hydroponics as an Advanced Technique forVegetable Production: An Overview. *J. Soil Water Conserv.* 2019, 17, 364–371, doi:10.5958/2455- 7145.2018.00056.5.
6. Chowdhury, M.; Khandakar, A. Design, Construction and Testing of IoT Based Automated Indoor Vertical Hydroponics Farming Test-Bed in Qatar. *Sensors* 2020, 20, doi:10.3390/s20195637.
7. Seni, S.; Audah, L. Automated Vertical Hydroponic Farming. *Evol. Electr. Electron. Eng.* 2020, 1, 219–225.
8. Chua Mason, H.S.; Paramasivam, S.; Thing Thing, G.; Gee, C. *Effects of Artificial Night LED Lighting on The Loose Head Lettuce Growth in Hydroponic System*; 2019;
9. Michael, G.; Tay, F.; Then, Y. Development of Automated Monitoring System for Hydroponics Vertical Farming. *J. Phys. Conf. Ser.* 2021, 1844, 012024, doi:10.1088/1742-6596/1844/1/012024.

Authors Introduction

Ammar Abdulaziz Al Talib

He received his B.Sc and M.Sc degrees in Mechanical Engineering from the University of Mosul Iraq.
He has finished his Ph.D degree from UPM University, Malaysia. Member of the Institute of Mechanical Engineers UK. (CEng. MIMechE). He has developed all the Postgraduate Programs at the Faculty of Engineering at UCSI University / Malaysia and worked asthe Head of Postgraduate and Research department at the same faculty for the years 2010-2018.

Tew Hwa Hui

He is currently pursuing Bachelor of Mechanical Engineering with Honours as final year student in Faculty of Engineering Technology University, Malaysia. His research interests are automation & control,mechanical design, and robotics.

Sarah 'Atifah Saruchi

She received her B.Eng. in Mechanical and Aerospace Engineering from Nagoya University, Japan. She received her Master and Doctoral degrees from Malaysia-Japan International Institute of Technology (MJIIT), Universiti Teknologi Malaysia. Currently, she is working as senior lecturer at Universiti Malaysia Pahang Al-Sultan Abdullah, Pahang, Malaysia. Her research interests are control, mechatronics and Ai.

Noor Idayu Mohd Tahir

She received her B.Eng in Mechanical Engineering from Universiti Teknologi MARA, Malaysia. She was a design engineer in oil and gas industries for 7 years. Currently, she works as a lecturer under the Department of Mechanical and Mechatronics, Faculty of Technology and Built Environment, UCSI Kuala Lumpur. Her research interests include design and robotics.

Nor Fazilah Abdullah

She received her Bachelor's degree in aerospace engineering (Hons) from IIUM, Gombak in 2010 and her Master's degree in Mechanical Engineering from UKM, Bangi in 2015. Currently she is pursuing Doctoral of Philosophy programme at UCSI University, Kuala Lumpur. Her research interest is in bio-based nanoparticle materials.

A Design and Fabrication of a Solar Agriculture Water Pumping System

Ammar A. M. Al-Talib

Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment,
UCSI University, 56000 Kuala Lumpur, Malaysia

Noor Idayu Mohd Tahir

Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment,
UCSI University, 56000 Kuala Lumpur, Malaysia

Ain Atiqa Mustapha

Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment,
UCSI University, 56000 Kuala Lumpur, Malaysia

Amar Ridzuan Abd Hamid

Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment,
UCSI University, 56000 Kuala Lumpur, Malaysia

Sarah 'Atifah Saruchi

Faculty of Manufacturing and Mechatronic Engineering Technology,
Universiti Malaysia Pahang Al-Sultan Abdullah, 26600 Pekan, Pahang, Malaysia

Yazan Shebli Ibrahim

Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment,
UCSI University, 56000 Kuala Lumpur, Malaysia
E-mail: ammart@ucsiuniversity.edu.my, nooridayu@ucsiuniversity.edu.my, ainatiqa@ucsiuniversity.edu.my,
amar@ucsiuniversity.edu.my, sarahatifah@umpsa.edu.my

Abstract

This study explores the use of a solar driven water pump. PV technology replaces conventional electricity and diesel pumps by using solar energy to power DC or AC water pumps. The main objective of this study is to design and construct a solar-powered agriculture water pumping system and to evaluate its performance. The solar agriculture water pumping system used in this project consists of a 40-watt monocrystalline solar cell with an efficiency conversion of between 23% – 24%. This can supply power to a 16.8 W DC Flow Submersible Pump. It could lift the water up to 5m and a flow range of 700 Liter/Hour. The system also includes a PWM 30A Solar Charge Controller to regulate the input power to a 12V, 7.2A Sealed Rechargeable Battery. Finally, a 20m long watering kit with nozzles irrigation system is connected to a 12V DC Submersible Pump to water the plants.

Keywords: Solar panel, water pumped, renewable energy, irrigation system, SPVWPS.

1. Introduction

Electricity is an essential element of nature and a widely utilized form of energy that is vital for various human activities, such as lighting and work. With the ever-increasing global population, energy consumption has surged significantly in recent times. From 2005 to 2014, global energy consumption increased by 18%, and it is expected to grow by 35% by 2035 [1]. Renewable energy sources, including biofuel, waste, hydro, solar, wind, geothermal, and thermal, Comprise merely 13% of global energy consumption. Fossil fuels, such as natural gas, coal, and oil, make up 81%, while nuclear power contributes only 5.7% [2]. Sustainable energy, such as water, sunlight, and wind, is a solution that has no harmful impact on the environment and contributes to achieving sustainable energy goals [3]. The sun is the most significant power source on Earth, providing more energy in one hour than all nations produce in a year. It delivers over 15,000 EJ of energy to the planet every day, which is more than 104 times the daily energy used by human activities [4]. However, only 0.1% of the sun's energy is consumed for electricity [5]. Alexandre-Edmond Becquerel observed the PV effect for the first time in 1839 [6]. The initial contemporary silicon solar

Ammar A. M. Al-Talib, Noor Idayu Mohd Tahir, Ain Atiqa Mustapha, Amar Ridzuan Abd Hamid, , Sarah 'Atifah Saruchi, Yazan Shebli Ibrahim

cell was invented by Russel Ohl in 1946 [7]. Despite technological advancements, solar power installation costs five times more than coal, gas, or nuclear sources for electricity generation [8]. Agriculture is one of the important sectors in emerging nations. According to a World Bank estimate, rural agriculture areas employ nearly 86 percent of the global population (World Bank,2008). High water consumption coincides with periods of high solar irradiation. As a result, PV water pumping can help to alleviate the water shortage problem [9]. Climate change and population growth cause droughts and agricultural losses, prompting farmers to adopt solar-powered irrigation systems as an alternative energy source [10]. The world is facing a crisis in water and energy, both of which are critical for agricultural productivity. To cater to the needs of an expanding populace, a sustainable solution is necessary. The employment of solar photovoltaic water pumping systems is a feasible and economical substitute, especially in distant and undeveloped regions of developing nations. With continual innovation, these systems have gained widespread use in the industrial, residential, and agricultural sectors [11]. Solar Powered Photovoltaic Water Pumping Systems (SPVWPS) offers an alternative to traditional water pumping systems that rely on non-renewable sources like diesel, coal, and gas for electricity. Diesel systems require a lot of fuel, cause noise pollution, and result in environmental problems such as acid rain and greenhouse emissions [12]. Replacing diesel pumps with solar PV pumps costs 2-4 times as much. SPVWPS have eco-friendly operations and low maintenance costs [13]. Solar water pumping systems are in the nascent phase of development and require addressing several obstacles, including intermittent performance, steep upfront expenses, low efficiency of 15%-16%, and weather conditions that affect sun availability. Sun availability is not constant, and it varies from day to day, making it difficult to rely solely on solar power [14]. The performance of SPVWPS relies on factors such as the number of solar panels used, the type of controllers employed, and the energy storage system utilized. The pump, PV array adjustment, and ambient conditions also impact system efficiency. The maximum conversion efficiency of solar panels is 18%, and controllers like MPP tracking and charge controllers can increase efficiency and battery life. As SPVWPS only generates electricity during daylight, energy storage systems are used [15].

2. Methodology and Experimental Setup

The experimental setup comprises a 40-watt monocrystalline solar panel that generates a 12V DC output and 1.5A current. This panel is linked to a 30A PWM solar charge controller featuring an LCD controller and dual USB, which regulates the charging of 12V sealed lead-acid batteries with a 7.2AH capacity connected to the system. In addition, the charge controller

is connected to a 12V DC submersible pump for water flow. This pump has a power range of 16.8W to 26.4W, a Nylon PA66 shell is placed in a 20L water tank. It can deliver a flow range of 700 to 800L/H and is programmed to operate at specific times via a timer. The pump is connected to a 20m water pipe and is part of a watering kit that includes spray nozzles, and tee connectors, designed for plant irrigations. Fig. 1 depicts the overall prototype.

Fig. 1 Solar agriculture water pumping system prototype

2.1. Design of solar agriculture water pumping system

The design of the system is presented in Fig. 2 SolidWorks has been employed to produce a detailed representation of the design. The external frame of the system is constructed using wood. Additionally, two sheets of plywood, each measuring 40cm by 45cm, are used as supporting pieces. One of these pieces accommodates the solar panels, while the other houses the battery, charge controller, and timer. A submersible pump located inside the water tank and the tank itself are other essential components of the system.
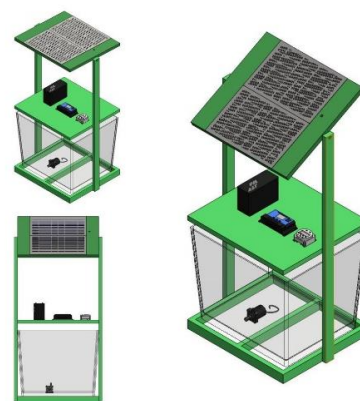
Fig. 2 Different views of the prototype

The overall connection is illustrated as a block diagram in Fig. 3. To acquire an accurate experimental result and compare it to the theoretical outcome.
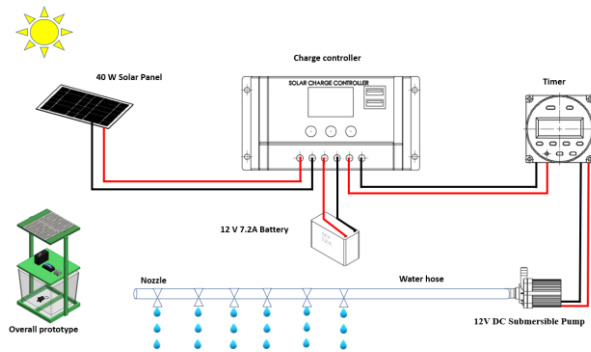
Fig. 3 Overall connection of the prototype.

3. Results and Discussion

In this section, Table 1 presents the average results for the three different solar panel orientation angles, with two different direct solar radiation of 854.3 W/m2 and 542.7 W/m2. Table 2 presents the average results of charging duration, solar panel voltage input, solar panel current input, and solar irradiations when the solar panel of three different tilt angles was placed directly under the sun. Additionally, Fig. 3 is displaying the average results of voltage drops, the discharge rate of water under two different conditions of discharge (nozzles connected to hoses of different lengths and free discharge), and the time required for the battery to be fully discharged. The readings shown in Table 3 are taken with a constant water volume of one litre per minute.

Table 1. The average readings throughout different tilt angles at Indoor two constant solar flux values.

Angle (°)	Direct solar radiation of 854.3 W/m^2				Direct solar radiation of 542.7 W/m^2			
	Charging Duration (Hour)	Solar Panel Voltage Input (V)	Solar Panel Current Input (A)	Light Intensity (W/m^2)	Charging Duration (Hour)	Solar Panel Voltage Input (V)	Solar Panel Current Input (A)	Light Intensity (W/m^2)
16	10	12.64	0.72	693.2	20	12.28	0.36	367.7
25	11.8	12.36	0.61	585.7	27.7	12.17	0.26	340.9
45	15.6	12.31	0.46	508.1	36.05	12.06	0.2	303.4

Table 2 .The average outdoor readings through different tilts angle of solar panel

Angle	Charging Duration (Hour)	Solar Panel Voltage Input (V)	Solar Panel Current Input (A)	Solar Irradiances (W/m^2)
16°	5.76	12.68	1.25	961.67
25°	6.79	12.64	1.06	896.3
45°	7.58	12.44	0.95	783.5

Table 3. Overall average results for hoses of different lengths

Tube length	Voltage output (V)	The discharge rate of water using a nozzles (L/min)	The discharge rate of water without a nozzle (L/min)	battery fully discharge (hours)
5m	13	0.866	1.44	3.5
10m	12.98	0.636	1.10	3.28
15m	12.94	0.566	0.75	3.03
20m	12.9	0.458	0.625	2.47

3.1. *Light Intensity*

Fig. 4 illustrates the light intensity recorded by the solar panel at different tilt angles. As the tilt angle increases, the amount of light received by the panel decreases significantly, particularly when the light source is positioned further away from the solar panel. This reduced light exposure results in a diminished input of voltage and current, ultimately leading to a lower power output generated by the solar panel.
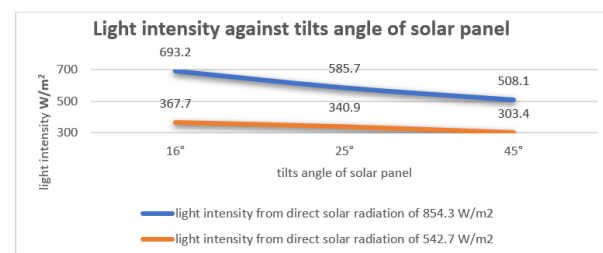
Fig. 4 Light intensity vs different tilt angles at two different indoor direct solar radiations

3.2. *Solar Irradiations*

Fig. 5 illustrates the solar panel's received solar irradiance at various tilt angles when placed directly under the sun.
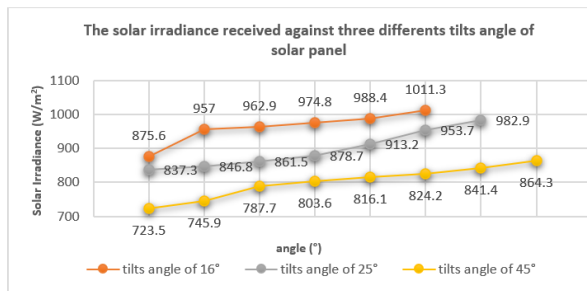
Fig. 5 Solar irradiance received at various tilt angles

Furthermore, it is observed that the solar panel receives the most radiation when placed at a smaller tilts angle, which leads to the generation of more power this can be seen in Fig. 6. Hence, it can be concluded that the tilts angle of the solar panel has a significant impact on the amount of radiation received and the subsequent power generated.
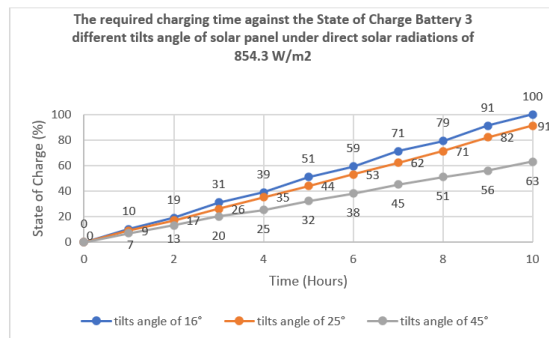
Fig. 6 Time required to charge battery against the state of charge at direct solar radiation of 854.3 W/m2

3.3. *Time required to charge the battery.*

In Fig. 6, the charging duration of the battery for three different tilts angles is shown as up to 10 hours when the height difference between the solar panel and the 1000-watt halogen bulb is 50 cm. Similarly, Fig. 7 illustrates the charging duration of the battery for the same tilt angles, but up to 20 hours and when the height difference is 70 cm. Furthermore, in Fig. 8, the duration to charge the battery for six hours is exhibited when it is situated directly under sunlight.
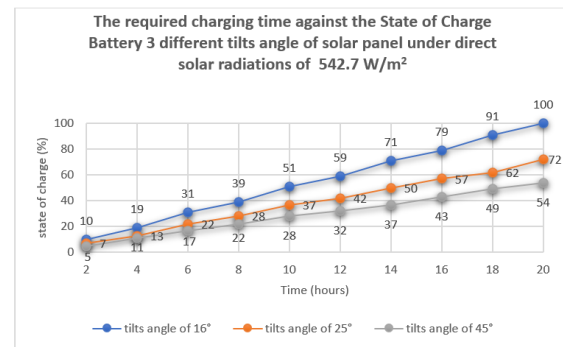
Fig 1. Time required to charge battery against the state of charge at direct solar radiation of 542.7 W/m^2
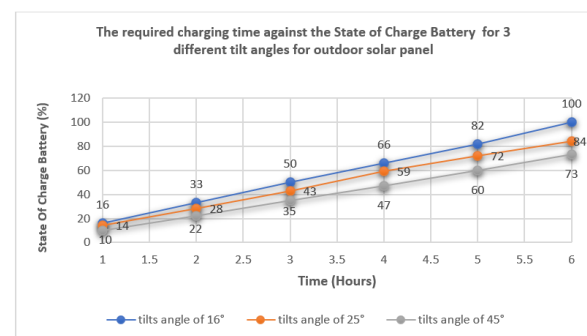
Fig 8. Time required to charge battery against the state of charge under outdoor sun light

Fig. 6 and Fig. 7 illustrate that the differences in the time required to charge the battery can be attributed to the amount of current generated by the solar panel, which is related to the amount of radiation received by the solar panel.

3.4. *The discharge rate of water*

Fig. 9 illustrates the collective findings of the water discharge experiment conducted with hoses of varying lengths and nozzle openings. The experiment involved the flow of one litre of water per minute from a 20L water tank. The presented results are based on the different percentages of nozzle openings tested.
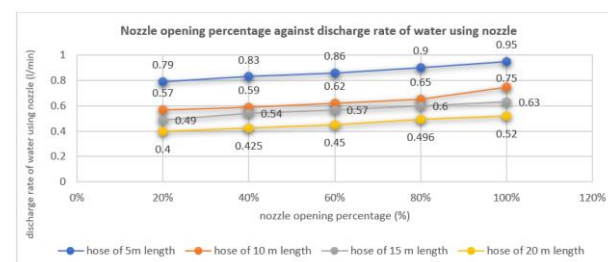
Fig 9. Discharge rate of water against nozzles of different percentage opening

According to Fig. 9, the highest water discharge rate achieved is approximately 0.95 L/min, which occurred when a 5m long hose is connected to the nozzle while maintaining a 100% opening. Conversely, the slowest water discharge rate of 0.4 L/min is observed when the hose is 20m long, and the nozzle opening percentage is 20%, unlike the hoses of 10m and 15m lengths. It is worth noting that friction between the water and the hose's inner surface, as well as changes in elevation or flow direction, contribute to losses in water pressure, known as head loss, during water pumping through a hose. Consequently, it can be inferred that hoses with smaller nozzle opening percentages offer more resistance to water flow, which leads to reduced water pressure and discharge rates.

3.5. *Maximum height water can reach*

Fig.10 presents the findings obtained by using hoses of different lengths which are 0.5m, 1m, 1.5m, 2.0m, 2.5m, and 3m to determine the maximum height that water can reach using a 12v DC submersible pump.
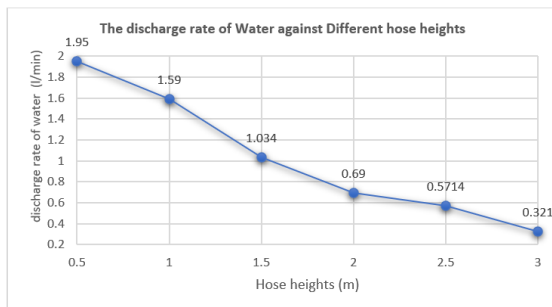
Fig. 2 The discharge rate of water against hoses of different heights

The graph shows that the pump can lift water to a maximum height of 3m with a discharge rate of approximately 0.321 l/min, while the minimum height of 0.5m has the fastest discharge rate of water at around 1.95 l/min. This indicates that the pump's lifting ability decreases as the height of the hose increases, resulting in a decrease in the water discharge rate. This is due to the longer hoses creating more resistance and reducing the flow rate.

4. Conclusion

The designed prototype for a solar irrigation system has been fabricated and tested. The study indicates that more solar radiation will result in higher charging power, which is useful for running the solar irrigation system. Additionally, the system is capable of pumping water up to 20 meters by fully utilizing solar energy.

To maximize the efficiency of the system, it has been found that the solar panel should be tilted at an appropriate angle for each location, which enables the collection of more solar energy and facilitates faster battery charging. Moreover, the system is capable of pumping water up to a height of 3 meters.

To maintain long battery life, it has been found that shorter irrigation pipes are more suitable. Furthermore, the study revealed that the solar collector efficiency is 80%. Overall, the designed solar irrigation system has demonstrated its capability in achieving its intended objectives, and the findings of the study have provided valuable insights for future system improvements.

Acknowledgements

References

1. E. Pérez-Denicia, F. Fernández-Luqueño, D. Vilariño-Ayala, L. Manuel Montaño-Zetina, and L. Alfonso Maldonado-López, "Renewable energy sources for electricity generation in Mexico: A review," Renewable and Sustainable Energy Reviews, vol. 78, pp. 597–613, Oct. 2017, doi: 10.1016/J.RSER.2017.05.009.
2. M. Hosenuzzaman, N. A. Rahim, J. Selvaraj, M. Hasanuzzaman, A. B. M. A. Malek, and A. Nahar, "Global prospects, progress, policies, and environmental impact of solar photovoltaic power generation," Renewable and Sustainable Energy Reviews, vol. 41, pp. 284–297, Jan. 2015, doi: 10.1016/J.RSER.2014.08.046.
3. A. Q. Al-Shetwi, "Sustainable development of renewable energy integrated power sector: Trends, environmental impacts, and recent challenges," Science of The Total Environment, vol. 822, p. 153645, May 2022, doi: 10.1016/J.SCITOTENV.2022.153645
4. S. Bilgen, "Structure and environmental impact of global energy consumption," Renewable and Sustainable Energy Reviews, vol. 38, pp. 890–902, Oct. 2014, doi: 10.1016/J.RSER.2014.07.004.
5. C. J. Chen, "Physics of Solar Energy," Physics of Solar Energy, Aug. 2011, doi: 10.1002/9781118172841.
6. L. M. Fraas, "History of Solar Cell Development," Low-Cost Solar Electric Power, pp. 1–12, 2014, doi: 10.1007/978-3-319-07530-3_1.
7. S. Sharma, K. K. Jain, and A. Sharma, "Solar Cells: In Research and Applications—A Review," Materials Sciences and Applications, vol. 06, no. 12, pp. 1145–1155, 2015, doi: 10.4236/msa.2015.612113.
8. N. Babakatcha, A. Yabagi, M. B. Ladan, and M. D. Oladipupo, "HARNESSING SOLAR ENERGY POTENTIAL AS AN ALTERNATIVE SOURCE OF ELECTRICAL ENERGY IN NORTH CENTRAL, NIGERIA," African Journal of Environment and Natural Science Research, vol. 3, p. 86, 2020, Accessed: Mar. 28, 2023. [Online]. Available: www.abjournals.org
9. D. Zhou and Abdullah, "The acceptance of solar water pump technology among rural farmers of northern Pakistan: A structural equation model," Cogent Food Agric, vol. 3, no. 1, 2017, doi:

10.1080/23311932.2017.1280882/SUPPL_FILE/OAFA_A_1280882_SM1744.DOCX.

10. J. B. Rutibabara and W. M. Budzianowski, "Environmental and Economic Cost Analysis of a Solar PV, Diesel and hybrid PV-Diesel water Pumping Systems for Agricultural Irrigation in Rwanda: Case study of Bugesera district INSTITUTE OF WATER AND ENERGY SCIENCES (Including climate change)," 2018.

11. G. Li, Y. Jin, M. W. Akram, and X. Chen, "Research and current status of the solar photovoltaic water pumping system – A review," Renewable and Sustainable Energy Reviews, vol. 79. Elsevier Ltd, pp. 440–458, 2017. doi: 10.1016/j.rser.2017.05.055.

12. A. Mokeddem, A. Midoun, D. Kadri, S. Hiadsi, and I. A. Raja, "Performance of a directly-coupled PV water pumping system," Energy Convers Manag, vol. 52, no. 10, pp. 3089–3095, Sep. 2011, doi: 10.1016/J.ENCONMAN.2011.04.024.

13. S. S. Chandel, M. Nagaraju Naik, and R. Chandel, "Review of solar photovoltaic water pumping system technology for irrigation and community drinking water supplies," Renewable and Sustainable Energy Reviews, vol. 49, pp. 1084–1099, Oct. 2015, doi: 10.1016/J.RSER.2015.04.083.

14. B. Singh, U. Sharma, and S. Kumar, "Standalone Photovoltaic Water Pumping System Using Induction Motor Drive with Reduced Sensors," IEEE Trans Ind Appl, vol. 54, no. 4, pp. 3645–3655, Jul. 2018, doi: 10.1109/TIA.2018.2825285.

15. R. Foster, M. Ghassemi, and A. Cota, "Solar energy: Renewable energy and the environment," Solar Energy: Renewable Energy and the Environment, pp. 1–337, Jan. 2009, doi: 10.5860/choice.47-5672.

Authors Introduction

Ammar Abdulaziz Al Talib

He received his B.Sc and M.Sc degrees in Mechanical Engineering from the University of Mosul Iraq.
He has finished his Ph.D degree from UPM University, Malaysia. Member of the Institute of Mechanical Engineers UK. (CEng. MIMechE). He has developed all the Postgraduate Programs at the Faculty of Engineering at UCSI University / Malaysia and worked as the Head of Postgraduate and Research department at the same faculty for the years 2010-2018.

Noor Idayu Mohd Tahir

She received her B.Eng in Mechanical Engineering from Universiti Teknologi MARA, Malaysia. She was design engineer in oil and gas industries for 7 years. Currently, she works as a lecturer under the Department of Mechanical and Mechatronics, Faculty of Technology and Built Environment, UCSI Kuala Lumpur. Her research interests include design and autonomous robotics.

Ain Atiqa Mustapha

She received her Master's degree from the Faculty of Electronic & Computer Engineering, Universiti Teknikal Malaysia Melaka, Malaysia in 2017. She is currently a Tutor under Department of Mechanical and Mechatronics Engineering, UCSI University, Kuala Lumpur. Her field of research is renewable energy and artificial intelligence.

Amar Ridzuan Abd.Hamid

He received his Bachelor's degree in automotive engineering (Hons) from UTeM, Melaka in 2011 and his Master's degree in Mechanical Engineering from UPM, Selangor in 2015. He is a Professional Technologist, Head of Programmes and lecturer of Mechanical and Mechatronic programmes from Department of Mechanical Engineering, UCSI University, Malaysia. His research interests are in the design and analysis of automotive compartments.

Sarah 'Atifah Saruchi

She received her B.Eng. in Mechanical and Aerospace Engineering from Nagoya University, Japan . She received her Master and Doctoral degrees from Malaysia-Japan International Institute of Technology (MJIIT), Universiti Teknologi Malaysia. Currently, she is working at Universiti Malaysia Pahang Al-Sultan Abdullah, Pahang, Malaysia. Her research interests include mechatronics and artificial intelligence.

Yazan Shebli Ibrahim

He has graduated form the Mechanical Engineering Department at UCSI University as a first class graduand in 2023 . Currently, he is working in the private sector.

Design and Performance of a Power Generating Manual Treadmill

Ammar A.M. Al-Talib

Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment,
UCSI University, 56000 Kuala Lumpur, Malaysia

Sarah 'Atifah Saruchi

Faculty of Manufacturing and Mechatronic Engineering Technology,
Universiti Malaysia Pahang Al-Sultan Abdullah, 26600 Pekan, Pahang, Malaysia

Cik Suhana Hassan

Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment,
UCSI University, 56000 Kuala Lumpur, Malaysia

Nor Fazilah Binti Abdullah

Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment,
UCSI University, 56000 Kuala Lumpur, Malaysia

Ain Atiqa Mustapha

Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment,
UCSI University, 56000 Kuala Lumpur, Malaysia

Ahmad Jelban

Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment,
UCSI University, 56000 Kuala Lumpur, Malaysia

E-mail: ammart@ucsiuniversity.edu.my, sarahatifah@umpsa.edu.my, suhana@ucsiuniversity.edu.my, ,
norfa@ucsiuniversity.edu.my, ainatiqa@ucsiuniversity.edu.my

Abstract

Treadmills are one of the most popular training equipment in the gym and at home. The working principle of treadmills is by moving the belt with the human knee bending, which creates mechanical energy to turn the belt. A gear or pulley and belt system connects to the generator along the axel line of the rolling bars. The power generated by the DC generator is stored in a battery pack and could be used to charge phones or other equipment. It has been found that treadmills can provide an efficiency of 95% when the DC motor is used and 92% when the AC motor is used. The main objective of this study is to design and fabricate a powder-generating manual treadmill and to analyze the performance of the system under different operation conditions.

Keywords: Manual Treadmills, Gym Renewable Energy

1. Introduction

In 2015, the United Nations General Assembly established the Sustainable Development Goals (SDGs), which consist of 17 objectives to be accomplished by 2030. SDG 7 aims to provide affordable, reliable, sustainable, and modern energy for all [1] This goal is especially relevant during the COVID-19 pandemic, as measures such as movement control orders (MCO) have significantly increased residential energy consumption while commercial and business energy consumption has

decreased [2]. Impact of the MCO on electricity consumption has shown a sharp increase in residential electricity consumption [3]. To address this issue, the treadmill, which has become a popular fitness equipment for both home and gym use, can be modified to generate electricity. Originally the treadmill consumed energy to operate [4]. The treadmill consists of a wide belt driven by a variable-speed motor. However, by connecting it to a generator or an inverter, the kinetic energy generated by a person walking or running on the belt can be converted into electrical energy. This energy can then be used to power various devices or fed back into the power grid,

Ammar A. M. Al-Talib, Sarah 'Atifah Saruchi, Cik Suhana Hassan, Nor Fazilah Abdullah, Ain Atiqa Mustapha, Ahmad Jelban

making the treadmill a potential solution to the increased energy consumption during the pandemic [5].

2. Methodology and Experimental Setup

As shown in Fig. 1, this experiment is set up with a 36 V DC generator that is driven by a pulley linked by a belt, with the main motion generated by the user walking or running on the treadmill belt [6]. The generator is connected to a charging control that regulates the voltage and current supplied to a 12 V, 7.2 Ah battery. The battery then supplies a 200 W inverter, which converts DC to AC power that can be used to charge devices such as phones and laptops. A Watt meter is connected between the generator and the charging controller to measure the current and voltage produced by the generator.
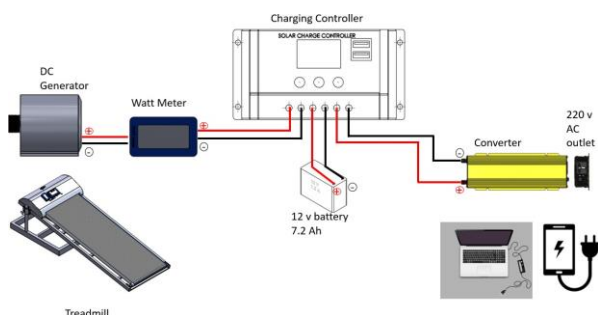
Fig. 1 Connection Illustration of The Setup

2.1. Design and Modifications

As the treadmill was already made, modifications are required to convert the regular treadmill to manual one that is able to generate electricity. The modifications can be summarized as an assembly of a jack to have an incline, support for the side, generator base, motor pulley and treadmill roller. Fig. 2 shows the treadmill before modifications.

Fig. 2 Treadmill Before the Modifications

Fig. 3 shows the modified treadmill at its charging status.

Fig. 3 Charging Laptop from The Treadmill Battery

3. Results and Discussion

In this section, Table 1 shows the average results of 3 different people who tested the prototype. The results are the readings from the watt meter for the voltage and current while the average speed is from the speedometer after 5 minutes of walking on the treadmill. As stated, before the treadmill is not motorized, therefore the speed is not constant, and the only way to maintain a constant speed is dependent on the person himself. However, the adjustment of pace was constant to maintain a relatively constant speed.

Table 1. Average Speed for The Three Testers

	Person 1	Person2	Person 3
Speed (km/hr)	2.24	3.12	5.05
Current (A)	0.56	0.86	1.19
Voltage (v)	12.18	12.53	13.24
Power (W)	6.81	10.63	15.43

3.1. Discussion

Fig. 4 shows the average result for the three-persons. The results are combined in one graph, in terms of speed (km/hr), current (A), Voltage, and power (W).
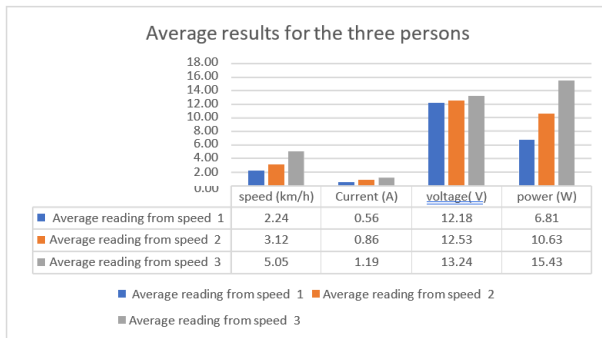
Fig. 4. Average Results for The Three Persons

From Fig. 4, it can be noticed that the higher the speed the higher the voltage and current, and accordingly the power generated. The data represents the average speed that has been achieved by each user. The voltage and current generated by the lowest speed of 2.24 km/h are 12.18 v and 0.56A respectively. This speed can be achieved by anyone as it is within the normal walking speed. Speed 2 of 3.2 km/h induced a voltage of 12.53 v and a current of 0.86 A. The results are the average of the 3 trials by each person. And lastly, speed 3 which is the average of the highest speed that can be achieved at 5.05 km/h induced a voltage of 13.24 v and current of 1.19 A. Taking into consideration that the maximum current to charge the battery as recommended by the manufacturer is 0.1 to 0.25 of the battery capacity. The current achieved falls within these values. Based on the results, it is evident that there is a direct relationship between the speed, voltage, current, and power output of the manual Treadmill. When the speed of the person increases, the voltage and current generated also increase, resulting in higher power output.

3.2. Theoretical Current

From the generator specifications the RPM rating is 3400, while the current is 9.5 A for the 36 v. the generator efficiency is stated to be 78%. Eq. (1) is used to calculate the current produced by 1 RPM.

$$Current = \frac{I}{RPM} \tag{1}$$

The treadmill belt is connected to the roller pulley with a ratio of 1:20 which is measured experimentally. The ratio between the roller pulley and the generator pulley can be calculated using Eq. (2) to Eq. (5):

$$\frac{D_r}{D_m} = \frac{95}{35} = 2.7 \tag{2}$$

$$Roller\ Pulley\ RPM = \frac{speed(m/s)}{\pi D} \tag{3}$$

$$For\ Motor\ RPM = Roller\ Pulley\ RPM \times 2.7 \tag{4}$$

$$For\ Current\ generated\ current = current\ (\frac{A}{RPM}) \times Motor\ RPM \tag{5}$$

From Eq. (1) for each rotation of the generator, a current of 0.0029 A will be produced. The motor efficiency is 78% as stated by the manufacturer, which will be considered when calculating the current. The average speed in Table 2 will be considered for these calculations.

Table 2. Average Speed Current Calculation

	Speed 1	Speed 2	Speed 3
Speed (km/hr)	2.24	3.12	5.05
Speed (m/s)	37.28	52.0	84.11
Roller (RPM)	125.0	174.14	282.0
Motor (RPM)	337.41	470.16	761.31
Current (A)	0.91	1.27	2.06
Current with 78% efficiency	0.71	0.99	1.61

Eq. (6) is used to calculate the efficiency for actual current and theoretical current.

$$current\ efficiency =$$
$$= (1 - \frac{theoretical\ current - actual\ current}{theoretical\ current}) \times 100 \tag{6}$$

Table 3. Actual Current and Theoretical Current Efficiency

Actual Current (A)	0.56	0.87	1.19
Current with efficiency	0.71	0.99	1.61
Efficiency	78.88%	87.88%	73.91%

From Table 3, it can be noticed that the produced current is less than the theoretical current and that is due to the losses due to friction during the experiments.

4. Conclusion

In conclusion, the prototype of the manual treadmill is able to operate and generate electricity of different values depending on the walking/ running speed of the user as it is directly connected. The results have shown that the faster the walking on the treadmill, the higher the current generated which is required to charge a battery for later use. The average 3 speeds are 2.24 km/h, 3.12 km/h, and

Ammar A. M. Al-Talib, Sarah 'Atifah Saruchi, Cik Suhana Hassan, Nor Fazilah Abdullah, Ain Atiqa Mustapha, Ahmad Jelban

5.05 km/h which have generated a current of 0.56A, 0.86A, and 1.19 A respectively. While the voltage for the 3 speeds is 12.18 v 12.53 v and 13.24 v. For the application of the manual treadmill, 3 devices with different battery capacities have been tested. The devices are a smartphone, a smartwatch, and a laptop, with batteries capacities of 4200 mAh, 455 mAh, and 56 Wh respectively. After testing using USB and a normal charger it has been found that the 7.2 Ah battery can charge the smartform 4.8 times, the smartwatch 44.6 times, and the laptop 1.5 times. However, it is important to note that the manual treadmill is designed to be an affordable and accessible option for exercise beside being successful in generating electrical power. The findings of this research align with target 13.2 of the SDG's, which emphasizes the integration of climate change measures into policies and planning.

Acknowledgements

References

1. M. Yu, J. Kubiczek, K. Ding, A. Jahanzeb, and N. Iqbal, "Revisiting SDG-7 under energy efficiency vision 2050: the role of new economic models and mass digitalization in OECD," Energy Effic, vol. 15, no. 1, p. 2, Jan. 2022, doi: 10.1007/s12053-021-10010-z.
2. S. I. Mustapa, R. Rasiah, A. H. Jaaffar, A. Abu Bakar, and Z. K. Kaman, "Implications of COVID- 19 pandemic for energy-use and energy saving household electrical appliances consumption behaviour in Malaysia," Energy Strategy Reviews, vol. 38, p. 100765, Nov. 2021, doi: 10.1016/j.esr.2021.100765
3. Md. T. Ullah, Md. A. Bin Karim, M. H. Uddin, and G. M. Tauseef, "Harvesting green energy from wastage energy of human activities using gymnasium bicycle at Chittagong city," in 2015 3rd International Conference on Green Energy and Technology (ICGET), IEEE, Sep. 2015, pp. 1–4. doi: 10.1109/ICGET.2015.7315085.
4. S. Kolgiri, "'Design & Fabrication of Treadmill Bicycle' Synopsis for B.E. (Mechanical Engineering) Project Analysis Ergonomics aspects of Powerloom Industry View project STATIC AND DYNAMIC ANALYSIS FOR ROTOR SHAFT OF ELECTRIC MOTOR " View project." [Online]. Available: https://www.researchgate.net/publication/331950519
5. Sahil, P. K. Sharma, N. Hari, N. Kumar, and D. Shahi, "An innovative technique of electricity generation and washing machine application using treadmill," in 2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES), IEEE, Jul. 2016, pp. 1–5. doi: 10.1109/ICPEICES.2016.7853524.
6. R.Harsha, "DESIGN AND FABRICATION OF TREADMILL BICYCLE," Chennai, T.Nagar, 2018. [Online]. Available: www.jetir.org
7. Mustapa, S. I., Rasiah, R., Jaaffar, A. H., Abu Bakar, A., & Kaman, Z. K. (2021). Implications of COVID-19 pandemic for energy-use and energy saving household electrical appliances consumption behaviour in Malaysia. Energy Strategy Reviews, 38, 100765. https://doi.org/10.1016/j.esr.2021.100765

Authors Introduction

Ammar Abdulaziz Al Talib

He received his B.Sc and M.Sc degrees in Mechanical Engineering from the University of Mosul Iraq.
He has finished his Ph.D degree from UPM University, Malaysia. Member of the Institute of Mechanical Engineers UK. (CEng. MIMechE). He has developed all the Postgraduate Programs at the Faculty of Engineering at UCSI University / Malaysia and worked as the Head of Postgraduate and Research department at the same faculty for the years 2010-2018.

Sarah 'Atifah Saruchi

She received her B.Eng. in Mechanical and Aerospace Engineering from Nagoya University, Japan She received her Master and Doctoral degrees from Malaysia-Japan International Institute of Technology (MJIIT), Universiti Teknologi Malaysia. Currently, she is working at Universiti Malaysia Pahang Al-Sultan Abdullah, Pahang, Malaysia. Her research interests include mechatronics and artificial intelligence.

Cik Suhana Hassan

She received her bachelor's and master's degrees in 2009 and 2011, respectively, from Universiti Teknologi PETRONAS, and her PhD in 2019 from Universiti Putra Malaysia. She is an Assistant Professor at the Department of Mechanical and Mechatronics Engineering of UCSI University. Her research interests include the investigation of bio-composites for use in automotive applications. She is also an active member of the materials community, having been recognized as a Professional Member of the Institutes of Materials Malaysia and a Professional Technologist of the Malaysian Board of Technologists in the Material Science Technology field .

Nor Fazilah Abdullah

She received her Bachelor's degree in aerospace engineering (Hons) from IIUM, Gombak in 2010 and her Master's degree in Mechanical Engineering from UKM, Bangi in 2015. Currently she is pursuing Doctoral of Philosophy programme at UCSI University, Kuala Lumpur. Her research interest in bio-based nanoparticles materials.

Ain Atiqa Mustapha

She received her Master's degree from the Faculty of Electronic & Computer Engineering, Universiti Teknikal Malaysia Melaka, Malaysia in 2017. She is currently a Tutor under Department of Mechanical and Mechatronics Engineering, UCSI University, Kuala Lumpur. Her field of research is renewable energy and artificial intelligence.

Ahmad Abu Jelban

He graduated from the Mechanical Engineering Department at UCSI University in 2022, and currently is working as an Engineer.

Smart Car Jack Using Internet of Things

Noor Idayu Mohd Tahir

Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment,
UCSI University, 56000 Kuala Lumpur, Malaysia

Ammar A.M. Al-Talib

Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment,
UCSI University, 56000 Kuala Lumpur, Malaysia

Zeiad Mahmoud Ebeid Abdelghaffar

Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment,
UCSI University, 56000 Kuala Lumpur, Malaysia

Cik Suhana Hassan

Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment,
UCSI University, 56000 Kuala Lumpur, Malaysia

E-mail: NoorIdayu@ucsiuniversity.edu.my, ammart@ucsiuniversity.edu.my, 1001851834@ucsiuniversity.edu.my,
suhana@ucsiuniversity.edu.my

Abstract

This research explores the integration of smart technologies in traditional car jacks, aiming to enhance efficiency and user safety. By incorporating sensors and automation, the smart car jack introduced in this study not only streamlines the lifting process but also provides real-time data on load distribution and stability. ethodologically, the study involves the design, prototyping, and testing of the smart car jack, assessing its performance in various scenarios. Results demonstrate a significant reduction in lifting time and improved safety measures, positioning the smart car jack as a viable innovation in the automotive industry. The implications of this technology extend to increased convenience for users and a potential reduction in roadside accidents. This research contributes to the ongoing evolution of automotive tools and underscores the benefits of merging smart technologies with traditional equipment.

Keywords: Internet of Things; Car Jack, Smart, ESP32

1. Introduction

This paper addresses the challenges of roadside emergencies, such as flat tires, and the limitations of conventional scissor car jacks that require physical effort and can lead to backache problems. To tackle these issues, the paper presents a modified car jack design that is safe, easy to operate, and reduces physical effort for lifting and lowering vehicles during automobile repair, especially for roadside situations. The study explores the integration of Internet of Things (IoT) technology in automotive tools, particularly in the development of smart car jacks. The aim is to revolutionize the traditional car lifting process, making it smarter, safer, and more efficient. By incorporating IoT, the proposed smart car jack offers convenience and enhanced control over vehicle maintenance tasks [1].This paper introduces a novel smart car jack design that leverages IoT technology to improve the efficiency and safety of roadside repairs, alleviating the burden on car owners during emergencies. The focus is on developing a user-friendly and cost-effective solution that can significantly reduce the physical effort required for car maintenance tasks [2].

2. Methodology and Experimental Setup

A. Software and microcontroller

In this project, software holds a pivotal position in designing and developing a smart car jack with Internet of Things (IoT) capabilities. To achieve the project's objectives, the software utilized includes the Blynk platform and the Arduino Integrated Development Environment (IDE) [3]. Blynk serves as a mobile platform, enabling remote control of various hardware modules like Arduino and ESP32 over the internet. With Blynk, developers gain access to an array of features, empowering them to control hardware components, store, and display data, visualize

Noor Idayu Mohd Tahir, Ammar A.M. Al-Talib, Zeiad Mahmoud Ebeid Abdelghaffar, Cik Suhana Hassan

information, and more. It simplifies the process of connecting input/output components with hardware devices and facilitates seamless data transfer between them. The Blynk server, functioning as a cloud-based backend service, establishes communication between the smartphone application and the hardware, making it highly convenient for developers to manage and control their IoT systems efficiently [4]. In this project, the ESP32 is utilized as the primary microcontroller. The ESP32 is a powerful and versatile microcontroller that operates using the Arduino IDE, an open-source software platform for writing and uploading code to Arduino boards.

The ESP32, integrated with the Blynk platform, allows for enhanced remote control and monitoring of the smart car jack's hardware components.
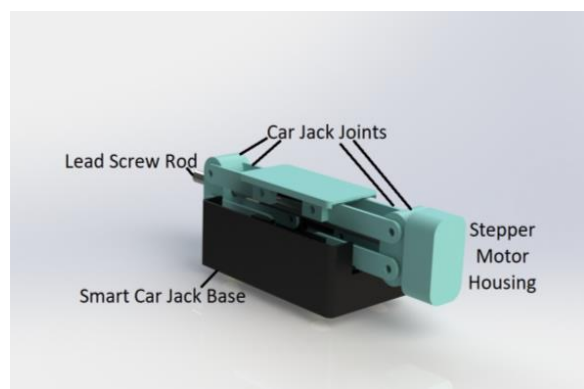
Fig. 1. Prototype Design

To establish the connection between the ESP32 and the Blynk platform, the Blynk library was installed in the Arduino IDE [5].This library provides an interface for seamless communication between the ESP32 and the Blynk cloud server and mobile application. Once the library is installed, developers can utilize the Blynk API to program the ESP32, enabling control and real-time monitoring of the smart car jack's various hardware functionalities [6].

B. Prototype Design

This section presents the final design for the smart car jack project, as depicted in Fig. 1. The design features a scissor jack prototype mounted on a movable base, allowing for increased mobility. The prototype is equipped with mecanum wheels driven by a DC gear motor, transforming it into a smart car jack. To control the lifting mechanism, a steel rod is connected to a stepper motor, which receives instructions from the microcontroller, ESP32. The components of the design, except for the steel rod, are 3D printed using robust PLA+ (Polylactic Acid) material, renowned for its improved mechanical properties like enhanced

toughness and impact resistance compared to standard PLA. This ensures the smart car jack's stability and durability during the car lifting process. The final design incorporates a movable scissor jack prototype with mecanum wheels, operated by a stepper motor controlled by the ESP32 microcontroller. The use of PLA+ material for 3D printing ensures the structure's strength and ability to withstand the pressure exerted during the lifting operation.
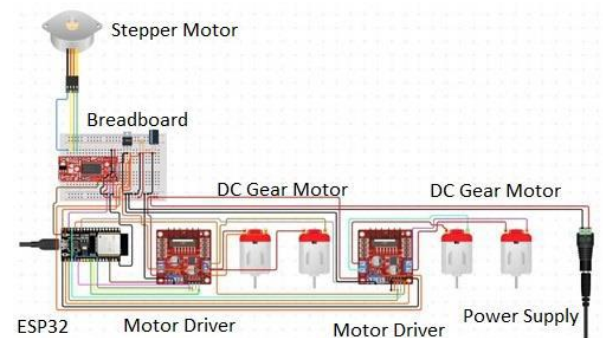
Fig. 2. System wiring diagram.

C. Coding and Programming

The code implemented in this project outlines the core functionalities and control logic of the Arduino sketch for the IoT-enabled car jack. It begins by including essential libraries for Wi-Fi communication, Blynk (the IoT platform), and Stepper motor control. The Blynk template ID, obtained from the Blynk website, along with the name and authentication token, is defined to establish communication with the Blynk app. Wi-Fi credentials are set to connect the ESP32 board to the network. In the setup function, the code connects to Wi-Fi, initializes Blynk, and sets up the motor control pins, ensuring the motors start in a stationary state. The loop function continually checks for incoming commands from the Blynk app through virtual pins. Corresponding functions are executed to control the car jack's movements or perform other desired actions. Additional functions are defined for precise motor control, including forward, backward, left, and right movements. The Stepper1 function governs the stepper motor's rotation based on specified parameters for direction and the number of rotations required [7]. This pseudocode provides a concise overview of how the IoT-enabled car jack interacts with the Blynk app, processes input commands, and controls the motors to achieve specific movements.

D. Connections

Fig. 2 illustrates the interconnected setup of the Smart Car Jack Prototype, comprising two power supplies, three motor drivers, one stepper motor, four DC gear

motors, and the ESP32 microcontroller. The motor drivers are connected to the power supply and the respective motors, with specific pins interfacing them to the ESP32 for control signals.

Regarding the stepper motor, pin connections were established using pins 25, 26, 21, and 19 as IN1, IN2, IN3, and IN4 for its motor driver. These specific pins allow precise control and coordination of the stepper motor, significantly contributing to the overall functionality and performance of the Smart Car Jack Prototype.

3. Results and discussion

Tests to analyse the performance of the smart car jack: A. Prototype efficiency on different ground surfaces. Table 1 shows the car jack performance while moving on different surfaces, The prototype demonstrated varying speeds on different surfaces. It achieved the highest average speed of 0.41 m/s on asphalt, performed moderately on grass with an average speed of 0.23 m/s, but encountered challenges on sand, getting stuck in all trials and unable to complete the one-meter test. Surface conditions significantly affected its mobility and effectiveness [8].

Table 1. Test Results

Surface	Time taken (s)	Average speed (m/s)
Asphalt	2.46	0.41
Grass	4.32	0.23
Sand	Failed to finish	-

B. Analysing the performance of the car jack with no load.

The data of the car jack with no load has been tabulated in Table 2 and illustrated in Fig. 3 which shows that the car jack's consistent and efficient performance in both jacking up and jacking down processes without any load. The mechanism is well-calibrated and operates smoothly, ensuring reliable and convenient operations for users during routine tasks [9].

Table 2. Test Results

Trial Number	Time taken (s)	
	Jack Up	Jack Down
1	14.23	13.34
2	13.58	12.56
3	13.45	13.13
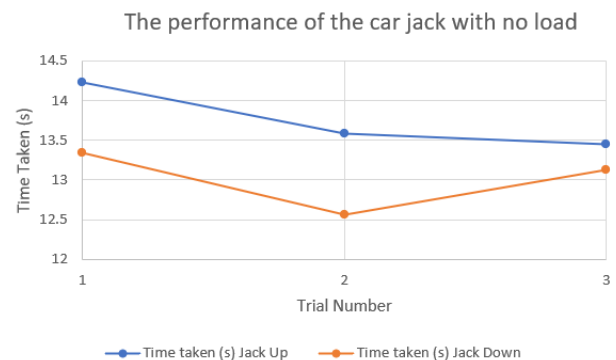
Fig. 3. The performance of the car jack with no load

C. Analysing the performance of the car jack with different loads.

The results reveal that the jacking up time increases with the load on the jack, indicating a direct relationship between them. Conversely, the jacking down time decreases as the weight applied to the prototype increases, showing an inverse relationship. These results provide valuable insights for optimizing the prototype's performance under varying load conditions, as illustrated in Table 3.

Table 3. Test Results

Weight (N)	Time taken (s)	
	Jack Up	Jack Down
150	23.58	11.54
200	27.34	10.28
250	31.27	9.56

D. Determining the maximum height of the prototype.

This test findings determine the prototype's maximum safe height as 30 cm, beyond which the stability and safety of the car lifting process could be compromised. Establishing this critical height limit provides valuable guidance for users to safely operate the smart car jack, ensuring both user and load safety during lifting operations. Table 4 shows the stability of the car jack with different heights.

Table 4. Test Results

Height (cm)	Observations
15	Completely Stable
25	Stable
30	Stable
35	Risky
40	Not Stable

Noor Idayu Mohd Tahir, Ammar A.M. Al-Talib, Zeiad Mahmoud Ebeid Abdelghaffar, Cik Suhana Hassan

E. Comparison between the jacking up speed of manual scissor car jack and the proposed prototype

The smart car jack prototype reaches its maximum height significantly faster than the manual scissor car jack, thanks to the use of a stepper motor that ensures precise and consistent movements. The elimination of human factors during operation contributes to enhanced efficiency and reliability, making the smart car jack an appealing choice for users seeking faster and more reliable lifting operations, as shown in Table 5.

Table 5. Test Results

Trial Number	Result to accomplish the task jack up and down (s)	
	Manual Scissor Jack	Proposed prototype
1	31.27	12.51
2	27.28	13.13
3	25.46	13.47

4. Conclusion

This study aimed to develop a smart scissor car jack with IoT capabilities, incorporating mecanum wheels and a stepper motor, controlled through the Blynk application. The prototype demonstrated remarkable performance, lifting a 200 N toy car to 30 cm in just 13.13 seconds, outperforming manual alternatives. The success of this project highlights the potential of IoT-integrated automotive tools in enhancing efficiency, safety, and convenience in vehicle maintenance. Future recommendations include improving mobility with larger wheels, implementing intelligent positioning systems, and enhancing safety mechanisms through IoT sensors. These enhancements will transform the smart car jack into an advanced and user-friendly automotive solution for various terrains and vehicles.

5. References

1. Madanhire, I., Chatindo, T., & Mbohwa, C. (2019b). Development of a Portable Motorized Car Jack. http://www.edutwin.com.
2. S. R. J. Ramson, S. Vishnu, and M. Shanmugam, "Applications of Internet of Things (IoT) – An Overview," in 2020 5th International Conference on Devices, Circuits and Systems (ICDCS), IEEE, Mar. 2020, pp. 92–95. doi: 10.1109/ICDCS48716.2020.243556.
3. ASEFA AND ADUGNA, "INSTITUTION OF TECHNOLOGY DEPARTMENT OF MECHANICAL ENGIENEERING MACHINE DESIGN PROJECT TWO THE DESIGN OF ACME THREADED SCREW JACK," 2022.
4. S. A. Pekhale and S. V Karanjkar, "International Journal of Modern Trends in Engineering and Research Design Modification and Analysis of Electrically Operated Screw Jack for Light Motor Vehicles," 2016. [Online]. Available: www.ijmter.com
5. B. Ezurike and M. Okwu, "Modified Screw Jack for Lifting Operation in Industrial Setting," International Journal of Engineering and Technologies, vol. 13, pp. 39–50, Dec. 2017, doi: 10.18052/www.scipress.com/ijet.13.39.
6. Pratheep, V. G., Tamilarasi, T., Ravichandran, K., Srinivasan, M., Someswaran, N., Vimal, S., & Praveen, S. (2023). Design and fabrication of IOT based scissor jack. 020002. https://doi.org/10.1063/5.0148663
7. H. S. Purnama, T. Sutikno, S. Alavandar, and A. C. Subrata, "Intelligent Control Strategies for Tuning PID of Speed Control of DC Motor - A Review," in 2019 IEEE Conference on Energy Conversion (CENCON), IEEE, Oct. 2019, pp. 24–30. doi: 10.1109/CENCON47160.2019.8974782.
8. H. Taheri and C. X. Zhao, "Omnidirectional mobile robots, mechanisms and navigation approaches," Mech Mach Theory, vol. 153, p. 103958, Nov. 2020, doi: 10.1016/j.mechmachtheory.2020.103958.
9. M. Kelemen et al., "Distance Measurement via Using of Ultrasonic Sensor," Journal of Automation and Control, Vol. 3, 2015, Pages 71-74, vol. 3, no. 3, pp. 71–74, Dec. 2015, doi: 10.12691/AUTOMATION-3-3-6.

Authors Introduction

Noor Idayu Mohd Tahir

She graduated from University Teknologi MARA, Malaysia in Mechanical Engineering. She was design engineer in oil and gas industries for 7 years. Currently, she works as a lecturer under the Department of Mechanical and Mechatronics, Faculty of Technology and Built Environment, UCSI Kuala Lumpur. Her research interests include design and autonomous robotics.

Ammar A.M. Al-Talib

He has finished his B.Sc and M.Sc degrees in Mechanical Engineering from the University of Mosul/Iraq. He has finished his Ph.D degree from UPM University / Malaysia. He is also a Chartered Engineer and Member of the Institute of Mechanical Engineers / UK. (CEng. MIMechE). He has developed all the Postgraduate Programs at the Faculty of Engineering at UCSI University / Malaysia and worked as the Head of Postgraduate and Research department at the same faculty for the years 2010-2018.

Mr. Zeiad Mahmoud Ebeid Abdelghaffar

He received his bachelor's degree in Mechatronic Engineering in 2023 at UCSI University, Kuala Lumpur. His field of research is Internet of Thing (IoT) technology.

Cik Suhana Hassan

She received her bachelor's and master's degrees in 2009 and 2011, respectively, from Universiti Teknologi PETRONAS, and her PhD in 2019 from Universiti Putra Malaysia. She is an Assistant Professor at the Department of Mechanical and Mechatronics Engineering of UCSI University. Her research interests include the investigation of bio-composites for use in automotive applications.

Gas Detection for Biogas System Using Internet-of-Things(IoT)

Ammar A.M. Al-Talib

Department of Mechanical & Mechatronics, Faculty of Engineering Technology and Built Environment, UCSI University, 56000 Kuala Lumpur, Malaysia

Ian Hoh Wen Yang

Department of Mechanical & Mechatronics, Faculty of Engineering Technology and Built Environment, UCSI University, 56000 Kuala Lumpur, Malaysia

Noor Idayu Mohd Tahir

Department of Mechanical & Mechatronics, Faculty of Engineering Technology and Built Environment, UCSI University, 56000 Kuala Lumpur, Malaysia

Ayu Haslija Binti Abu Bakar

Department of Mechanical & Mechatronics, Faculty of Engineering Technology and Built Environment, UCSI University, 56000 Kuala Lumpur, Malaysia

Nur Muhammad Afifi bin Zainal
Tenaga National Berhad
E-mail: ammart@ucsiuniversity.edu.my, 1001955881@ucsiuniversity.edu.my, NoorIdayu@ucsiuniversity.edu.my, ayuhaslija@ucsiuniversity.edu.my, afifi.zainal@tnb.com.my

Abstract

Global warming remains one of the most detrimental by-products of industrialization. Fossil fuels contribute to the majority of greenhouse gases emitted but remain a popular option for the generation of energy. An easy fix for this conundrum is to utilize other forms of fuel for energy generation which burns cleaner and renewable as opposed to fossil fuels. The aforementioned solution would be to use biomethane generated from waste products which burn cleaner and comes from a renewable source. In this paper, an IoT based biogas monitoring system for biogas reactors is proposed. An ESP-32 microcontroller system is deployed and tested to detect the presence of gas production. A dashboard plotting the data obtained from sensors is designed to help user monitor parameters. Data obtained is automatically uploaded to a Google Spreadsheet for data.

Keywords: Sensors; Biogas monitoring; Microcontroller; IoT

1. Introduction

Fossil fuels have been lambasted for being one of the main contributors of global warming. The burning of fossil fuels releases harmful chemicals into the atmosphere such as carbon dioxide and nitrogen oxides into the atmosphere. Carbon dioxide contributes to the greenhouse effect in which the heat from the sun is trapped in the atmosphere. Since the mid-19th century, the United States alone were responsible for a whopping 29% of carbon dioxide emissions by humans or 328,000,000 metric tons of carbon dioxide [1]. With the increase in temperature comes the melting of the glaciers which results in an increase in sea levels. The world population continues to grow every year, and with that the demand for energy also continues to grow. In big countries such as China, it was estimated that the population in 2021 was 1.43 billion and 543.98 million in 1950 [2]. The increase in population also contributes to the increase in urbanization of a country. With the increase in urbanization comes the increase in energy demand as well. China's energy consumption in 2003

Ian Hoh Wen Yang, Ammar A.M. Al-Talib, Noor Idayu Mohd Tahir, Ayu Haslija Binti Abu Bakar, Nur Muhammad Afifi bin Zainal

reached 1678 million tonnes coal equivalent (or MTCE) where coal consumption took up 67% of the total consumption, oil taking up 23% and hydroelectricity and natural gas being 7% and 3%, respectively [3]. This shows that the burning of fossil fuels remains one of the preferred fuels for energy generation.

With the continuous use of fossil fuels, comes the detriment of the environment and eventual depletion of this fuel source as fossil fuels are not renewable. A promising solution for this is to use biogas as an alternative to fossil fuels. The decomposition of waste products such as animal manure and food wastes through anaerobic digestion of different microorganisms creates biogas as one of the end products of the decomposition process [4]. All that is required by the average household to create their own biogas is an affordable way to monitor the biogas production.

IoT helps connect previously unconnected and dumb physical devices while simultaneously giving them the intelligence to act on a command or situation. It is likely that the integration of IoT technology and other industries can change the mode of economic development, achieve green growth and low-carbon economy [5]. IoT technology has been widely used in environment protection, industry monitoring, food tracing source, logistic trading and other fields [6]. In this paper, a system that utilizes IoT to monitor components in a biogas reactor will be presented.

2. Methodology and Experimental Setup

2.1. *Materials*

Biogas samples were obtained by mixing animal manure and food waste in a blender and then adding inoculum to the solution. Desiccant beads were included in the biogas chamber to remove moisture.

2.2. *Apparatus and Equipment*

The microcontroller used for the data acquisition device is an ESP-32. An MQ-4 semiconductor sensor was used for the detection of methane. An MQ-135 sensor was used to detect the presence of carbon dioxide in the air while a DHT-11 sensor is used for detecting the presence of temperature and humidity. A soldering set was used to create a circuit board for the data acquisition device. IFTTT protocol was used to constantly update a Google Sheets file with the data obtained from the sensors.

2.3. *Prototype Setup*

The basic idea of the system is to monitor the biogas generated in the biogas reactor and update the remote database. The microcontroller collects data from the sensors and then sends it to the Blynk mobile application and Google spreadsheets via IFTTT. Fig. 1 shows the overall architecture of the system proposed while Fig. 2 shows the system flowchart. The reactor contains a mixture of inoculum, food waste and rabbit dung. Biogas produced in the reactor travels from the reactor to a mason jar containing the sensors of the data acquisition device.
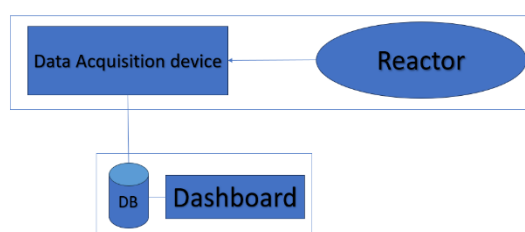
Fig. 1 Overall System Architecture
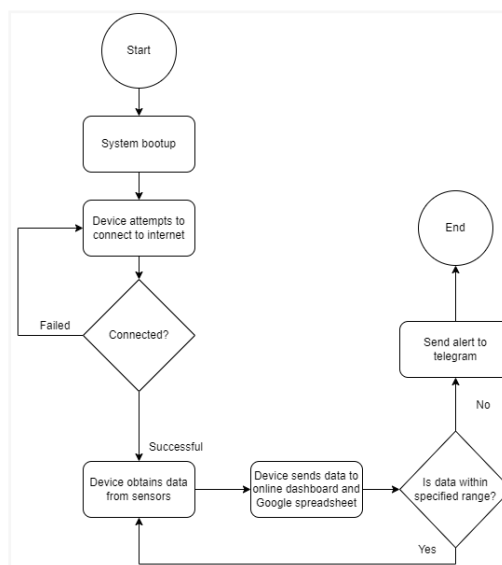
Fig. 2 System Flowchart

The data acquisition device is placed together with the biogas reactor. The device measures the methane gas as well as carbon dioxide produced by the biogas reactor and then uploads the data to Blynk's database and a Google spreadsheet via IFTTT. The data can then be viewed on Blynk's dashboard. The data acquisition

device requires an active internet connection to send data. A Telegram bot was created to receive alerts from the device if the data is out of the specified range.

2.3.1 Preheating Process

The MQ-4 and MQ-135 sensors require initial preheating before the sensors provide stable readings. The datasheet advises that the sensors be given a 24-hour period to preheat. The sensors were found to provide stable readings after 12-hours from the start of power on. The subsequent use of the sensors requires roughly 10-minutes of preheating.

2.3.2 Sensor Calibration Process

The MQ-4 and MQ-135 sensors require calibration. After the sensors were preheated and readings stabilized, data from a reference sensor was used to assist in the calibration process. Both readings were compared and the difference in data was compensated for in the code written for the ESP-32 microcontroller.

2.3.3 Data Acquisition Device fabrication

The data acquisition device consists of an MQ-4 sensor to detect methane, MQ-135 sensor to detect carbon dioxide, and DHT-11 to detect temperature including humidity. The microcontroller used for the device is an ESP-32. Fig. 3 shows the circuit diagram of the device.
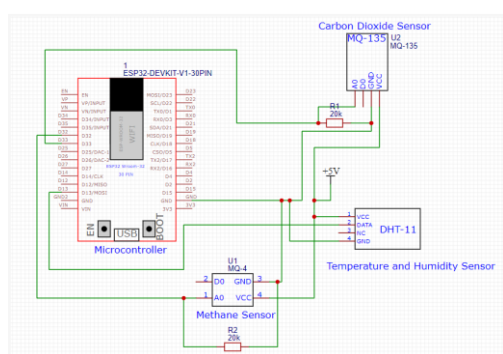
Fig. 3 Data Acquisition Device Circuit Diagram

2.3.4 Blynk Dashboard and IFTTT

A dashboard was designed to show the data obtained by the device. The dashboard displays the data in the form of a graph and the value at an instance on a gauge. The data is provided in PPM (Parts Per Million) over time. Fig. 4 shows the dashboard as viewed on a computer and Fig. 5 shows the dashboard on the mobile application.
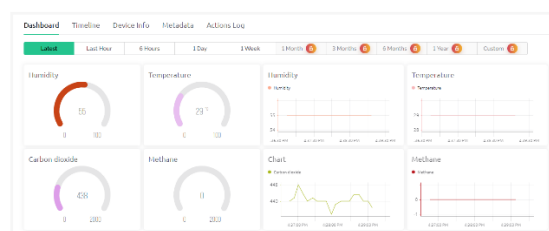
Fig. 4 The dashboard on the computer

Fig. 5 The dashboard on the mobile application

2.3.5 Telegram Alert Function

A Telegram message bot was created to receive alerts from the device in an event where a specific condition has been met (Fig. 6). An example of this would be when the methane or carbon dioxide levels drop below a certain level or when methane is detected. When the methane or carbon dioxide levels drop from a high concentration to a lower concentration, it may signify that a leak is present in the chamber and allows the user to perform data-driven decision making.
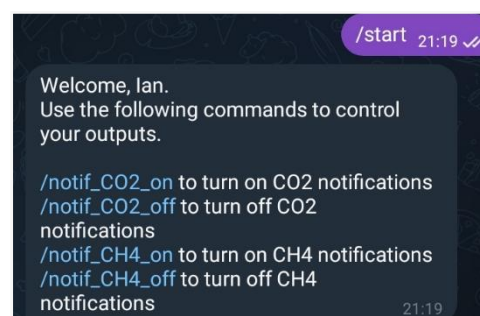
Fig. 6 Telegram Alert Function

2.3.6 Test Setup

The first prototype was built and tested as shown in Fig. 7. From the labels in Fig. 7, one represents the reference sensor, two represents the biogas reactor containing food waste and animal manure, three represents the sensors in a mason jar, and four represents the microcontroller.

Ian Hoh Wen Yang, Ammar A.M. Al-Talib, Noor Idayu Mohd Tahir, Ayu Haslija Binti Abu Bakar, Nur Muhammad Afifi bin Zainal

Fig. 7 Prototype test setup

3. Result and Discussion

3.1. *Data Acquisition Device*

In determining the sensor's accuracy, the sensor readings of the MQ-4 and MQ-135 were compared to a secondary set of data obtained in clean air and in biogas. The DHT-11 sensor data was compared with another measuring device that measures temperature and humidity.

Fig. 8 shows results of the MQ-4 sensor in clean air. The sensor readings fluctuate slightly when compared to the secondary set of data due to noise. Fig. 9 shows the results of the MQ-135 sensor in clean air. The data shows a similar trend when compared to the secondary set of data in clean air. Fig. 10 and Fig. 11 shows the comparison between the data obtained by the DHT-11 sensor and the reference measuring device. It was found that the temperature data from the DHT-11 sensor deviated about two degrees Celsius while the humidity data was found to deviate about three percent. As the data of the MQ-4 and MQ-135 fluctuates drastically, a tabulation of data was created for comparison (Table 1).
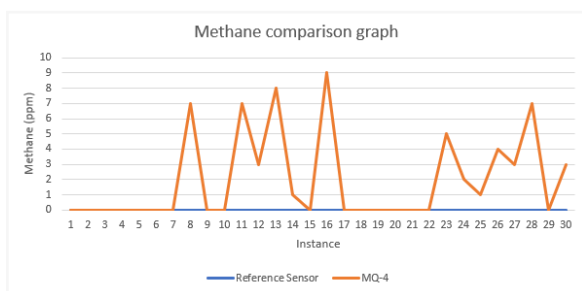
Fig. 8 Methane comparison graph
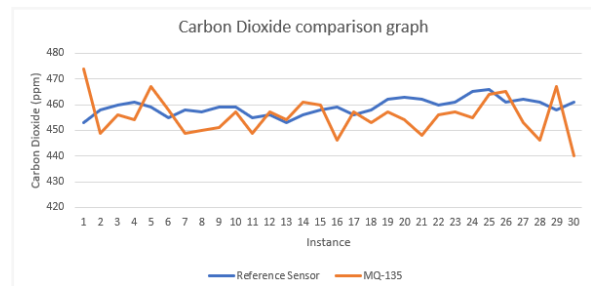
Fig. 9 Carbon dioxide comparison graph
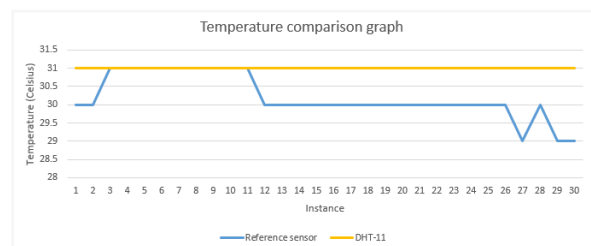
Fig. 10 Temperature comparison graph
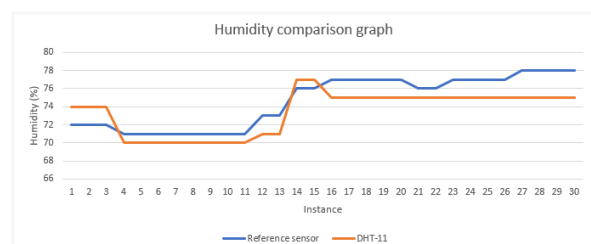
Fig. 11 Humidity comparison graph

Table 1 Sensor Data in Clean Air

| No. | Sensor Data | | | |
	MQ-4	MQ-135	Reference (methane)	Reference (carbon dioxide)
1	0	474	0	453
2	0	449	0	458
3	0	456	0	460
4	0	454	0	461
5	0	467	0	459
6	0	458	0	455
7	0	449	0	458
8	7	450	0	457
9	0	451	0	459
10	0	457	0	459
11	7	449	0	455

No.	Sensor Data			
	MQ-4	*MQ-135*	*Reference (methane)*	*Reference (carbon dioxide)*
12	*3*	*457*	*0*	*456*
13	*8*	*454*	*0*	*453*
14	*1*	*461*	*0*	*456*
15	*0*	*460*	*0*	*458*
16	*9*	*446*	*0*	*459*
17	*0*	*457*	*0*	*456*
18	*0*	*453*	*0*	*458*
19	*0*	*457*	*0*	*462*
20	*0*	*454*	*0*	*463*
21	*0*	*448*	*0*	*462*
22	*0*	*456*	*0*	*460*
23	*5*	*457*	*0*	*461*
24	*2*	*455*	*0*	*465*
25	*1*	*464*	*0*	*466*
26	*4*	*465*	*0*	*461*
27	*3*	*453*	*0*	*462*
28	*7*	*446*	*0*	*461*
29	*0*	*467*	*0*	*458*
30	*3*	*440*	*0*	*461*
31	*4*	*465*	*0*	*463*

Fig. 12 shows the results of the MQ-4 sensor in biogas. It was observed that the MQ-4 sensor provided near similar readings compared to the reference sensor until near the end where a leak was likely the cause of the large deviation. Fig. 13 shows the MQ-135 sensor in biogas. The percentage of error was calculated using Eq. (1).

$$Percent\ error = \frac{Act-Meas}{Act} \times 100\% \qquad (1)$$

Where;

Act= Actual reading
Meas= Measured reading

It was discovered that the MQ-4 sensor had an error of less than 6% and was deemed accurate. It was discovered that in higher carbon dioxide concentration, the MQ-135 becomes less sensitive and thus has a 41% error. However, the MQ-135 did detect the presence of carbon

dioxide and is deemed acceptable. The data for Fig. 12 and Fig. 13 were tabulated in Table 2. The readings from the DHT-11 were affected by the heat generated by the MQ-4 and MQ-135. Due to the reference sensor being placed in a separate compartment, the data of the DHT-11 and reference sensor was different and was thus omitted.
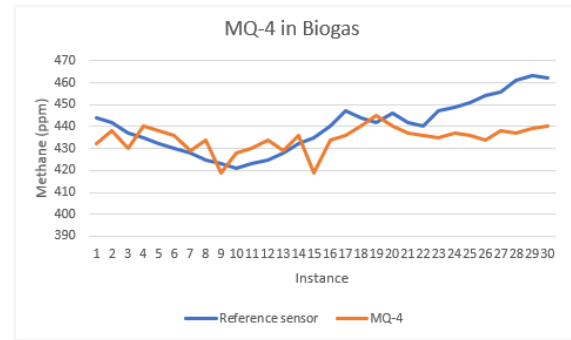
Fig. 12 Comparison Between MQ-4 and Reference in Biogas
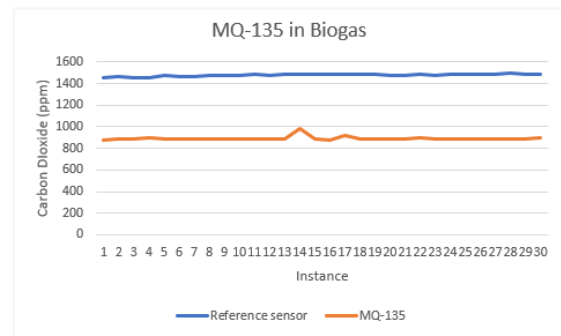
Fig. 13 Comparison Between MQ-135 and Reference in Biogas

Table 2 Sensor Data in Biogas

No.	Sensor Data					
	MQ-4	*MQ-135*	*Ref. (CH₄)*	*Ref. (CO₂)*	*%error (CH₄)*	*%error (CO₂)*
1	432	881	444	1451	2.703	39.283
2	438	882	442	1462	0.905	39.672
3	430	887	437	1459	1.602	39.205
4	440	897	435	1452	1.149	38.223
5	438	885	432	1471	1.389	39.837
6	436	884	430	1466	1.395	39.700
7	429	883	428	1468	0.234	39.850
8	434	885	425	1476	2.118	40.041
9	419	892	423	1473	0.946	39.443

Ian Hoh Wen Yang, Ammar A.M. Al-Talib, Noor Idayu Mohd Tahir, Ayu Haslija Binti Abu Bakar, Nur Muhammad Afifi bin Zainal

No.	Sensor Data					
	MQ-4	*MQ-135*	*Ref. (CH₄)*	*Ref. (CO₂)*	*%error (CH₄)*	*%error (CO₂)*
10	428	890	421	1479	1.663	39.824
11	430	884	423	1482	1.655	40.351
12	434	886	425	1479	2.118	40.095
13	429	888	428	1485	0.234	40.202
14	436	980	432	1482	0.926	33.873
15	419	890	435	1488	3.678	40.188
16	434	877	440	1485	1.364	40.943
17	436	922	447	1484	2.461	37.871
18	440	890	444	1487	0.901	40.148
19	445	885	442	1482	0.679	40.283
20	440	887	446	1475	1.345	39.864
21	437	889	442	1479	1.131	39.892
22	436	895	440	1484	0.909	39.690
23	435	892	447	1479	2.685	39.689
24	437	889	449	1485	2.673	40.135
25	436	889	451	1484	3.326	40.094
26	434	890	454	1490	4.405	40.268
27	438	892	456	1487	3.947	40.013
28	437	890	461	1493	5.206	40.388
29	439	892	463	1491	5.184	40.174
30	440	902	462	1486	4.762	39.300

*To save space methane and carbon dioxide has been abbreviated to CH₄ and CO₂ respectively.

3.2. Power Consumption

The current draw of the sensors including the ESP-32 were pulled from the data sheet. The power consumption was calculated using Eq. (2).

$$Power = \ Current \times Voltage \qquad (2)$$

Table 3 shows that the estimated power consumption of the device totals up to 2.702W. Although the power was calculated using the maximum rated current draw from the component's respective datasheet, the actual power consumption of the device could potentially be higher due to energy loss in the form of heat.

Table 3 Power Consumption

Component	*Current draw (mA)*	*Power (W)*
MQ-4	*150*	*0.75*
MQ-135	*150*	*0.75*
DHT-11	*0.3*	*0.0015*
ESP-32	*240*	*1.2*
Total	*540.3*	*2.7015*

3.3. Project Cost Analysis

The project cost analysis breakdown is shown in Table 4. Some of the components of the project such as the wires and soldering iron were not taken into account for as the cost is negligible due to only using a small amount.

Table 4 Cost Breakdown

Part name	*Individual Price (MYR)*	*Quantity*	*Total (MYR)*
ESP-32	*15.00*	*1*	*15.00*
MQ-4	*5.20*	*1*	*5.20*
MQ-135	*6.50*	*1*	*6.50*
DHT-11	*4.90*	*1*	*4.90*
PCB	*1.00*	*1*	*1.00*
Mason Jar	*10.00*	*1*	*10.00*
Total			*42.60*

4. Conclusion

The design proposed in this paper is a scalable solution to monitor the production of methane and carbon dioxide within the biogas which are important parameters to determine the efficiency of the biogas reactor. The proposed design was tested and was deemed capable in detecting the presence of biogas and carbon dioxide in the biogas reactor.

Acknowledgment

Ayu Haslija from the chemical side of the university for preparing the biogas samples. A massive thank you to Ir. Noor Idayu Mohd Tahir for supervising this project.

References

1. K. A. Hossain, Global warming and impact to third world countries, http://www.ieomsociety.org/imeom/44.pdf (accessed Jul. 27, 2023).
2. H. Ritchie et al., "Population growth," Our world in Data, https://ourworldindata.org/population-growth?utm_campaign=Weekly%2Bnewsletter%2Bof%2BIgor%2BChykalov&utm_medium=email&utm_source=Revue%2Bnewsletter#citation (accessed Jul.20, 2023).
3. P. Crompton and Y. Wu, "Energy consumption in China: Past trends and future directions," *Energy Economics*, vol. 27, no. 1, pp. 195–208, 2005. doi:10.1016/j.eneco.2004.10.006
4. P. G. Kougias and I. Angelidaki, "Biogas and its opportunities—a review," *Frontiers of Environmental Science & Engineering*, vol. 12, no. 3, 2018. doi:10.1007/s11783-018-1037-8
5. R. Liu and J. Wang, "Internet of things: Application and Prospect," *MATEC Web of Conferences*, vol. 100, p. 02034, 2017. doi:10.1051/matecconf/201710002034
6. L. Wan, D. Sun, and J. Deng, "Application of IOT in Building Energy Consumption Supervision," *2010 International Conference on Anti-Counterfeiting, Security and Identification*, 2010. doi:10.1109/icasid.2010.5551509

Authors Introduction

Ammar Abdulaziz Al Talib

He has finished his B.Sc and M.Sc degrees in Mechanical Engineering from the University of Mosul/Iraq. He has finished his Ph.D degree from UPM University / Malaysia. He is also a Chartered Engineer and Member of the Institute of Mechanical Engineers / UK. (CEng. MIMechE). He has developed all the Postgraduate Programs at the Faculty of Engineering at UCSI University / Malaysia and worked as the Head of Postgraduate and Research department at the same faculty for the years 2010-2018.

Ian Hoh Wen Yang

He completed his Bachelor's degree in Mechatronic Engineering from UCSI University, Kuala Lumpur. He is slated to graduate in 2024.

Noor Idayu Binti Mohd Tahir

She graduated from University Teknologi MARA, Malaysia in Mechanical Engineering. She was a design engineer in oil and gas industries for 7 years. Currently, she works as a lecturer under the Department of Mechanical and Mechatronics, Faculty of Technology and Built Environment, UCSI Kuala Lumpur. Her research interests include design and autonomous robotics.

Ayu Haslija binti Abu Bakar

She received her B. Eng in Chemical Engineering in Universiti Teknologi Malaysia. She then received her Master's degree and PhD in University Putra Malaysia. Currently, she is a lecturer under the Department of Chemical & Petroleum Engineering in UCSI Kuala Lumpur. Her research interests include Waste Management (solid & Liquid waste), Pup & Paper (Development of antibacterial packaging, nanocomposite), and Wastewater Treatment (Absorption of pollutant in water & wastewater).

Nur Muhammad Afifi bin Zainal

He finished his B. Eng degree in Chemical Engineering in University Malaysia Pahang. He then continued his M. Sc degree in University Kebangsaan Malaysia. He is currently pursuing a PhD in University Putra Malaysia while working for Tenaga Nasional Berhad Research as a researcher under the Emission and Waste Management Unit.

IoT Based Smart Mushroom Growing Kit

Ammar A.M. Al-Talib

Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment,
UCSI University, 56000 Kuala Lumpur, Malaysia

Cynthia Kuan Jing Ting

Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment,
UCSI University, 56000 Kuala Lumpur, Malaysia

Noor Idayu Mohd Tahir

Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment,
UCSI University, 56000 Kuala Lumpur, Malaysia

Ain Atiqa Binti Mustafa

Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment,
UCSI University, 56000 Kuala Lumpur, Malaysia

Tan Yong Hui

Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment,
UCSI University, 56000 Kuala Lumpur, Malaysia
E-mail: ammart@ucsiuniversity.edu.my, 1001956246@ucsiuniversity.edu.my, NoorIdayu@ucsiuniversity.edu.my,
AinAtiqa@ucsiuniversity.edu.my, tanyh@ucsiuniversity.edu.my

Abstract

This project introduces an IoT-based smart mushroom growing kit to meet the rising global demand for high-quality mushrooms. Various species of mushrooms can be efficiently grown in different environmental conditions with the help of IoT devices that enable farmers to regulate the climate condition according to the specific needs of each type of mushroom. The kit employs sensors and actuators, including temperature, humidity, MQ-135, and ultrasonic sensors, along with an ESP32 camera, controlled by a microcontroller. The collected data is transmitted to an IoT platform via Wi-Fi, facilitating real-time monitoring and control through a user-friendly dashboard on Blynk website and Blynk app. This innovative system optimizes mushroom cultivation by adjusting environmental conditions, offering efficiency and profitability. Users can remotely monitor and regulate the growth environment through their smartphones, enhancing the overall mushroom cultivation experience.

Keywords: Internet of Things; Environmental Condition Control, Mushroom Growing Kit

1. Introduction

With the growth of cities and urbanization in Malaysia, there is a growing demand for sustainable food production within the urban areas. Mushrooms are types of fungi that are known as high protein food which is beneficial for health. Mushrooms were mostly used in health supplements and used as an ingredient in meals to provide nutritional properties to the human body [1]. The demand for mushrooms in Malaysia was expected to increase due to the health awareness among the people increases [2]. Based on Hedley [3], the duration of the complete mushroom life cycle can vary from as short as one day to couple of years based on the types of the mushroom species. Thus, it is best to consume the mushroom fresh within a day. Other than that, drying, freezing, canning, pickling are the common processing and preservation methods to keep

the mushrooms fresh [4]. Despite the fact that, Malaysia is a tropical country with suitable climates condition to grow and produce various mushrooms locally, Malaysia had imported more than five tons of mushrooms every year since 2009 and has reached up to 10 million tons in 2012 which cost more than USD 3.0 million from China [2]. This shows that, the mushroom industry in Malaysia is inadequate as the local mushroom production can barely supply enough mushroom for local consumers [5]. There are several techniques that can be used in mushroom cultivation. Machine learning technique is one of the advanced technologies used in large-scale farming for crops monitoring and pest control. Fuzzy logic technique and command control programming were also used in urban farming and greenhouse to monitor the environmental parameters [6]. However, the cost of implementing advanced technology can be relatively

Ammar A.M. Al-Talib, Cynthia Kuan Jing Ting, Noor Idayu Mohd Tahir, Ain Atiqa Binti Mustafa, Tan Yong Hui

high, making it difficult for small-scale farmers to afford [7].

The paper shows the development of an IoT-based smart mushroom growing kits that allows beginners and urban farmers to monitor and remotely control the growing condition of mushroom easily through smart device without being worried or being at the presence using Blynk IoT application and sensors. The automated IoT smart mushroom growing system can help users to take care of the mushrooms and save user time while enjoying growing mushrooms at home for cooking or to earn extra income.

2. Methodology and Experimental Setup

ESP32 microcontroller was used in this project. The IoT platform chosen for this project is Blynk as it is a mobile application-based platform, easy to use and it offers a user-friendly interface to control and monitor the IoT devices. The controlling parameters and output devices of this project consist of temperature, relative humidity, carbon dioxide level, water level, LED light, fan, misting system and camera.

A. System block diagram

The IoT-based smart mushroom growing kit was designed to monitor the growth of mushrooms under the desirable environmental condition by obtaining data from the sensors and controlling the actuators using IoT technology. Fig. 1 shows the block diagram of the IoT system where the input and output devices were connected to the microcontroller and the ESP32 microcontroller communicates with Blynk IoT platform through Wi-Fi internet connection.
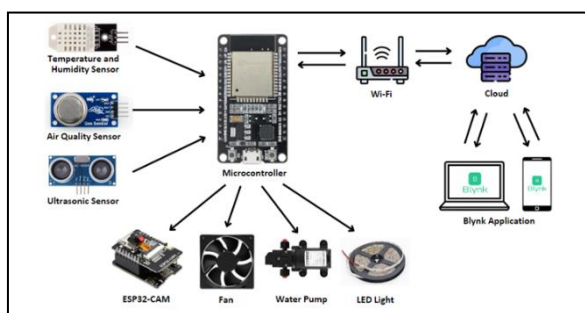
Fig. 1 Block diagram of the system

B. Control System wiring diagram

The control system consists of three subsystems which are the lighting, ventilation and misting system. Fig. 2 shows the wiring connection of the system which

consists of the microcontroller, sensors, actuators and power supply. A three-pin plug was used to connect the 240V AC input power generated from the power station to the house to the power supply transformer to convert voltage 240V AC to 12V DC. Then, the 12V DC output voltage positive and negative pole were connected to the terminal block to supply 12V DC to the fan, water pump, LED light and buck converter. In addition, the ESP32 microcontroller, MQ-135 air quality sensor, HC-SR04 ultrasonic sensor, ESP32-CAM and relay were connected to the output of the buck converter where 12V DC is step down to 5V DC. Lastly, the DHT22 temperature and humidity sensor was powered with 3.3V DC from the ESP32 microcontroller.
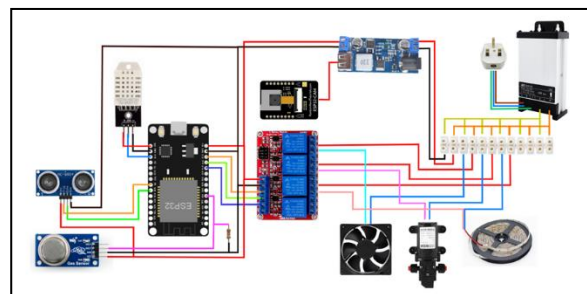
Fig. 2 System wiring diagram

C. System structure design

Fig. 3 shows the structure design of the mushroom growing kit using SolidWorks software which consists of the misting system, LED light, ESP32-CAM for video streaming, MQ-135 air quality sensor, DHT22 temperature and humidity sensor, a DC fan and an ultrasonic sensor. The material used for the mushroom growing kit structure is stainless steel with a dimension of 450mm length, 360mm width and 690mm height. The kit was designed to fit 10 mushroom bags in pyramid arrangement. The electrical component section was designed to keep the electrical components safe on top of the kit and the water tank section was used to place the plastic container used as the water tank for the misting system, while the mushroom room section was used to place the mushroom bags for mushroom development. The components used in the IoT-based smart mushroom growing kit were arranged and labeled in Fig. 3 and Fig. 4.
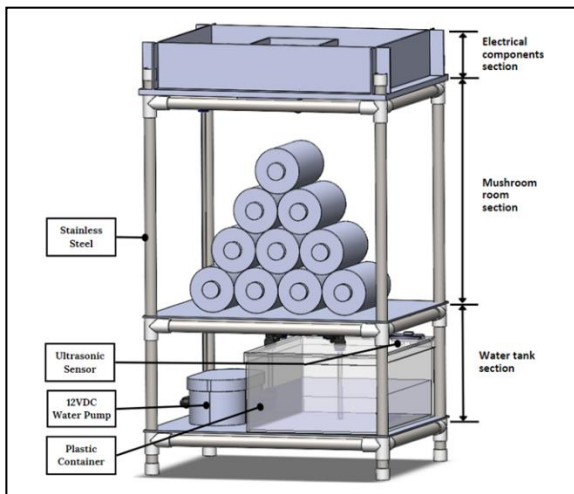
Fig. 3 IoT-based smart mushroom growing kit section
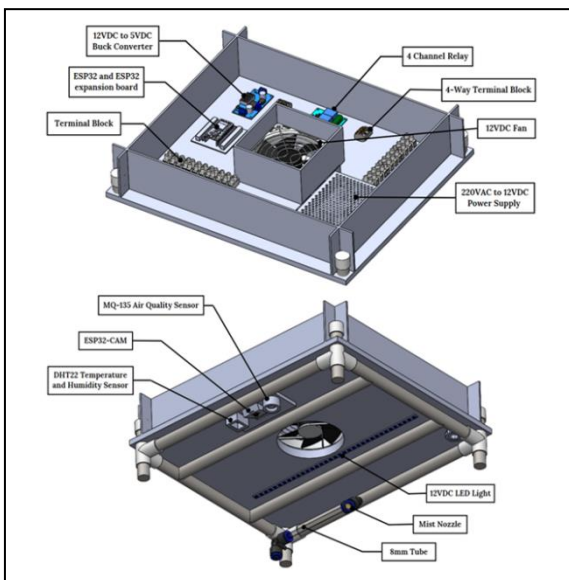
Fig. 4 IoT-based smart mushroom growing kit components labels

3. Results and discussion

A. *IoT-based smart mushroom growing kit prototype*

Fig. 5 shows the kit prototype when the LED light button is pressed as shown in Fig. 6. on Blynk IoT application. Blue light spectrum was used in the prototype due to blue-colored LED light is the most effective color in enhancing the development and production of the oyster mushroom [8].

Fig. 5 Mushroom growing kit prototype

Fig. 6 LED Light setting on Blynk app interface

Desired value for temperature, relative humidity and CO_2 concentration can be set by user in the Blynk app as shown in Fig. 7. Once the desired values are set, the data will be read by ESP32 microcontroller, and the actuators will operate according to the command given by the ESP32 microcontroller to alter the kit environmental temperature, relative humidity and CO_2 concentration. The user is able to monitor and control the kit from the Blynk IoT application as measured readings from the DHT22, MQ-135 and ultrasonic sensors are displayed on the Blynk app IoT interface as shown in Fig. 8 and user could view the live-stream video by clicking on the video stream button on the Blynk app as shown in Fig. 9.
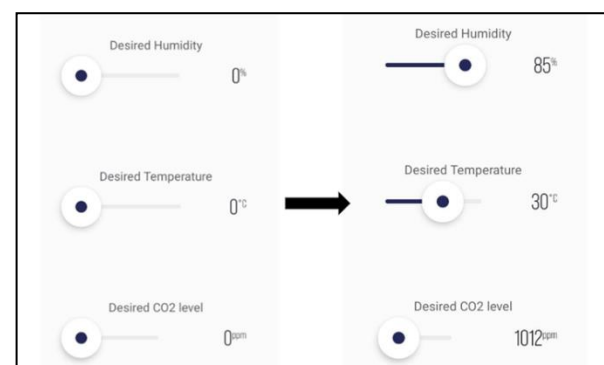
Fig. 7 Desired controlling variables setting on Blynk app interface

Ammar A.M. Al-Talib, Cynthia Kuan Jing Ting, Noor Idayu Mohd Tahir, Ain Atiqa Binti Mustafa, Tan Yong Hui
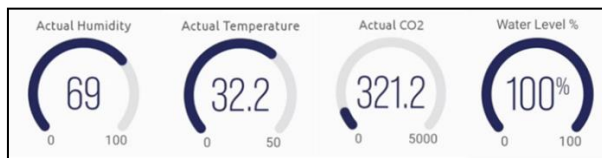
Fig. 8 Measured readings from sensors displayed on Blynk app interface
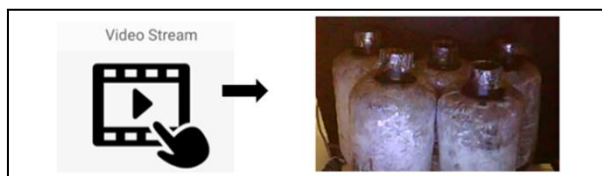
Fig. 9 Video stream button

B. Sensor detection accuracy results

To analyze the accuracy of the sensors, the temperature, relative humidity, CO_2 concentration level and water level readings obtained from a commercial air detector, and a ruler were compared with the ultrasonic sensors used in the kit. After eight attempts of testing, the result shown that temperature and humidity sensor (DHT22), air quality sensor (MQ-135) and ultrasonic sensors (HC-SR04) have high accuracy in detecting temperature, relative humidity, CO_2 concentration level and water level as it has low percentage error below 6%.

C. Results before and after IoT system implementation

Desired temperature value was set at 27°C. Based on Fig. 10, temperature value detected outside the prototype kit remain constant at 28°C and the temperature value measured from the DHT22 sensor inside the prototype kit reduced from 28.1°C to 27.8°C throughout the 10 minutes testing. This has proven that the prototype kit was able to work according to the microcontroller command by reading the desired value from the Blynk IoT application to turn on the fan. However, the prototype kit was only able to reduce the temperature up to 0.3°C due to there is no cooler and heater implemented in the mushroom growing kit.
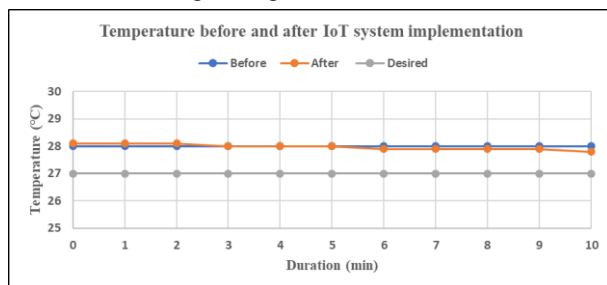
Fig. 10 Temperature before and after IoT system implementation graph

Moreover, the desired value for relative humidity was set at 70%. The relative humidity value detected outside the prototype kit has a slight change from 61% to 63% as shown in Fig. 11. However, it does not reach the desired humidity value. The relative humidity value measured from the DHT22 sensor inside the prototype kit increases from 60.4°C to 69.5°C throughout the 10 minutes testing. This has proven that the misting system in the prototype kit was able to work effectively by proving sufficient humidity required for the mushroom development when the humidity detected is low.
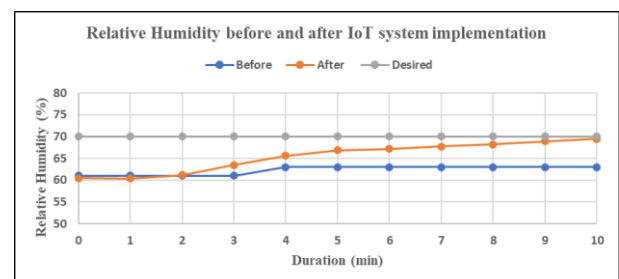
Fig. 11 Relative humidity before and after IoT system implementation graph

D. Mushroom harvest result

A mushroom harvest cycle was conducted to observe the performance of the IoT-based smart mushroom growing prototype kit and the traditional mushroom cultivation methods. Fig. 12 shows the mushroom harvest using traditional method and Fig. 13 shows the mushroom harvest using the IoT-based smart mushroom growing prototype kit.

Fig. 12 Mushroom harvest using traditional method

Fig. 13 Mushroom harvest using prototype kit

The traditional harvest method took eight days for the primordia formation while the mushroom harvest using prototype kit only took five days which is three days faster than the traditional harvest method. Hence, the duration for the mushroom fruitbody development using traditional method is longer compared to the mushroom fruitbody development using prototype kit. Other than that, the weight of the mushroom produced using the traditional method was only 49 grams which is 11 grams lesser than the 60 grams weight of the mushroom produced using the prototype kit. The mushroom produced from the prototype kit is bigger in size and the color is more even compared to the mushroom grown with the traditional method. This shows that the mushroom harvest using prototype kit is more efficient compared to the traditional method. By using the traditional method, user must continuously damp cloth in water and place the cloth on top of the mushroom bag once the cloth is dry. The prototype kit could automatically control the environment condition to ensure the surrounding has enough moisture which could help in speeding up the mushroom development process and to produce good quality mushrooms.

4. Conclusion

The proposed IoT-based smart mushroom growing prototype kit has achieved all requirements to control the environmental condition for the mushroom's development. The kit was capable of controlling the environmental variables based on the desired value set by the user. An IoT system was created with Blynk application to control and monitor the growth of mushrooms through the smartphone. Other than that, users could obtain the data stored on the google sheet for data analysis. The kit was considered a successful innovation to support the sustainable development goals (SDGs) to reduce the use of resources and waste in the mushroom growing process and increases the efficiency and productivity of the mushroom growing operation.

References

1. N. I. P. Samsudin and N. Abdullah, "Edible mushrooms from Malaysia; a literature review on their nutritional and medicinal properties," *International Food Research Journal*, vol. 26, no. 1, pp. 11–31, 2019, Available: http://www.ifrj.upm.edu.my/26%20(01)%202019/(2).pdf.

2. M. Zaffrie, M. Amin, and A. Harun, "Competitiveness of the Mushroom Industry in Malaysia," 2015.

3. E. Hedley, "The Life Cycle of a Mushroom," *Teelixir*, Aug. 02, 2021. https://teelixir.com/blogs/news/mushroom-life-cycle.

4. K. Sharma, "Mushroom: Cultivation and Processing," *International Journal of Food Processing Technology*, vol. 5, pp. 9–12, 2015.

5. M. Zaffrie, M. Amin, A. Harun, M. Mukmin, and A. Wahab, "Status and potential of mushroom industry in Malaysia (Status dan potensi industri cendawan di Malaysia)," *Economic and Technology Management Review*, vol. 9, pp. 103–111, 2014.

6. M. A. Mohd Ariffin et al., "Enhanced IoT-Based Climate Control for Oyster Mushroom Cultivation Using Fuzzy Logic Approach and NodeMCU Microcontroller," *Pertanika Journal of Science and Technology*, vol. 29, no. 4, Oct. 2021, doi: https://doi.org/10.47836/pjst.29.4.34.

7. J. Ruan et al., "A Life Cycle Framework of Green IoT-Based Agriculture and Its Finance, Operation, and Management Issues," *IEEE Communications Magazine*, vol. 57, no. 3, pp. 90–96, Mar. 2019, doi: https://doi.org/10.1109/mcom.2019.1800332.

8. I. Roshita and S. Y. Goh, "Effect of exposure to different colors light emitting diode on the yield and physical properties of grey and white oyster

mushrooms," *AIP Conference Proceedings*, 2018, doi: https://doi.org/10.1063/1.5066751.

Authors Introduction

Ammar Abdulaziz Al Talib

Dr. Ammar Al Talib has finished his B.Sc and M.Sc degrees in Mechanical Engineering from the University of Mosul/Iraq. He has finished his Ph.D degree from UPM University / Malaysia. He is also a Chartered Engineer and Member of the Institute of Mechanical Engineers / UK. (CEng. MIMechE). He has developed all the Postgraduate Programs at the Faculty of Engineering at UCSI University / Malaysia and worked as the Head of Postgraduate and Research department at the same faculty for the years 2010-2018.

Ms. Cynthia Kuan Jing Ting

She received her Bachelor's degree in Mechatronic Engineering in 2023 at UCSI University, Kuala Lumpur. Her field of research is Internet of Thing (IoT) technology.

Ir. Noor Idayu Binti Mohd Tahir

She graduated from University Teknologi MARA, Malaysia in Mechanical Engineering. She was design engineer in oil and gas industries for 7 years. Currently, she works as a lecturer under the Department of Mechanical and Mechatronics, Faculty of Technology and Built Environment, UCSI Kuala Lumpur. Her research interests include design and autonomous robotics.

Ms. Ain Atiqa Binti Mustafa

She received her Master's degree from the Faculty of Electronic & Computer Engineering, Universiti Teknikal Malaysia Melaka, Malaysia in 2017. She is currently a Tutor under department of Mechanical and Mechatronics Engineering, UCSI University, Kuala Lumpur.

Dr. Tan Yong Hui

Dr. Tan Yong Hui finished her B.Sc and M.Sc degrees in Biotechnology from UTAR in 2011. She has finished her Ph.D in mushroom biotechnology from University of Malaya in 2022. She is currently a Lecturer under department of Biotechnology.

Design and Analysis of Artificial Magnetic Conductor for Metal Shielding Applications in RFID Car Detection Applications

Eryana Hussin[1,2]
[1]*Faculty of Engineering, UCSI University, Malaysia*
[2]*Electrical and Electronic Engineering, Universiti Teknologi PETRONAS, Malaysia*
Email: [1]EryanaEiyda@ucsiuniversity.edu.my, [2]eryana_17007877@utp.edu.my

Azman Zakariya
Electrical and Electronic Engineering, Universiti Teknologi PETRONAS, Malaysia
Email: mazman_zakariya@utp.edu.my

Md. Ashraful Haque[2,3]
[2]*Electrical and Electronic Engineering, Universiti Teknologi PETRONAS, Malaysia*
[3]*Electrical and Electronic Engineering, Daffodil International University, Dhaka, Bangladesh*
Email: md_21000832@utp.edu.my

Nur Izzati Ali
Faculty of Electronic Engineering & Technology, University Malaysia Perlis, Malaysia
Email: nurizzatimohdali@unimap.edu.my

Abstract

In this paper, the process of developing the reformulation equation of the Square AMC is conducted. The reformulation yielded results with less than a 5% difference for an AMC using FR-4 substrate with thickness of 3 mm. The equation for calculating capacitance was developed using the conformal mapping theory, with the assumption that inductance remains fixed at the thickness of the AMC. A comprehensive study analysis and integration of an Artificial Magnetic Conductor (AMC) tailored explicitly for metal shielding in Radio Frequency Identification (RFID) applications. The primary focus is to establish a stable and efficient communication link between RFID readers and tags, ensuring reliable data transmission within metallic environments. The proposed AMC structure exhibits a symmetrical design to facilitate seamless integration with RFID tags. This integration strategy aims to optimize communication stability and signal fidelity despite challenging metallic surroundings.

Keywords: Artificial Magnetic Conductor (AMC), RFID, Dipole Antenna.

1. Introduction

The Artificial Magnetic Conductor (AMC) belongs to the metamaterial structure group, specifically designed to enhance the performance of various antenna types, including dipole, monopole, and patch antennas [1], [2], [3]. Employing AMC as the ground plane serves to enhance the antenna's gain by minimizing undesirable back radiation and mitigating mutual coupling effects. The AMC is characterized by its high surface impedance, exceeding a thousand ohms, rendering it akin to an open circuit or lossless structure under these conditions. The reflection graph, which exhibits a phase angle of zero degrees and a magnitude of +1 when the system is in resonance, serves as a defining feature of AMC behaviour [4], [5], [6].

In [7] and [8], AMC-based antenna designs were explored to enhance gain and optimize radiation patterns by manipulating the top layer of the Perfect Electric Conductor (PEC) patch on a conductive substrate material. By applying the lumped circuit concept in [9], the modelling strategy based on characteristic modes, a generic smartphone antenna's impedance characteristic can be achieved. The research was further extended by [10], who employed a similar circuit model to calculate the resonant frequency and reflection phase of structures based on Jerusalem crosses for artificial magnetic conductors (JC-AMC) to intercept waves that would typically strike the structure.

In the paper presented by [11] includes the analysis of a new equation model of resonant frequency for absorber based on the AMC structure composed an array of square patch. Using an equivalent LC circuit model to analyse the square patch array's physical dimensions and material properties, the resonant frequency is determined. The comparable LC circuit model formula was derived using the conformal mapping theory for neighbouring co-planar square patches. While the Gauss' law and the Faraday's law are applied to the square patch that is parallel to the ground plane. This paper proposed a modification of the standard equation of square patch

Eryana Hussin, Azman Zakariya, Md. Ashraful Haque, Nur Izzati Ali

AMC for different parameters and produce smaller discrepancy between the simulated and calculated resonant frequency. The new equation was modified by introducing the correction factor based on the parametric analysis conducted. The difference of formulation for capacitive element of the proposed structure based on the parametric analysis of patch width, substrate width and gap size.

2. Methodology

The method of RFID tag for metal object detection presented in the paper is divided into two parts; single AMC design and array AMC with dipole antenna integrations.

2.1. Single 2.45 GHz AMC design

The single AMC structure consists of three layers of patch, substrate, and ground plane. The size of the patch must be smaller than the substrate to allow some gap between the patches when integrating with the dipole antenna in array arrangement. The resonant frequency and bandwidth of the standard AMC structure presented in square shape can be calculated using Eq. (1) and Eq. (2) where the L and C represents the inductive and conductive elements of the structure which determine by the Eq. (3) and Eq. (4). Fig. 1 shows the structure of the basic square AMC with square patch on a FR-4 substrate with 3 mm thickness and backed with a full ground plane.

$$f = \frac{1}{2\pi\sqrt{LC}} \quad (1)$$

$$Bw = \frac{1}{\eta_0}\sqrt{\frac{L}{C}} \quad (2)$$

$$C = \frac{w_p\varepsilon_\circ(1+\varepsilon_r)}{\pi}cosh^{-1}\left(\frac{w_s}{g}\right) \quad (3)$$

$$L = \mu_\circ h \quad (4)$$

where w_p= patch width, w_s= substrate width, g=gap around patches, ε_r=substrate permittivity, h =substrate thickness.
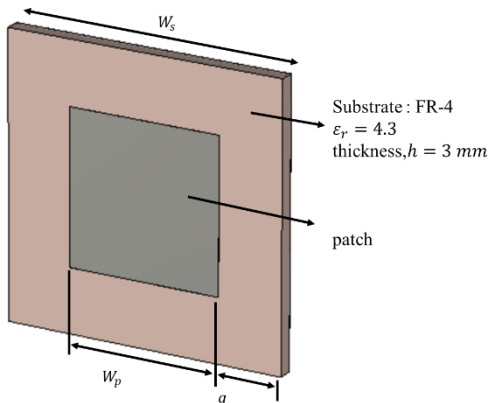
Fig. 1 Structure of the basic square patch AMC, SP-AMC

The flowchart in Fig. 2 show the design steps of designing the Square AMC using CST software. The single structure is simulated using the frequency domain solve with modified boundary setting shows in the flowchart. The simulated result should meet the characteristics of the AMC by achieving the zero-reflection phase and high impedance at resonant [12], [13].
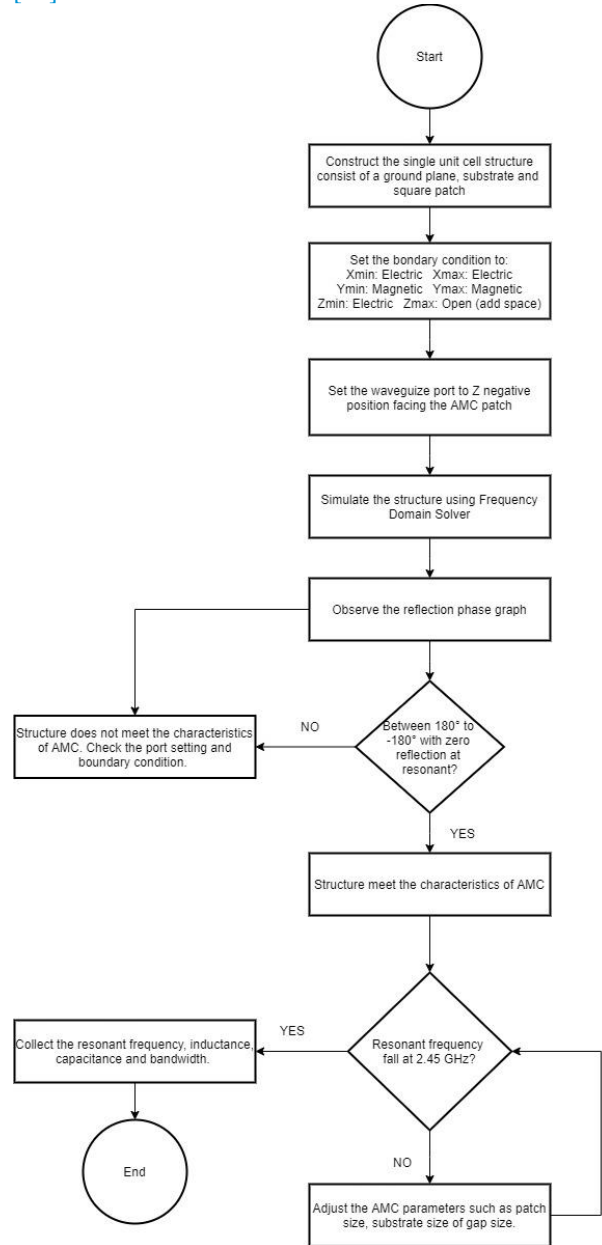
Fig. 2 The flowchart of single unit AMC design using CST software.

2.2. Dipole antenna and Array AMC integration

The second part discussed about the proses of dipole and array AMC integration. A simple dipole antenna will be used to replicate the RFID Tag antenna. The 50Ω discrete port is used to replace the RFID chip at the center of the dipole antenna [14]. The single unit cell

AMC will be arranged in 2×2 and 3×3 arrangement and will be presented Section 3.

Fig. 3 shows the flow chart of the dipole antenna and array AMC integration. To maintain the symmetrical structure, the array arrangement used in this chapter will be in constant multiplication of 2×2 and 3×3 arrangement. It is important to analyze the effect of adding the dipole antenna at different positions on the two different array SP-AMC to see the effect of different metallization around the dipole antenna.
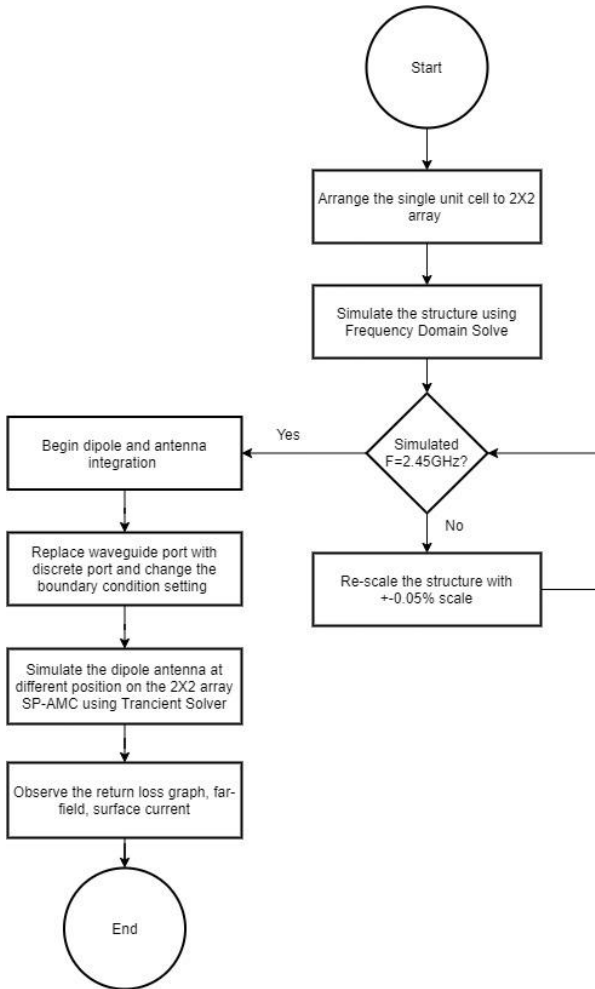
Fig. 3 Flow chart of the dipole and array AMC

3. Results and Discussion

3.1. *Analysis of single unit cell AMC*

In this section, the characteristics of the capacitive and inductive components will be investigated through the simulated frequency. Then, the capacitance value is calculated based on the resonant frequency. By applying the conformal mapping theory, the inductance of the Square AMC use in this paper will be based on the fixed thickness of 3 mm and Fr-4 as substrate. By using the Eq. (4), the inductance value is maintained at 3.77 nH. With this, the capacitance value can be calculated using the Eq. (1) and Eq. (3). The capacitance obtained from both equations will be compared and the Eq. (3) will be

reformulated again using the curve fit method. The reformulation of Square Based AMC has been conducted by [11] and produced a maximum deviation of 9.59%. In this paper, the new formulated proposed is expected to be less than 5% by using the MATLAB curve tools.

Table 1 shows the comparison of simulated and calculated resonant frequency of the AMC with substrate size of 50 mm and patch size varied from 18 mm to 21 mm. The calculated frequency using the standard formula displays a significant disparity with percentage different more than 30% when compared to the simulated frequency. The reformulation of the Eq. (5) was conducted using the curve fit tools in MATLAB. The reformulated resonant frequency (R) shown in Table 1 produce smaller discrepancy less than 1%. Fig. 4 shows the comparison between the standard formula, simulated and reformulation resonant frequency of the square AMC with different patch size.

Table 1: Comparison of the S=simulated, C=Calculated, R=Reformulated and percentage different CS= calculated and simulated and RS = reformulated and simulated.

Patch size	Frequency (GHz)			Percentage Different	
	S	C	R	CS	RS
18	3.62	4.81	3.58	33.0%	0.9%
19	3.44	4.59	3.43	33.5%	0.3%
20	3.27	4.38	3.28	34.0%	0.3%
21	3.15	4.19	3.15	33.1%	0.1%

$$C = \left(\frac{1.63 w_p \varepsilon_0 (1 + \varepsilon_r)}{\pi} \right) \times$$
$$cosh^{-1} \left(\frac{ws}{0.9(w_p - w_{s)}} \right) \quad (5)$$
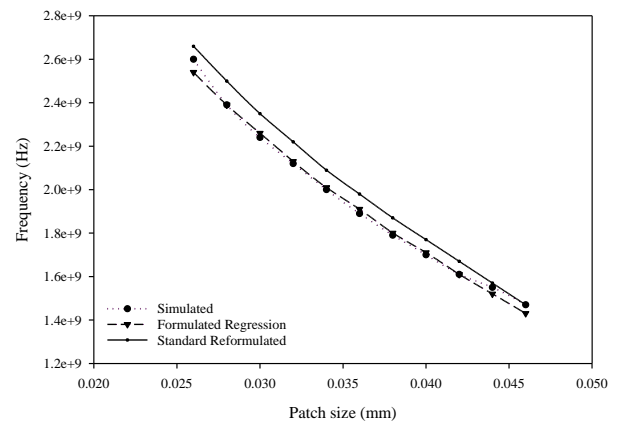
Fig. 4 Comparison of frequency based on standard formula, simulated, and reformulated for the Square AMC with different patch (w_p) and gap (g) size.

The Eq. (5) is used to recalculate the frequency based on the proposed structure in [15] and [16] respectively. Table 2 shows the comparison of calculated frequency based on the proposed equation and the actual resonant frequency presented in the paper. Both use the same substrate but with different thicknesses. The comparison yields a percentage difference below 15%.

Table 2: Comparison of frequency the Square Patch AMC based on the previous papers.

Paper	Structure dimension	F_R, GHz	F_A, GHz	Percentage Different
[15]	10.8 × 12.4 × 1.6	3.40	3.90	12.8%
[16]	25.6 × 28.1 × 1.25	2.22	2.45	9.54%

where F_R = calculated frequency based proposed reformulation in Eq. (5) and F_A =actual resonant frequency

3.2. *Dipole and AMC Integration and Analysis*

This section will discuss the performance of the dipole antenna on two different AMC array integrations. Dipole antenna is placed at different positions on the array antenna to obtain optimized results in term of gain, directivity and efficiency. Fig. 5 shows the structure of dipole antenna placed at two different positions on the 2×2 array AMC. Based on the simulated graph shows in Fig. 8, the dipole antenna on the 2×2 array SP-AMC affects its interaction with the metallization components of the structure, leading to changes in impedance, radiation pattern, near-field effects, and resonance behavior. The overall structure for 2×2 array SP-AMC is measured at $92 \, mm \times 92 \, mm$.

(a)   (b)
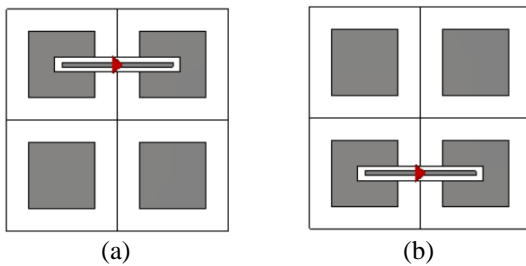
Fig. 5 Placement of the dipole antenna with optimized performance at the top center and bottom center on the 2×2 array AMC

As the dipole antenna position changed around the 2×2 array SP-AMC, the coupling and interaction with the near-field area due to the patches of gap between the patches has different effect to the efficiency of the energy and radiation characteristics. For example, when the dipole is placed at the top center and bottom center of the 2×2 array SP-AMC, the center of the dipole is in contact

with the gap between the patches. The placement of the dipole antenna on the 2×2 array SP-AMC also affects the effective capacitance and inductance of the SP-AMC, hence modifying the mutual coupling and impedance of the structure. Based on the simulated graph, the optimized result obtained when the dipole is place at the center top and bottom of the 2×2 array SP-AMC with return loss -28.03 dB.

Fig. 6 shows the surface current distribution of the dipole antenna on the 2×2 array SP-AMC. More inducted current from the dipole antenna was distributed around the top AMC structure compared to the bottom. The AMC structure placed at the back of the dipole antenna helps to realign the surface current to be in-phase with the dipole antenna. Hence, overcome the problem of metal object detection in RFID application.
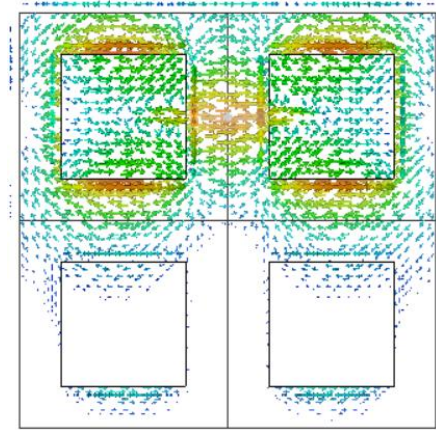
Fig. 6 Surface current distribution of the dipole antenna backed with 2×2 array AMC.

The next analysis discussed on the performance of dipole antenna on the 3×3 array AMC. The dipole is placed at the top and bottom of the gap between patched on left and right of the array the structure as shown in Fig. 7. The overall structure for 2×2 array SP-AMC is measured at $138 \, mm \times 138 \, mm$. Table 3 shows the comparison of the performance of the dipole antenna on two different square AMC array arrangements.

(a)   (b)
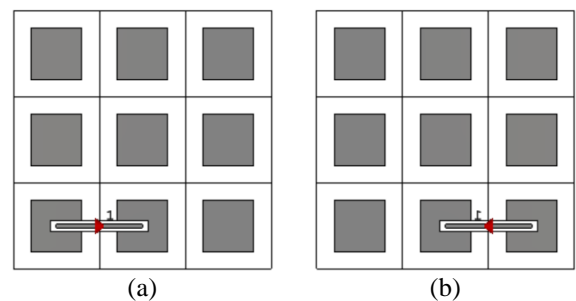
Fig. 7 Placement of the dipole antenna with optimized performance on the 3×3 array AMC

Table 3: Comparison of the dipole antenna performance with different AMC array arrangements.

Structure	Gain (dB)	Directivity (dB)
Dipole only	1.87	2.02
Dipole with 2×2 SP-AMC	3.32	8.21
Dipole with 3×3 SP-AMC	3.75	7.58

Fig. 9 shows the surface current distribution of the dipole antenna on the 3×3 array SP-AMC. More inducted current from the dipole antenna was distributed around the top AMC structure compared to the bottom. The induced current are lesser as the distance between the dipole and array cell increased due to the large gap used in this structure. Therefore, for structure with big gap, it is recommended to limit the array cell to achieve compact structure.

The return loss in Fig.8 shows the comparison between the dipole antenna with and without the array AMC. It is observed that the use of AMC degrades the bandwidth of the dipole antenna. However, the integration of dipole and AMC shows in Table 3 improved 77% of gain and 275% of directivity compared to dipole antenna without AMC.
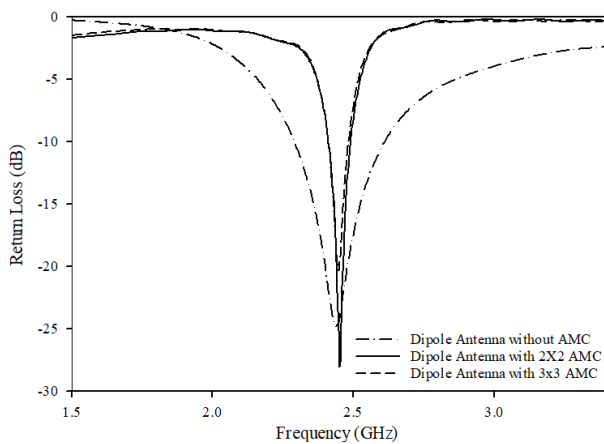
Fig.8: Return loss of the dipole antenna with and without the AMC
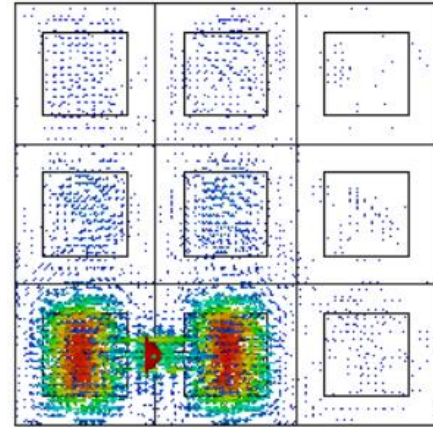
Fig.9: Surface current distribution of the dipole antenna backed with 3×3 array AMC

4. Conclusion and Recommendations

To conclude, the first part of this paper the new reformulation of square AMC has been proposed with smaller discrepancy of less than 1% compared to the simulated resonant frequency in CST software. The second part discussed the performance of the simple dipole antenna on two different arrays with optimized results obtained at the top and bottom center of the SP-AMC structure.

The dipole and array AMC structure proposed in the paper can be used as metal shielding in RFID applications such as car detection for parking or toll-system. The use of AMC to the dipole RFID tag can improve the detection range as the AMC helps to reduce the back radiation of the dipole and improve the directivity and gain of the structure. The array arrangement in this paper is limited to 3×3 to minimize the overall structure.

A bigger array is recommended for any application which has no objection on the size limits. For low frequency such as UHF RFID, the AMC structure with slots can be introduced to achieve compact size.

References

1. Capolino, Filippo. Theory and phenomena of metamaterials. CRC Press, 2017.
2. Zharov, Alexander A., Ilya V. Shadrivov, and Yuri S. Kivshar. "Nonlinear properties of left-handed metamaterials." Physical Review Letters 91, no. 3 (2003): 037401.
3. G. V. Trentini, "Partially reflecting sheet arrays," IRE Trans. Antennas Propag., vol. AP-4, pp. 666–671, 1956.
4. E. Carrubba, A, Monorchio, and G. Manara "Artificial Magnetic Surface for Circular Polarization Movemenent" Microwave and Optical Technology Letters, Vol 5, pp.1782-1786, Aug 2010.
5. M Elena de Cos, Yuri Alvarez and Fernando Las-HEras, "Design and Characteristics of Planar Artificial Magnetic Conductor in the RFID SHF Band" Proceeding of Forth European Conference on Antenna and Propagation 2010, pp. 1-5, Apr 2010

6. Sievenpiper D. F. "High Impedance Electromagnetic Surface" PHD Thesis, University of California at Los Angeles 1999.

7. D. SievenpiperLijun Zhang, Romulo F. Jimenez Broas, Nicholas G. Alex´opolous, and Eli Yablonovitch, "High-Impedance Electromagnetic Surfaces with a Forbidden Frequency Band" IEEE Transactions On Microwave Theory And Techniques, Vol. 47, No. 11, November 1999, pp: 2059 – 2074.

8. Constantin R. Simovski, Peter de Maagt, and Irina V. Melchakova, "High-Impedance Surfaces Having Stable Resonance With Respect to Polarization and Incidence Angle" IEEE Transactions On Antennas And Propagation, Vol. 53, No. 3, March 2005, pp: 908-914.

9. Jafargholi, A,Atlasbaf, Z. And Kamyab, M., "Miniaturized dual-band dipole antenna loaded with metamaterial based structure," 2011 19th Iranian Conference Electrical Engineering (ICEE), pp.1-4, 17-19 May 2011.

10. Sujatha, M. N., and K. J. Vinoy. "A stacked ring-patch artificial substrate for surface waves suppression and in phase reflection." In Industrial and Information Systems (ICIIS), 2010 International Conference on, pp. 75-79. IEEE, 2010.

11. Kurniawan, Adit, Achmad Munir, and Levy Olivia Nur. "Theoretical Analysis of Resonant Frequency for AMC-based Absorber Composed of Square Patch Array." International Journal on Electrical Engineering & Informatics 7.2 (2015).

12. Abu, M., E. E. Hussin, A. R. Othman, Fauzi Johar, Norhidayah M. Yatim, and Rose F. Munawar. "Design Of 0.92 GHz Artificial Magnetic Conductor For Metal Object Detection In RFID Tag Application With Little Sensitivity To Incidence Of Angle." Journal of Theoretical & Applied Information Technology (2014).

13. Abu, M., E. E. Hussin, A. R. Othman, N. M. Yatim, F. M. Johar, and R. F. Munawar. "Design of stacked wafers AMC at 920 MHz for metallic object detection in RFID application." In 2013 IEEE Symposium on Wireless Technology & Applications (ISWTA), pp. 236-239. IEEE, 2013.

14. Abu, M., E. E. Hussin, M. S. M. Isa, Z. Zakaria, and Z. A. Baharudin. "Designing artificial magnetic conductor at 2.45 GHz for metallic detection in RFID tag Application." International Journal of Engineering and Technology (IJET) 6, no. 1 (2014): 427-435.

15. Askari, Hussain, Niamat Hussain, Domin Choi, Md Abu Sufian, Anees Abbas, and Nam Kim. "An AMC-Based Circularly Polarized Antenna for 5G sub-6 GHz Communications." Computers, Materials & Continua 69, no. 3 (2021).

16. Eardprab, Sanchai, Chuwong Phongcharoenpanich, and Danai Torrungrueng. "Improvement of a circular microstrip antenna excited by four feeds and suspended with artificial magnetic conductors." International Journal of Antennas and Propagation 2013 (2013).

Dr. Azman Zakariya

He joined Universiti Teknologi PETRONAS in 1999. He is holding a postion as Senior Lecturer in the Electrical and Electronic Engineering Department. He obtained the B.Eng. degree in Electrical & Electronic Engineering in 1999 from Universiti Teknologi Malaysia, Malaysia, Masters of Science Degree in Communication And Signal Processing in 2001 from University of Newcastle Upon Tyne, and PhD in RF and Microwave in 2015 from Universiti Sains Malaysia. Currently he is assigned as Head of Research Group of Algriculture IOT (AIOT), under Institute of Tranportation and Infrasture (ITI). His research interest is in RF and Microwave Circuit Design, Antenna and Wireless System and IOT. Systems.

Mr Md Ashraful Haque.

He is doing Ph.D. at the Department of Electrical and Electronic Engineering, Universiti Teknologi PETRONAS, Malaysia, He got his B.Sc. in Electronics and Electronic Engineering (EEE) from Bangladesh's Rajshahi University of Engineering and Technology (RUET) and his M.Sc. in the same field from Bangladesh's Islamic University of Technology (IUT). He is currently on leave from Daffodil International University (DIU) in Bangladesh. His research interest includes microstrip patch antenna, sub 6 5G application, regression model machine learning on antenna design.

Dr. Nur Izzati Mohd Ali

Shei received her bachelor's and master's degrees from Universiti Teknikal Malaysia Melaka (UTeM). She pursued her Ph.D. at Universiti Kebangsaan Malaysia (UKM). Currently, she is a senior lecturer in the Faculty of Electronic Engineering & Technology at Universiti Malaysia Perlis. Her expertise lies in antenna design, particularly in transparent antennas, solar-integrated antennas, and wireless MIMO communication systems.

Authors Introduction

Ms. Eryana Hussin

She he received her Master's and Bachelor's degree in Electrical and Electronic Engineering from University Teknikal Malaysia Melaka. She is currently pursuing her study under Doctor of Philosophy in Electrical and Electronic Engineering in Universiti Teknology Petronas, Malaysia. She is also a lecturer from the Department of Electrical and Electronics of UCSI University, Malaysia. Her research interest covers in Metamaterical structures, Antenna and Wireless System.

Effect of Fiber Orientation on the Mechanical Performance of Natural Fiber Polymer Composite Bicycle Frame using Finite Element Analysis

Kok Sem Too
Faculty of Engineering, Technology and Built Environment,
UCSI University, Kuala Lumpur, 56000, Malaysia

Cik Suhana Hassan
Faculty of Engineering, Technology and Built Environment,
UCSI University, Kuala Lumpur, 56000, Malaysia

Nor Fazilah Abdullah
Faculty of Engineering, Technology and Built Environment,
UCSI University, Kuala Lumpur, 56000, Malaysia

Ammar Abdulaziz Majeed Al-Talib
Faculty of Engineering, Technology and Built Environment,
UCSI University, Kuala Lumpur, 56000, Malaysia

Email: 1001748617@ucsiuniversity.edu.my, suhana@ucsiuniversity.edu.my,
norfa@ucsiuniversity.edu.my, ammart@ucsiuniversity.edu.my

Abstract

In this research, the performance of oil palm empty fruit bunch (OPEFB) fiber-reinforced epoxy composite with varying fiber orientation and stacking sequence as the material for mountain bike frame was studied utilizing ANSYS software. The choice of OPEFB fiber was motivated by the fact that the waste by-product of oil palm extraction in Malaysia alone might reach 70-80 million tons per year, with 90% of oil palm biomass lost as waste. The properties of epoxy OPEFB composite in principal 1, 2, and 3 directions were calculated using Whitney and Riley estimates. 10 stacking sequences and five loading conditions were taken. The results show that the fibre orientation of epoxy OPEFB composite on the bicycle frame had little effect on the performance contrary to the number of plies in the laminate or number of laminates which had major effects.

Keywords: Biocomposites, Bicycle frame, Oil palm empty fruit bunch, Polymer composites

1. Introduction

In the 21st century, companies and governments always push towards green technologies, eco-friendly products, and sustainability. In fact, in 2015, Sustainable Development Goals (SDGs) were established by the United Nations (UN) to push for all countries to take part in sustainability for the conservation and preservation of nature and resources for current and future generations [1]. SDGs were fully adopted by the Malaysian government and taught in universities. With this awareness, to reduce the use of resources like fuels and cut carbon emissions, bicycles become the ideal vehicle for short-distance travel.

While it is true that bicycles do not produce emissions, the manufacturing of steel, aluminium, or titanium bicycle frames produces a lot of carbon emissions. With the increasing popularity of bicycles, more bicycles are going to be manufactured and more carbon emissions are going to be produced. In Malaysia, there is also another problem involving the oil palm industry. Oil palm empty fruit bunches (OPEFB) are discarded by the industry as waste in large quantities without utilization [2]. OPEFB fibre has the potential to be used as reinforcing fibre in composite materials. Therefore, utilizing OPEFB fibre in a composite bicycle frame would ease both problems at the same time.

Many researchers have investigated the properties of OPEFB fibre and OPEFB fibre in matrix. Gunawan et al. [3] performed experiment to find out mechanical properties of 40 strands of OPEFB fibre and found that the average Young's modulus is 11.88 GPa and the average tensile strength is 156.3 MPa. Zuhri [4] experimented on just a single strand of OPEFB fibre and found that Young's modulus is 1.7 GPa and tensile strength is 71 MPa. Hassan [5] experimented on

Kok Sem Too, Nor Fazilah Abdullah, Ammar Abdulaziz Majeed Al-Talib, Cik Suhana Hassan

unidirectional OPEFB fibre-reinforced epoxy composite in 0º, 45º, and 90º directions and found that the tensile strengths are 30.5 MPa, 10 MPa, and 9.2 MPa respectively.

Utilizing unidirectional continuous OPEFB fibre in a composite bicycle frame requires the study of the fibre orientation and stacking sequence to produce a bicycle frame that can withstand all loading conditions. As investigated by the researchers, the properties of OPEFB fibre are adequate for the development of composite bicycle frames.

2. Methodology

2.1. *Mountain bicycle frame*

The dimensions and geometry of the mountain bicycle frame was adopted from a journal by Sajimsha B et al. [6]. The modelling of the frame was completed on SOLIDWORKS. Table 1 shows the parameters of dimensions and geometry of the bicycle frame.

Table 1. Parameters of dimensions and geometry of the mountain bicycle frame

Parameter	Value
Head tube angle	73.5º
Seat tube angle	73.5º
Top tube length	580 mm
Seat tube length	570 mm
Chain stay length	360 mm
Head tube length	120 mm

For composite material simulation, the shell model of the bicycle frame was exported to ANSYS. The meshing was done on ANSYS Mechanical and the completed meshed model has 12830 nodes and 12306 elements. Fig. 1 shows the completed meshed model of the bicycle frame.
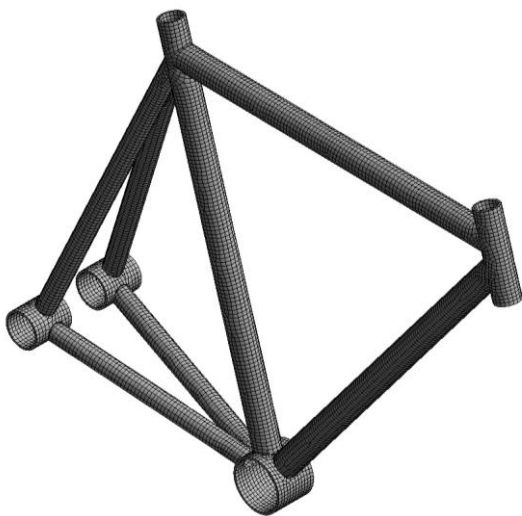
Fig.1 Shell model of mountain bicycle frame

2.2. *OPEFB fiber reinforced epoxy composite properties*

Although many researchers have conducted experiments on the properties of OPEFB fibre reinforced epoxy composite, only the properties of Young's modulus, tensile strength, and flexural strength were found. Properties such as shear modulus and Poisson's ratio in principal 1-2, 2-3, and 1-3 directions and compressive strength and shear strength in principal 1, 2, and 3 directions were not found. To find the moduli and Poisson's ratios in all three directions of composite, many theoretical estimates were developed. One such theoretical estimates is Whitney and Riley estimates developed by Air Force Materials Laboratory [7-8]. The calculations of the moduli and Poisson's ratios using Whitney and Riley theory are less rigorous than other theoretical estimates and only require Young's modulus, shear moduli, bulk moduli and Poisson's ratios of the fibre and matrix. The theory also regards that the fibre and matrix are isotropic. The compressive strength and shear strength of the composite were calculated using the equations in Chapter 5 of the book of Engineering Mechanics of Composite Materials by Daniel and Ishai [9]. The properties of OPEFB fibre were obtained from an experiment done by Gunawan et al. [3] and the properties of epoxy were obtained from Appendix A in the book [9]. Table 2 shows the calculated properties of the epoxy OPEFB composite. Note that the fibre-to-matrix volume ratio is 60:40.

Table 2. OPEFB fibre properties

Young's modulus 1 (GPa)	8.848
Young's modulus 2 (GPa)	6.347
Young's modulus 3 (GPa)	6.347
Poisson's ratio 1-2	0.3184
Poisson's ratio 2-3	0.32
Poisson's ratio 1-3	0.3184
Shear modulus 1-2 (GPa)	2.899
Shear modulus 2-3 (GPa)	2.404
Shear modulus 1-3 (GPa)	2.899
Bulk Modulus (GPa)	5.937
Tensile strength 1 (MPa)	142 MPa
Tensile strength 2 (MPa)	49.46 MPa
Tensile strength 3 (MPa)	49.46 MPa
Compressive Strength 1 (MPa)	400 MPa
Compressive Strength 2 (MPa)	143.4
Compressive Strength 3 (MPa)	143.4
Shear Strength 1-2 (MPa)	70.82
Shear Strength 2-3 (MPa)	48
Shear Strength 1-3 (MPa)	70.82

2.3. *Stacking sequence*

The stacking sequences of the plies were determined from the journal of Thomas Jin-Chee LIU et al. [10] who found 10 stacking sequences of 8-ply laminate that

produced better designs for composite bicycle frames. The 10 stacking sequences are shown in Table 3. Each stacking sequence is designated with a letter with 0, 45 and 90 indicating the fibre orientation angle in the composite. Fig. 2 shows examples of 0º fibre orientation and 90º fibre orientation on the bottom bracket.

Table 3. Composite stacking sequence

Designation	Stacking Sequence
A	[0/90/45/-45]s
B	[0/90/-45/45]s
C	[90/0/45/-45]s
D	[90/0/-45/45]s
E	[-45/45/0/90]s
F	[-45/45/90/0]s
G	[45/-45/0/90]s
H	[45/-45/90/0]s
I	[0/90/90/0]s
J	[90/0/0/90]s

Thomas Jin-Chee LIU [10] used the same stacking sequence of 8-ply laminate for the entire bicycle frame. Other studies are more dynamic with having different stacking sequence and number of plies for each section of the bicycle frame. For instance, Hu et al. [11] separated the bicycle frame into two sections: section A with 8-ply laminate consisting of head tube, top tube, seat tube, and bottom bracket and section B with 5-ply laminate consisting of seat stays and chain stays. Jung et al. [12] initially used the same stacking sequence of 8-ply laminate for the entire bicycle frame then separated the bicycle frame based on the weak regions and failure indexes into three sections with different stacking sequences and 16-ply laminate only for section 3. Jung et al. [12] grouped the head tube into Group 1, top tube, seat stays, and chain stays into Group 2, and seat tube, top tube, and bottom bracket into Group 3. Similarly, Chun et al. [13] initially employed the same stacking sequence of 20-ply laminate throughout the bicycle frame and then optimised different stacking sequences for each member of the frame based on the initial results.
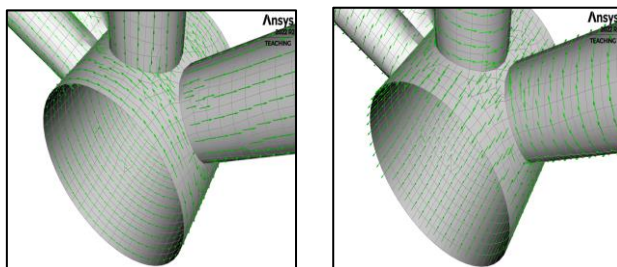
a. 0°   b. 90°
Fig. 2. Fibre orientation angle

2.4. *Loading conditions*

The bicycle frame design and loading conditions applied were based on Sajimsha B et al. [7] work. There were five loading conditions simulated and analysed. All the loading conditions are shown in Fig. 3. The first loading condition is static start-up. It simulates the bicycle in a resting state with a rider sitting on the saddle and the rider is about to start pedalling. A force of 700 N is applied to the top of the seat tube implying the weight of the rider and a force of 200 N is applied to the bottom bracket implying pedalling force. The head tube and rear brackets are fixed supported. The second loading condition is steady state pedalling. It simulates the bicycle is being pedalled out of saddle, meaning that the rider is not sitting on the saddle instead standing up while pedalling. Force of 1000 N is applied to the top of the head tube implying the force of the rider's hands pushing down on the steering and force of 200 N is applied to the bottom bracket implying the pedalling force. The head tube and rear brackets are fixed supported. The third loading condition is vertical impact. It is to simulate an event of impact to the top of the seat tube. Downward force of 2250 N is applied to the top of the seat tube. The head tube and rear brackets are fixed supported. The fourth loading condition is horizontal impact. Similarly to third loading condition, it simulates an event of impact to the top tube. Horizontal force of 2250 N is applied along the top tube. The head tube and rear brackets are fixed supported. The fifth loading condition is rear wheel braking. It simulates the rider is applying brake to the rear wheel when coming to a stop. Horizontal force of 750 N is applied to each of the rear bracket. The bottom bracket is fixed supported.

To analyse whether the bicycle frame fails or not in a loading condition, multiple composite failure theories were selected in composite failure tool on ANSYS. ANSYS calculates the failure theories based on inverse reserve factor (IRF) which tells a number that indicates failure. IRF between 0 and 0.99 indicates that the bicycle frame is still intact while IRF that exceeds 1 indicates the bicycle frame already fails. Composite failure theories can be divided into three groups. Limit or noninteractive theories like maximum stress and maximum strain theories disregard interaction among different stress components and only compare individual ply stresses or strain with the corresponding ultimate strengths or strains [9]. Interactive theories like Tsai-Hill and Tsai-Wu theories include all stress component in one expression without reference to particular failure modes 9. Partially interactive or failure-mode-based theories like Hashin and Puck theories give separate criteria for fibre and interfibre failures [9]. Five failure theories are chosen in the composite failure tool: Maximum stress theory, Tsai-Wu theory, Tsai-Hill theory, Hashin theory, and Puck theory.

(a)   Static start-up

(b)   Steady-state pedalling

(c)   Vertical impact

(d)   Horizontal impact

(e)   Rear-wheel braking
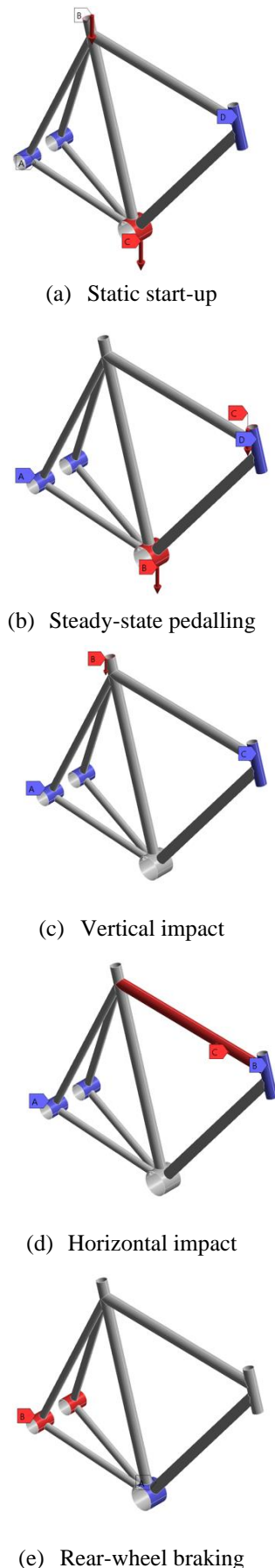
Fig. 3. Loading conditions applied

3.   Results and Discussion

3.1.   *Initial analysis*

The initial equivalent (von-mises) stress results show good or even better performances than those of the conventional materials for some stacking sequences and loading conditions. All the stress values for corresponding stacking sequences and loading conditions are shown in Table 4. For comparison, the equivalent (von-mises) stress results produced by Deepak [14] were listed in Table 5.

Table 4.   Initial equivalent (von-Mises) stress

	Static Start-up	Steady State Pedalling	Vertical Impact	Horizon-tal Impact	Rear Wheel Braking
A	11.242	4.5884	35.034	11.127	40.049
B	11.245	4.5864	35.051	11.128	40.148
C	9.7708	3.7435	30.249	8.0981	**32.523**
D	**9.7704**	3.7417	**30.248**	8.0976	33.298
E	10.248	4.149	31.733	9.3719	34.823
F	10.259	4.2497	31.768	9.3758	35.377
G	10.235	4.1446	31.695	9.4727	38.412
H	10.245	4.2379	31.728	9.4773	38.717
I	11.196	4.6299	34.896	10.994	40.239
J	9.7798	**3.6939**	30.279	**7.9794**	32.685

Table 5.   Equivalent (von-Mises) stress of conventional materials

	Static Start-up	Steady State Pedalli-ng	Vertical Impact	Horizontal Impact	Rear Wheel Brakin-g
Steel	9.8212	6.3773	31.216	28.069	17.746
Aluminium	10.915	7.4270	33.722	30.364	17.780
Titanium	10.896	7.3168	34.607	31.071	18.022
Carbon Fibre	10.940	10.940	34.839	33.941	19.498

In static start-up, stacking sequence C, D, and J have lower stresses than all the conventional materials. Stacking sequence E, F, G, and H have lower stresses than aluminium 6061 T6, titanium grade 9 and carbon fibre while stacking A, B, and I have higher stresses than all the conventional materials. In steady state pedalling, all the stacking sequences show lower stresses than all the conventional materials. In vertical impact, stacking sequence C, D, and J have lower stresses than steel while stacking sequence E, F, G, and H have lower stresses than aluminium 6061 T6, titanium grade 9 and carbon fibre. Stacking sequence A, B, and H have higher stresses than all the conventional materials. In Horizontal impact, all the stacking sequences have lower stresses than all the conventional materials contrary to rear wheel braking

condition which displays that all the stacking sequences have higher stresses than all the conventional materials.

In each of the loading condition, the maximum stress is always concentrated at the same region despite the different stacking sequences. As can be seen in Fig. 4, the stress concentration is at the connection between seat tube and seat stays in static start-up; connection between down tube and bottom bracket in steady state pedalling; connection between seat tube and seat stays in vertical impact; connection between heat tube and top tube in horizontal impact; intersection between chain stays in rear wheel braking.

The bicycle frame will not necessarily fail at the stress concentration region. The region may be able to hold the maximum stress while other regions may not be able to hold their lower stresses. This is further seen with the location of maximum inverse reserve factor.

The initial total deformation results in Table 6 are not comparable to the results of the conventional materials in Table 7. As expected, the deformation in each loading condition and stacking sequence is way higher than the deformations for all the conventional materials. This is due to the epoxy OPEFB composite having way lower tensile and compressive moduli than all the conventional materials. Thus, producing way higher deformations. The deformations in rear wheel braking condition are unacceptable as 6 mm deformation is too high and will possibly affect riding performance and stability. As observed as well the different in deformations between the stacking sequences in each loading condition is insignificant.

Table 6. Initial total deformation induced.

	Static Start-up	Steady State Pedalling	Vertical Impact	Horizon-tal Impact	Rear Wheel Braking
A	0.29322	**0.13132**	0.8982	0.25591	6.8875
B	0.29323	0.1314	0.89823	0.25591	6.8863
C	0.29325	0.13792	0.89828	0.25626	6.8907
D	0.29325	0.13801	0.89829	0.25626	6.8895
E	0.29334	0.13446	0.89861	0.25608	6.8894
F	0.2932	0.13663	0.89812	0.25618	6.8824
G	0.29333	0.1342	0.89859	0.25608	6.8927
H	0.29318	0.13636	0.89808	0.25617	6.886
I	**0.29066**	0.13167	**0.89057**	**0.25316**	**6.8304**
J	0.29083	0.13614	0.89113	0.25341	6.8401

Table 7. Total deformation of conventional materials

	Static Start-up	Steady State Pedalling	Vertical Impact	Horizontal Impact	Rear Wheel Braking
Steel	0.0081404	0.0084963	0.024997	0.030902	0.18428
Aluminium 6061 6T	0.023512	0.024481	0.072276	0.090021	0.52604
Titanium Grade 9	0.017838	0.018578	0.054793	0.068393	0.39739
Carbon Fibre	0.0040985	0.0042668	0.012585	0.015749	0.089395
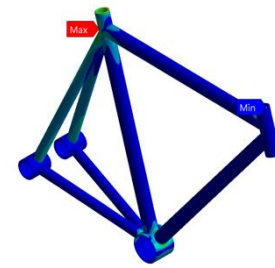
(a) Static start-up
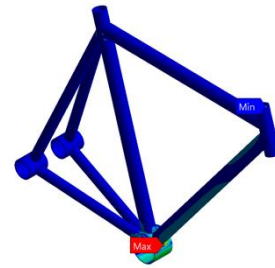
(b) Steady-state pedalling
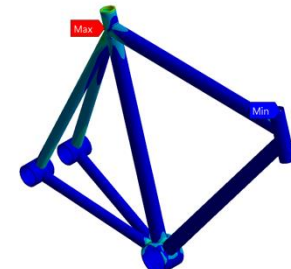
(c) Vertical impact
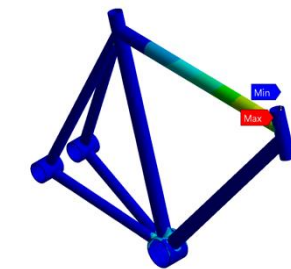
(d) Horizontal impact
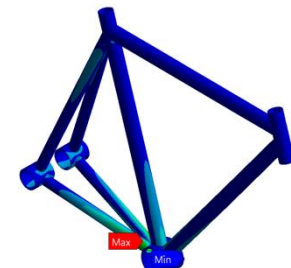
(e) Rear-wheel braking

Fig. 4. Stress concentration of each loading condition

Similarly to stress concentration, the bicycle frame deforms at the same location in each loading condition regardless of the stacking sequences. The bicycle frame always deforms the most at the highest load location if there is no fixed support at the location. The bicycle frame deforms at the tip of the seat tube where force of 700 N is applied in static start-up; bottom bracket where force of 200 N is applied in steady state pedalling; tip of the seat tube where force of 2250 N is applied in vertical impact, connection between top tube and seat tube where force of 2250 N is applied to the top tube in horizontal impact; both rear brackets where force of 750 N is applied to each in rear wheel braking.

Inverse reserve factor (IRF) is an indication if the bicycle frame will fail under the loading conditions. If the inverse reserve factor exceeds 1, the bicycle frame will fail. As can be observed in Table 8, the bicycle frame has relatively low IRF in static start-up and steady state pedalling but high IRF in vertical impact, horizontal impact, and rear wheel braking, indicating that the frame will most likely fail under these loading conditions.

0.8. General trend can be seen here is that stacking sequence A, B, G, H, and I produce lower IRF than all the other stacking sequences in all loading conditions. For initial results, stacking sequence A, B, G, H, and I are chosen to be general good stacking sequences. Table 9 shows the top five stacking sequences in each loading condition and they are composed of stacking sequences A, B, G, H and I except for steady state pedalling that swaps stacking sequence H for E. The (a), (b), (c) and (d) are indicators for the location of the maximum IRF which occurs at only four locations; as shown in Fig. 5. Location a is the connection between seat tube and chain stays; location b is the connection between bottom bracket and seat tube; location c is the connection between bottom bracket and down tube; location d is the chain stays intersection. These locations are also the weak regions of the bicycle frame. The designation inside the bracket beside the IRF value is the location of the IRF. The maximum IRF is greatly influenced by loading conditions but not the stacking sequences except for static start-up which shows variation in the maximum IRF locations based on the stacking sequences.

Table 8. Initial inverse reserve factor

	Static Start-up	Steady State Pedalling	Vertical Impact	Horizon-tal Impact	Rear Wheel Braking
A	0.15252	0.09564	0.52981	0.2719	0.6079
B	0.15182	0.095529	0.54129	0.27586	0.61455
C	0.20657	0.13149	0.80163	0.34544	0.84844
D	0.20922	0.13133	0.81165	0.34912	0.85809
E	0.20234	0.11412	0.80113	0.35434	0.80075
F	0.20584	0.11701	0.81488	0.35861	0.81305
G	0.15299	0.11521	0.5253	0.27102	0.71251
H	0.1565	0.11821	0.53666	0.27416	0.72179
I	0.15032	0.096605	0.53767	0.27396	0.61409
J	0.20484	0.12935	0.7955	0.34325	0.84165

The stacking sequence A, B, G, H, and I have slightly higher IRF than 0.15 compared to stacking sequence C, D, E, F, and J that have IRF higher than 0.2 in static start-up. Stacking sequence A, B and I have IRF lower than 0.1, stacking sequence E, F, G, H have IRF slight higher than 0.11, and stacking sequence C, D, and J have IRF higher than 0.12 in steady state pedalling.

Stacking sequence A, B, G, H, I have slightly higher IRF than 0.5 while stacking sequence C, D, E, F, and J have IRF around 0.8 in vertical impact. Stacking sequence A, B, G, H, and I have IRF slightly higher than 0.27 while Stacking sequence C, D, E, F, and J have IRF higher than 0.34. Lastly, Stacking sequence A, B, and I have slightly higher IRF than 0.6, stacking sequence G and H have IRF slight higher than 0.7, and stacking sequence C, D, E, F, and J have slight higher IRF than

Table 9. Top five stacking sequences of each loading condition for IRF

Static Start-up	
[0/90/90/0]s	0.15032 (a)
[0/90/-45/45]s	0.15182 (a)
[0/90/45/-45]s	0.15252 (a)
[45/-45/0/90]s	0.15299 (c)
[45/-45/90/0]s	0.1565 (c)
Steady State Pedalling	
[0/90/-45/45]s	0.09552 (c)
[0/90/45/-45]s	0.09564 (c)
[0/90/90/0]s	0.09660 (c)
[-45/45/0/90]s	0.11412 (c)
[45/-45/0/90]s	0.11521 (c)
Vertical Impact	
[45/-45/0/90]s	0.5253 (b)
[0/90/45/-45]s	0.52981 (b)
[45/-45/90/0]s	0.53666 (b)
[0/90/90/0]s	0.53767 (b)
[0/90/-45/45]s	0.54129 (b)
Horizontal Impact	
[45/-45/0/90]s	0.27102 (b)
[0/90/45/-45]s	0.2719 (b)
[0/90/90/0]s	0.27396 (b)
[45/-45/90/0]s	0.27416 (b)
[0/90/-45/45]s	0.27586 (b)
Rear Wheel Braking	
[0/90/45/-45]s	0.6079 (d)
[0/90/90/0]s	0.61409 (d)
[0/90/-45/45]s	0.61455 (d)
[45/-45/0/90]s	0.71251 (d)
[45/-45/90/0]s	0.72179 (d)
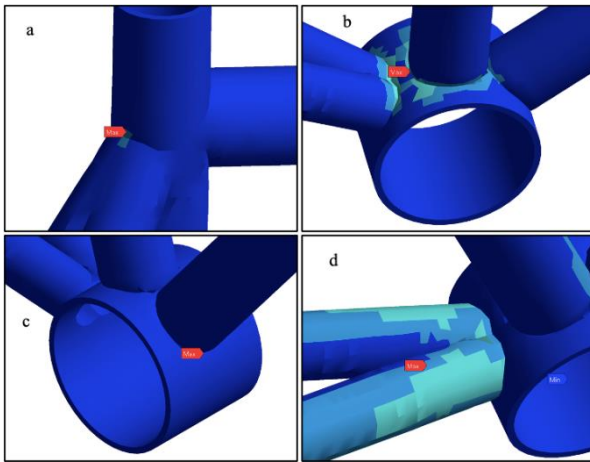
Fig. 5. Locations of maximum inverse reserve factor

3.2. *Optimization*

It was obvious that only one laminate of epoxy OPEFB composite cannot be feasible for a good performance bicycle frame. The deformation was too high compared to those of conventional materials and especially in rear wheel braking with one laminate. The IRFs in vertical impact, horizontal impact, and rear wheel braking were also high with one laminate. To reduce the deformation and the IRF, another laminate of 8-ply was introduced, creating two 8-ply laminates or 16 plies all together for every member of the bicycle frame.

The results, as shown in Table 10, 11 and 12, show reduction in half for all the stresses, deformations, and IRFs in all the loading conditions. The equivalent (von-Mises) stress was averagely reduced by 50.63% in static start-up; 65.39% in steady state pedalling; 49.95% in vertical impact; 41.55% in horizontal impact; 47.44% in rear wheel braking. The deformation was averagely reduced by 57.13% in static start-up; 71.56% in steady state pedalling; 57.68% in vertical impact; 51.79% in horizontal impact; 60.56% in rear wheel braking. The IRF was averagely reduced by 61% in static start-up; 70.98% in steady state pedalling; 59% in vertical impact; 63.69% in horizontal impact; 53.79% in rear wheel braking.

Table 10. Equivalent (von-Mises) stress results for double-laminate bicycle frame

Designation	Stacking Sequence	Equivalent (von-Mises) Stress (MPa)				
		Static Start-up	Steady State Pedalling	Vertical Impact	Horizontal Impact	Rear Wheel Braking
A_2	$[(0/90/45/-45)_2]_s$	5.1762	1.4255	16.136	5.3502	18.09
B_2	$[(0/90/-45/45)_2]_s$	5.1765	1.4256	16.138	5.3502	18.101
G_2	$[(45/-45/0/90)_2]_s$	4.5443	1.2326	14.397	4.5508	17.234
H_2	$[(45/-45/90/0)_2]_s$	4.5447	1.2382	14.458	4.5512	17.242
I_2	$[(0/90/90/0)_2]_s$	5.1473	1.4184	16.047	5.2871	18.002

Table 11. Total deformation results for double-laminate bicycle frame

Designation	Stacking Sequence	Total Deformation (mm)				
		Static Start-up	Steady State Pedalling	Vertical Impact	Horizontal Impact	Rear Wheel Braking
A_2	$[(0/90/45/-45)_2]_s$	0.12546	0.037708	0.37937	0.12315	2.7109
B_2	$[(0/90/-45/45)_2]_s$	0.12546	0.037705	0.37937	0.12315	2.7109
G_2	$[(45/-45/0/90)_2]_s$	0.12554	0.037838	0.37961	0.12317	2.7131
H_2	$[(45/-45/90/0)_2]_s$	0.12557	0.037923	0.37972	0.12319	2.7139
I_2	$[(0/90/90/0)_2]_s$	0.12437	0.037415	0.37618	0.1218	2.6907

Table 12. Inverse reference factor for double-laminate bicycle frame

Designation	Stacking Sequence	Inverse Reference factor				
		Static Start-up	Steady State Pedalling	Vertical Impact	Horizontal Impact`	Rear Wheel Braking
A_2	$[(0/90/45/-45)_2]_s$	0.060172 (b)	0.029942 (c)	0.2185 (b)	0.097273 (b)	0.30854 (d)
B_2	$[(0/90/-45/45)_2]_s$	0.060416 (b)	0.029935 (c)	0.21947 (b)	0.097597 (b)	0.30946 (d)
G_2	$[(45/-45/0/90)_2]_s$	0.0583 (b)	0.030153 (c)	0.21787 (b)	0.10033 (b)	0.29258 (d)
H_2	$[(45/-45/90/0)_2]_s$	0.058772 (b)	0.030249 (c)	0.21958 (b)	0.10081 (b)	0.29281 (d)
I_2	$[(0/90/90/0)_2]_s$	0.060352 (b)	0.029821 (c)	0.21942 (c)	0.097409 (c)	0.30814 (d)

While all the stresses, deformations, and IRFs were reduced half by just adding an additional laminate, the bicycle frame still had the same stacking sequence throughout. Different stacking sequences for different members of the bicycle frame is proven to be beneficial for sustaining different loading conditions. Therefore, based on the weak regions, the bicycle frame was sectioned into three sections: section 1 consists of top tube, head tube, and down tube; section 2 consists of seat tube alone; section 3 consists of seat stays, chain stays, bottom bracket, and rear brackets. Section 1 would employ stacking sequence B_2 and I_2, section 2 would employ stacking sequence A_2, B_2, G_2, and I_2, and section 3 would employ stacking sequence A_2, and H_2. The combinations of all the stacking sequences of each section are clearly listed in Table 13.

Table 13. Stacking sequences for section 1, 2, and 3

Stac-king	Section 1	Section 2	Section 3
K	$[(0/90/90/0)_2]_s$	$[(0/90/90/0)_2]_s$	$[(45/-45/90/0)_2]_s$
L	$[(0/90/90/0)_2]_s$	$[(0/90/90/0)_2]_s$	$[(0/90/45/-45)_2]_s$
M	$[(0/90/90/0)_2]_s$	$[(0/90/-45/45)_2]_s$	$[(45/-45/90/0)_2]_s$
N	$[(0/90/90/0)_2]_s$	$[(0/90/-45/45)_2]_s$	$[(0/90/45/-45)_2]_s$
O	$[(0/90/90/0)_2]_s$	$[(0/90/45/-45)_2]_s$	$[(45/-45/90/0)_2]_s$
P	$[(0/90/90/0)_2]_s$	$[(0/90/45/-45)_2]_s$	$[(0/90/45/-45)_2]_s$
Q	$[(0/90/90/0)_2]_s$	$[(45/-45/0/90)_2]_s$	$[(45/-45/90/0)_2]_s$
R	$[(0/90/90/0)_2]_s$	$[(45/-45/0/90)_2]_s$	$[(0/90/45/-45)_2]_s$
S	$[(0/90/-45/45)_2]_s$	$[(0/90/90/0)_2]_s$	$[(45/-45/90/0)_2]_s$
T	$[(0/90/-45/45)_2]_s$	$[(0/90/90/0)_2]_s$	$[(0/90/45/-45)_2]_s$
U	$[(0/90/-45/45)_2]_s$	$[(0/90/-45/45)_2]_s$	$[(45/-45/90/0)_2]_s$
V	$[(0/90/-45/45)_2]_s$	$[(0/90/-45/45)_2]_s$	$[(0/90/45/-45)_2]_s$
W	$[(0/90/-45/45)_2]_s$	$[(0/90/45/-45)_2]_s$	$[(45/-45/90/0)_2]_s$
X	$[(0/90/-45/45)_2]_s$	$[(0/90/-45/45)_2]_s$	$[(0/90/45/-45)_2]_s$
Y	$[(0/90/-45/45)_2]_s$	$[(45/-45/0/90)_2]_s$	$[(45/-45/90/0)_2]_s$
Z	$[(0/90/-45/45)_2]_s$	$[(45/-45/0/90)_2]_s$	$[(0/90/45/-45)_2]_s$

Equivalent (von-Mises) stress results, as shown in Table 14, show insignificant difference between the stacking sequences except for a few stacking sequences in some loading condition. Stacking sequence Q and Y have lower stresses slightly above 4.5 MPa compared to the rest of the stacking sequences that have stresses slightly above 5 MPa in static start-up. All the stacking sequences in steady state pedalling have stresses slightly above 1.4 MPa and the difference between each other is insignificant. Stacking sequence Q and Y have stresses slightly above 14.3 MPa, stacking sequence R and Z have stresses slightly above 15.5 MPa, and all other stacking sequences have stresses slightly above 16 MPa in vertical impact. In horizontal impact, stacking sequence L, K, N,

P, R, M, O, and Q have stresses slight above 5.28 MPa and stacking sequence T, S, V, X, Z, and U have stresses slightly above 5.35. In rear wheel braking, stacking sequence K, S, M, O, Q, U, W, and Y have stresses slightly above 17.24 MPa and stacking sequence V, X, Z, N, P, R, T, and L have stresses slightly above 18.09 MPa. The stresses are not improved over the stresses of the same case of double-laminate bicycle frame.

Table 14. Equivalent (von-Mises) stress results for double-laminate sectioned bicycle frame

	Static Start-up	Steady State Pedalling	Vertical Impact	Horizon-tal Impact	Rear Wheel Braking
K	5.1354	1.4279	16.009	5.2887	17.241
L	5.1469	1.4218	16.044	5.2881	18.093
M	5.1607	1.429	16.091	5.2892	17.242
N	5.1722	1.4229	16.126	5.2887	18.091
O	5.159	1.429	16.085	5.2892	17.242
P	5.1709	1.4229	16.121	5.2887	18.091
Q	4.5435	1.429	14.397	5.2893	17.242
R	5.0339	1.4229	15.543	5.2887	18.091
S	5.1407	1.4306	16.023	5.3502	17.241
T	5.1521	1.4245	16.059	5.3497	18.091
U	5.166	1.4316	16.105	5.3508	17.242
V	5.1775	1.4255	16.141	5.3502	18.09
W	5.1643	1.4316	16.1	5.3508	17.242
X	5.1762	1.4255	16.136	5.3502	18.09
Y	5.1354	1.4279	16.009	5.2887	17.241
Z	5.1469	1.4218	16.044	5.2881	18.093

Similar trend is observed for the total deformation for double-laminate sectioned bicycle frame, as shown in Table 15. The deformation difference between the stacking sequences in each loading condition is very insignificant. The deformations do not see any improvement over the deformations of the same case of double-laminate bicycle frame. IRFs also do not see any improvement over the IRFs of the same case of double-laminate bicycle frame. The IRF difference between stacking sequences in each loading condition is very insignificant, as shown in Table 16.

Table 15. Total deformation results for double-laminate sectioned bicycle frame

	Static Start-up	Steady State Pedalling	Vertical Impact	Horizon-tal Impact	Rear Wheel Braking
K	0.12515	0.037681	0.37848	0.12185	2.7085
L	0.1251	0.037485	0.3783	0.12181	2.7057
M	0.12536	0.037711	0.37918	0.12188	2.7113
N	0.12531	0.037515	0.379	0.12184	2.7085
O	0.12536	0.037709	0.37918	0.12188	2.7113
P	0.12531	0.037514	0.379	0.12184	2.7085
Q	0.1254	0.037709	0.37929	0.12188	2.7114
R	0.12534	0.037514	0.37911	0.12184	2.7086
S	0.1253	0.037874	0.37884	0.12315	2.7108
T	0.12525	0.037678	0.37867	0.12311	2.708
U	0.12552	0.037904	0.37955	0.12318	2.7137
V	0.12546	0.037708	0.37937	0.12315	2.7109
W	0.12552	0.037903	0.37954	0.12319	2.7137
X	0.12546	0.037707	0.37937	0.12315	2.7109
Y	0.12555	0.037902	0.37965	0.12319	2.7138
Z	0.1255	0.037707	0.37948	0.12315	2.711

Table 16. Inverse reserve factor for double-laminate sectioned bicycle frame

	Static Start-up	Steady State Pedalling	Vertical Impact	Horizon-tal Impact	Rear Wheel Braking
K	0.05958 (b)	0.029985 (c)	0.22218 (b)	0.10058 (b)	0.29283 (d)
L	0.06072 (b)	0.029894 (c)	0.22028 (b)	0.09677 (b)	0.3086 (d)
M	0.05919 (b)	0.03 (c)	0.22057 (b)	0.10025 (b)	0.29282 (d)
N	0.06034 (b)	0.02991 (c)	0.21869 (b)	0.09647 (b)	0.30858 (d)
O	0.05918 (b)	0.03 (c)	0.22061 (b)	0.10024 (b)	0.29282 (d)
P	0.06034 (b)	0.029909 (c)	0.21873 (b)	0.09645 (b)	0.30857 (d)
Q	0.05900 (b)	0.03 (c)	0.22007 (b)	0.10006 (b)	0.29282 (d)
R	0.06016 (b)	0.029909 (c)	0.21818 (b)	0.09626 (b)	0.30857 (d)
S	0.05942 (b)	0.030027 (c)	0.22194 (b)	0.10144 (b)	0.29282 (d)
T	0.06056 (b)	0.029935 (c)	0.22004 (b)	0.09759 (b)	0.30857 (d)
U	0.05902 (b)	0.030042 (c)	0.22033 (b)	0.1011 (b)	0.29281 (d)
V	0.06017 (b)	0.029951 (c)	0.21845 (b)	0.09729 (b)	0.30855 (d)
W	0.05902 (b)	0.030042 (c)	0.22037 (b)	0.10109 (b)	0.29281 (d)
X	0.0601 (b)	0.02995 (c)	0.21849 (b)	0.09727 (b)	0.30854 (d)
Y	0.0595 (b)	0.029985 (c)	0.22218 (b)	0.10058 (b)	0.29283 (d)
Z	0.0607 (b)	0.029894 (c)	0.22028 (b)	0.09677 (b)	0.3086 (d)

Overall, different stacking sequence for each section of the bicycle frame does not produce any improvement at all for epoxy OPEFB composite mountain bicycle frame in this study. Major improvement of equivalent (von-Mises) stresses, total deformations, and inverse reserve factor are seen with addition of laminate or ply. Unlike other studies which found that major improvement of performance by applying different stacking sequence to different member of the bicycle frame, all the other studies used synthetic fibre such as carbon fibre and glass fibre which, in a composite, possess big difference between the Young's modulus 1 and 2, and tensile strength 1 and 2. As calculated by using Whitney and Riley theory, epoxy OPEFB composite does not have very big difference between the properties in principal 1 and 2 direction. Therefore, the fibre orientation of epoxy OPEFB composite in each section of the bicycle frame does not contribute much to the performance compared to the fibre orientation of the entire bicycle frame which saw some improvement in the initial results.

4. Conclusion

The fibre orientation and stacking sequences of epoxy OPEFB composite on mountain bicycle frame have been studied thoroughly. A total of 10 stacking sequences were adopted from previous research. It was found that of the

10 stacking sequences, five stacking sequences, [0/90/45/-45]ₛ, [0/90/-45/45]ₛ, [45/-45/0/90]ₛ, [45/-45/90/0]ₛ, and [0/90/90/0]ₛ, were found to be general good stacking sequences. But the deformations and IRFs of the five stacking sequences in some loading conditions were still high. Therefore, the additional laminate was introduced and the results showed significant reduction by half in stresses, deformations, and IRF. The bicycle frame was also sectioned for employment of different stacking sequence in each section but the results showed no improvement. Hence, for epoxy OPEFB composite bicycle frame, additional laminate or ply will improve the performance significantly but not employment of different stacking sequence in each member of the frame.

Acknowledgements

The authors would like to express their gratitude towards the engineering faculty of UCSI University.

References

1. P. Pradhan, L. Costa, D. Rybski, W. Lucht, and J.P. Kropp, A Systematic Study of Sustainable Development Goal (SDG) Interactions, Earth's Future, 5, 2017, pp. 1169-1179.
2. M.K. Faizi, A.B. Shahriman, M.S. Abdul Majid, B.M.T. Shamsul, Y.G. Ng, S.N. Basah, E.M. Cheng, M. Afendi, M.R. Zuradzman, Khairunizam Wan and D. Hazry, An Overview of the Oil Palm Empty Fruit Bunch (OPEFB) Potential as Reinforcing Fibre in Polymer Composite for Energy Absorption Applications, MATEC Web of Conferences, 2017, 90. 01064.
3. F. E. Gunawan, H. Homma, S.S. Brodjonegoro, A.B. Hudin and A. Zainuddin, Mechanical Properties of Oil Palm Empty Fruit Bunch Fiber, Journal of Solid Mechanics and Materials Engineering, 3(7), 2009, pp. 943-951.
4. M.Z. Mohamed Yusoff, S.M Sapuan, and N. Ismail, Tensile Properties of Single Oil Palm Empty Fruit Bunch (OPEFB) Fibre, Sains Malaysiana. 38(4), 2009, pp. 525–529.
5. C.S. Hassan, C.W. Yeo, B.B. Sahari, S.M. Sapuan, and N. Abdul Aziz, Mechanical Properties of Unidirectional Oil Palm Empty Fruit Bunch (OPEFB) Fiber Reinforced Epoxy Composite, IOP Conference Series: Materials Science and Engineering, 206, 2017, 012045.
6. B, Sajimsha, Analysis of Mountain Bike Frames by ANSYS. International Journal for Research in Applied Science and Engineering Technology, 7, 2019, pp. 1167-1175.
7. J.M. Whitney and M.B. Riley, (1996). Elastic Properties of Fibre Reinforced Composite Materials, AIAA Journal, 4(9), 1966, pp. 1537–1542.
8. A.P.S. Selvadurai and H. Nikopour, Transverse Elasticity of a Unidirectionally Reinforced Composite with an Irregular Fibre Arrangement: Experiments, Theory and Computations, Composite Structures. 94(6), 2012, pp. 1973–1981.
9. M.D. Isaac and I. Ori, Engineering Mechanics of Composite Materials, 2nd Edition, Oxford University Press, 2006.
10. T.J-C. Liu and H-C. Wu, Fiber Direction and Stacking Sequence Design for Bicycle Frame Made of Carbon/Epoxy Composite Laminate, Materials & Design, 31(4), 2010, pp. 1971-1980.
11. Y. Hu, Y. Xiao, W. Shang, and J. Zhang, Effect of Fibre Direction and Stacking Sequence on Dynamic Impact Performance of Composite Bicycle Frame, International Journal of Crashworthiness, 22(5), 2017, pp. 556-564.
12. H. Jung, J.A. Lee, K.Y. Kim, and H. Chun, Determination of Number of Layers and Stacking Sequence of Composite Bicycle Frame., Advanced Materials Research, 123-125, 2010, pp. 531-534.
13. H.J. Chun, J.A. Lee, and K.T. Kang, Study of Design Variable with Loading Condition for Composite Laminate Bicycle Frame. In ECCM 2012 - Composites at Venice, Proceedings of the 15th European Conference on Composite Materials (ECCM 2012 - Composites at Venice, Proceedings of the 15th European Conference on Composite Materials), European Conference on Composite Materials, ECCM, 2012.
14. P. Sarath, A. Deepak, H, Hrishikesh, N.S. Daniel and Jinuchandran, Stress Analysis of Bicycle Frame using Different Materials by FEA, GRD Journals for Engineering, 6(7), 2021, pp. 14-20.

Authors Introduction

Mr. Too Kok Sem

He received his Bachelor's degree in Mechanical Engineering in 2022 from the Faculty of Engineering, Technology & Built Environment, UCSI University, Malaysia.

Dr. Cik Suhana Hassan

She received her bachelor's and master's degrees in 2009 and 2011, respectively, from Universiti Teknologi PETRONAS, and her PhD in 2019 from Universiti Putra Malaysia. She is currently an Assistant Professor at the Department of Mechanical and Mechatronics Engineering of UCSI University. Her research interests include the investigation of bio-composites characteristics for applications, particularly automotive.

Ms. Nor Fazilah Abdullah

She received her Bachelor's degree in aerospace engineering (Hons) from IIUM, Gombak in 2010 and her Master's degree in Mechanical Engineering from UKM, Bangi in 2015. Currently she is pursuing Doctoral of Philosophy programme at UCSI University, Kuala Lumpur. Her research interest in bio-based nanoparticles materials.

Ammar Abdulaziz Al Talib

Dr. Ammar Al Talib has finished his B.Sc and M.Sc degrees in Mechanical Engineering from the University of Mosul/Iraq.
He has finished his Ph.D degree from UPM University / Malaysia.
He is also is a Chartered Engineer and Member of the Institute of Mechanical Engineers / UK. (CEng. MIMechE). He has developed all the Postgraduate Programs at the Faculty of Engineering at UCSI University / Malaysia and worked as the Head of Postgraduate and Research department at the same faculty for the years 2010-2018.

Modelling of Short-Circuit Protection for A Residential Grid-Connected BESS

Kong De Kang
Faculty of Engineering, Technology and Built Environment,
UCSI University, Kuala Lumpur, 56000, Malaysia

Farah Adilah Jamaludin
Faculty of Engineering, Technology and Built Environment,
UCSI University, Kuala Lumpur, 56000, Malaysia

Rodney H.G. Tan
Faculty of Engineering, Technology and Built Environment,
UCSI University, Kuala Lumpur, 56000, Malaysia

E-mail: farahadilah@ucsiuniversity.edu.my

Abstract

This research paper presents the power protection study on a grid-connected Battery Energy Storage System (BESS) in a typical Malaysia low-voltage (LV) residential network. The BESS model and control algorithm is developed in MATLAB/Simulink environment. The BESS model can charge and discharge its energy with an algorithm-controlled bidirectional AC/DC converter. The paper also presents two cases for BESS short circuit evaluation to investigate stability of internal and external systems of BESS. The protection level on BESS is also optimized by introducing a time-delay characteristics model to coordinate circuit breakers in compliance with standards outlined in IEEE Std 1375-1998, IEC/EN 60898-2, IEC/EN 60947-2, UL 1077, and CSA 22.2. No.235. The paper presents BESS system stability with and without overcurrent protection. As a conclusion, BESS was interrupted within stipulated time. In case of internal short circuit protection failure, a backup protection will act on isolating BESS from the grid provided

Keywords: Battery Energy Storage System, Fault analysis, Power system stability, Protection scheme, Optimized protection

1. Introduction

Lately, more than 200 energy storage demonstration projects, at megawatt level and above, have been set up across the world, indicating huge potential for further expanded application of 154 Grid-scale energy storage systems and applications energy storage technologies [1]. According to U.S. department of energy, ESS increases grid resistant to disruptions, promotes clean electricity, and increases the economic value of wind and solar power [2]. ESS is also essential to provide power at optimal response speed to overcome the intermittence of renewable sources and contribute to the mains voltage regulation [3]. There are various types of ESS technologies proposed by [4] used for transportation, emergency, and small-to-large scale power generation. BESS may be the most versatile of the several storage systems according to [5].

BESS has three primary components: an AC grid, a battery, and a battery management system [6]. The kind of battery used in the system determines the power output and energy levels of BESS. Various research has been previously conducted for modelling of BESS. In [7], the author has proposed a detailed modelling of BESS using various power electronic components and control modules have been discussed. On the other hand, a simplified BESS model using MATLAB/Simulink logical-numerical modelling approach has been proposed by authors in [8] that can simulate the behaviors of a typical BESS.

Kong De Kang, Farah Adilah Jamaludin, Rodney H.G. Tan

Several issues have been faced by BESS due to its fault behaviours. The primary issue is the high fault current contributed from BESS [9]. The second issue is the rate of current rising [10], [11]. Regarding the two issues, the power electronic switches used with BESS may also be harmed by the current and its increasing rate [12]. The sudden decline in battery state of charge (SOC), which had an adverse effect on battery lifetime, is the final issue [13]. BESS can pose a significant safety risk to both users and installers, with the potential to cause hazards such as fire, flash burns, and explosion if it suffers from short circuit or fault if installed and used improperly [14]. Short circuits are a major contributor to the electrical and temperature risks brought on by Li-ion batteries [15]. Accordingly, it is crucial to achieve power protection of BESS to ensure its proper functions during critical time. Therefore, two main areas of BESS protection are protection of BESS itself for internal fault and isolating BESS from a system for external fault [6]. Several options to protection of stationary battery systems such as circuit breakers and fuses have also been provided in [16] to DC system designer. The paper presents an LV grid-connected BESS with and without overcurrent protection of both internal and external circuit of BESS alongside with fault analysis.

2. Methodology

2.1. *System Design*

As displayed in Fig. 1, the BESS is connected to a typical Malaysia LV network. The connection between BESS and Grid is straight forward since no load is connected to the network bus. BESS consists of a battery module and a bidirectional AC/DC converter. The BESS has utilized a battery module from manufacturer's datasheet. The battery module is referred to technical specifications outlined in manufacturer's datasheet. With the aid of a converter, the BESS can charge or discharge the battery at constant voltage and current. The BESS is protected by internal and external protection units against internal and external short circuit faults respectively. Short circuit and power system stability can be analyzed through DC and AC protection breaker monitoring interface.

research, similar fault locations have been determined in a block diagram as shown in Fig. 2. Three-phase fault blocks from MATLAB/Simulink library are placed
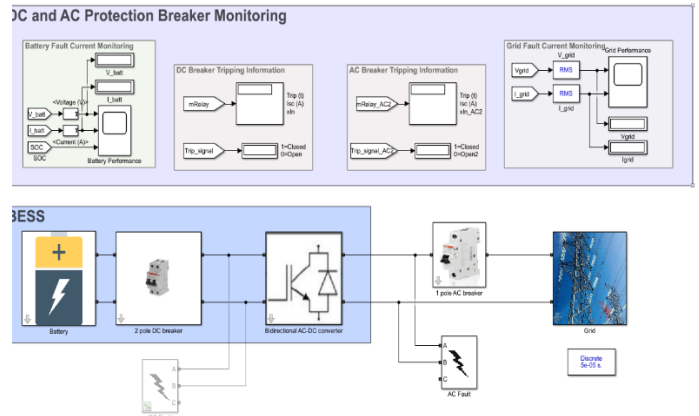
Fig. 1 Overview of BESS system protection model in Simulink

2.2. *BESS model*

The three primary components of BESS are an AC grid, a battery, and a battery management system as presented in [6]. The BESS consists of a Li-Ion battery and a bidirectional AC/DC converter. The nominal voltage and ratings of the battery are 450V and 33.33Ah respectively. The converter is designed using an average model based VSC as the converter bridge. A subsystem that consists of a series of mathematical blocks is built to generate pulse-width modulation (PWM) signal for the converter output as well as regulating the grid current. In addition, the subsystem also provides charging and discharging modes for BESS through current settings. The converter circuit has a DC link capacitor for smoothing the DC voltage as well as series parallel LCL filter to reduce the Total Harmonic Distortion (THD) of AC waveforms. Being the main charging source for BESS, AC grid is designed at 230V which is a typical LV residential network level in Malaysia. Simulation is run to analyse the behaviours of BESS when there is no short circuit.

2.3. *BESS short circuit evaluation*

According to previous research on protection of BESS in a DC microgrid, fault on the bus at BESS side and fault on DC terminals at battery side are the two fault locations as identified in [12]. Based on the project model in this

across the identified fault locations as shown in Fig. 3. The fault blocks are set to provide line-to-line fault as F1 and line-to-line-to-ground fault as F2. F1 and F2 are

identified as internal fault and external fault, respectively. The simulation is run to analyze the short circuit behaviors of BESS in terms of voltage, current, and SoC characteristics of its battery. The simulation is repeated

with several short-circuit resistance adjusted in the fault block parameter settings to observe the impacts of different network short-circuit levels on system stability of BESS and AC grid.
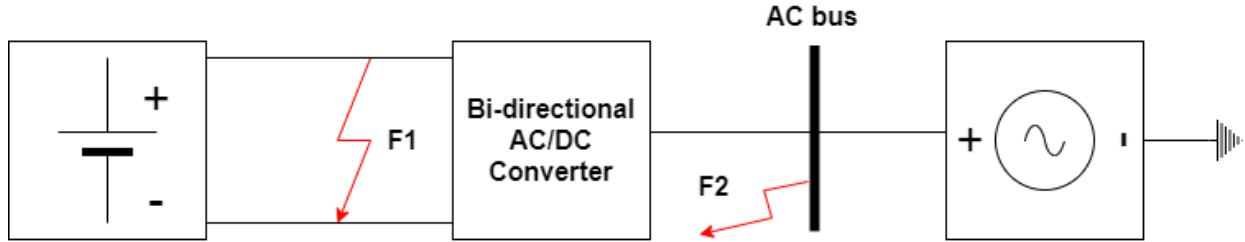
Fig. 2 Block diagram of short-circuited BESS network
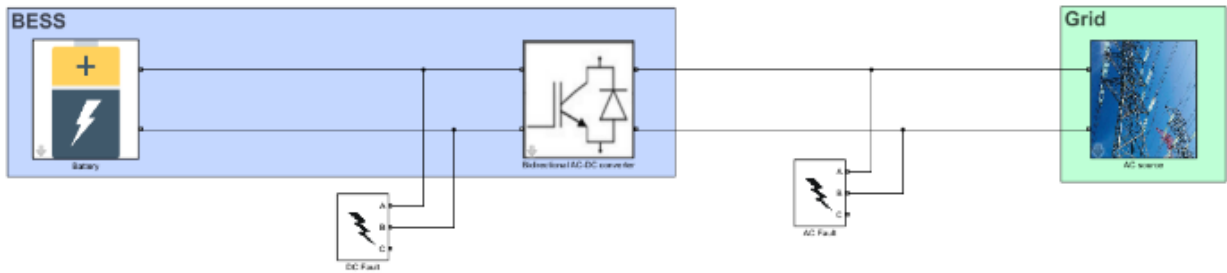
Fig. 3 A short-circuited BESS model.

2.4 Proposed design of BESS protection scheme with optimization method

Circuit breaker as a main protection scheme has been provided in [6]. However, AC breaker for DC use must be modified by adjusting the magnetic trip characteristics of an AC circuit breaker [16]. In this paper, a combination of AC breaker (ACCB) and DC breaker (DCCB) are used for the external and internal power protection of BESS, respectively as shown in Fig 4. During an external fault, F2, BESS will be isolated from AC bus whereas during an internal fault, F1, battery will be isolated within the BESS system. An ideal switch is used as a basic circuit breaker to isolate battery or BESS from the short circuit when the switch is opened. A circuit breaker trips according to its trip characteristics curve based on the guidelines provided in [16]. A model that simulates B and C trip characteristics curves has been developed using look-up table block from MATLAB/Simulink library. The trip characteristics used are referenced from a manufacturer datasheet for a miniature circuit breaker (MCB). As a common practice, B type curve is employed in DCCB whereas C type curve is employed in ACCB.

Kong De Kang, Farah Adilah Jamaludin, Rodney H.G. Tan
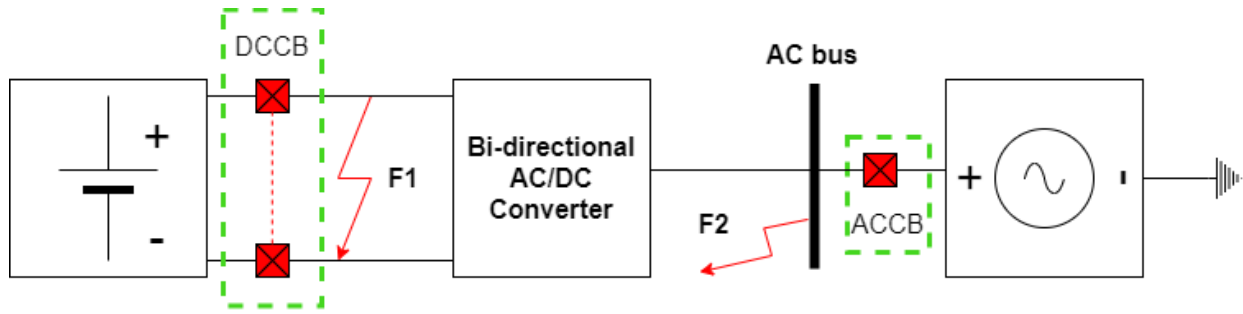
Fig. 4 Locations of circuit breaker.

Once the trip characteristics curves are verified, the curve model is employed in a subsystem. This subsystem is built to generate signal as if it was a "relay" for opening the circuit breaker. Some of the main parameters like current ratings and breaking capacity are defined in this system. Trip time limit is used to define the maximum and minimum trip time as indicated from trip curve from manufacturer datasheet. The current ratings of a circuit breaker are defined in [16]. Based on parameters of generic battery model predefined from MATLAB/Simulink, BESS can discharge current at approximate 14.5A. Whereas based on simulation, the AC bus is regulated by the converter at approximate 6A. Therefore, the DCCB and ACCB current ratings from calculation are 18.1A and 7.5A respectively. The current ratings selection should be at least or above the calculated values for common practice. From the available range of current ratings found in datasheet, DCCB and ACCB current ratings are determined as 20A and 8A respectively. In trip signal generator, MATLAB functional codes are developed to compare the current detected on the system the breaker is connected to with current multiples of the trip curve. A time-delayed signal that indicates trip will be sent to circuit breaker model to open the switch. The tripping sequence of the circuit breaker as coordinated by the trip signal generator subsystem is summarized in a flowchart shown in Fig 5. The trip curve model with trip signal generator subsystem can optimize the protection scheme by introducing delayed tripping in circuit breaker model approved with standards outlined in IEEE Std 1375-1998, IEC/EN 60898-2, IEC/EN 60947-2, UL 1077, and CSA 22.2. No.235. After installing the protection schemes in the project model, system stability with protection is analyzed.
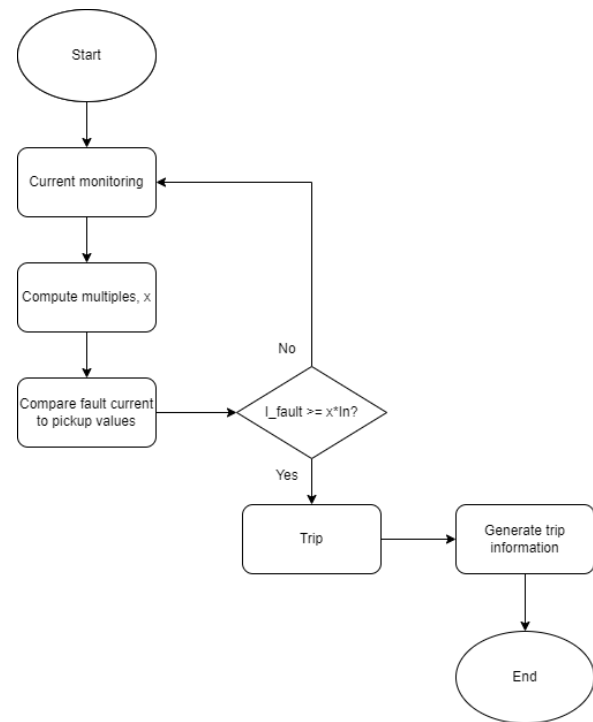
Fig. 5 Tripping sequence of a circuit breaker model.

3. Results and discussion

3.1 BESS behaviors under normal circumstances

Under normal circumstances, the voltage, current, and SoC of BESS are displayed on graphs shown in Fig 6(a) and Fig 6(b). From the graphs, BESS is operating at a constant voltage of 485V which is higher than its nominal voltage, 450V due to other settings assumed by default in generic battery model. However, the operating voltage is still within operating range based on datasheet. The BESS is also charging and discharging at 3A which is lower than current source from the grid due to impedance losses. The current when BESS is discharging opposes the polarity of current during charging as the direction of current flow differs in both states. The initial spike of current is due to initialization of the simulation. However, it does not effect on system stability as it happens only within 0.05s. SoC of BESS is increasing steadily during charging state whereas decreasing steadily during discharging state.
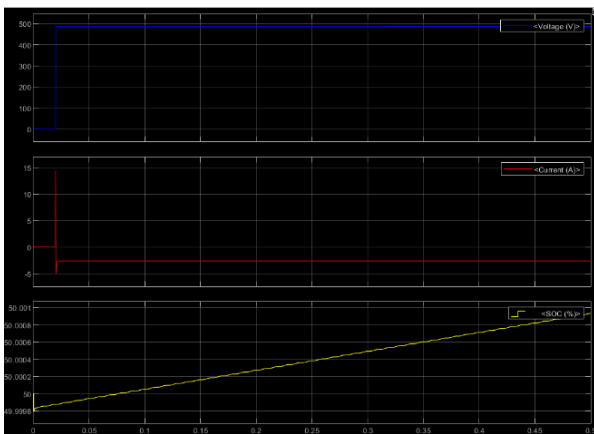
Fig. 6(a) BESS when charging.
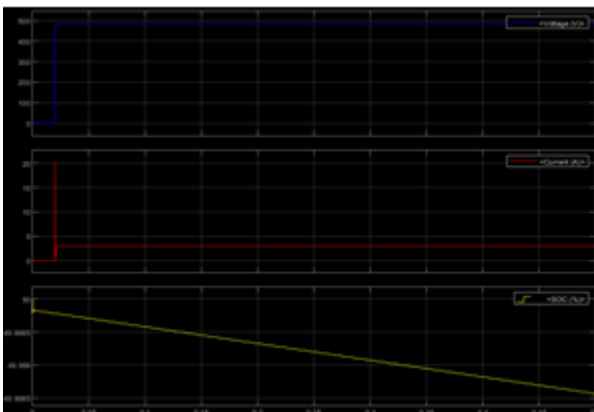
Fig. 6(b) BESS when discharging.

As seen from Fig 6(a) and Fig 6 (b), after a fault occurs at 0.1s, the battery starts to discharge quickly as observed

from the graph of SoC during BESS charging and discharging states. Moreover, the voltage across the terminals of battery begins to decrease due to voltage drop through the resistance. The gradual decrease in voltage is caused by chemical changes occurring at the surface of electrodes. Furthermore, the battery current rises rapidly within 20ms to its peak at approximate 1.4kA. The gradual increase in the current is due to battery inductance.

3.2. Analysis of BESS under external short circuit

From the graphs in Fig 7(a) and Fig 7(b), voltage of battery remains stable after external short circuit effects at 0.1s. Therefore, the battery voltage does not drop due to lower resistance on external circuit of BESS. Moreover, the battery current decreases in both charging and discharging states of BESS. Furthermore, the rate of charge and discharge of BESS also decreases. Therefore, it is said that BESS does not display several issues as discussed in [9]–[13] by neither having high discharged current, voltage drop, nor sharp decrease in SoC except a decrease in power efficiency or performance of BESS. The converter shows current blocking effect by functioning as a diode rectifier and pass the current through only the reverse diodes. The grid filter between ac grid and converter then limits the fault current through the diodes [10].
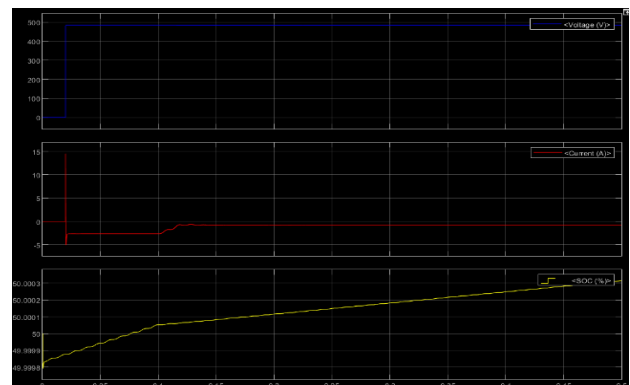
Fig. 7(a) External short circuit of BESS during charging.

Kong De Kang, Farah Adilah Jamaludin, Rodney H.G. Tan
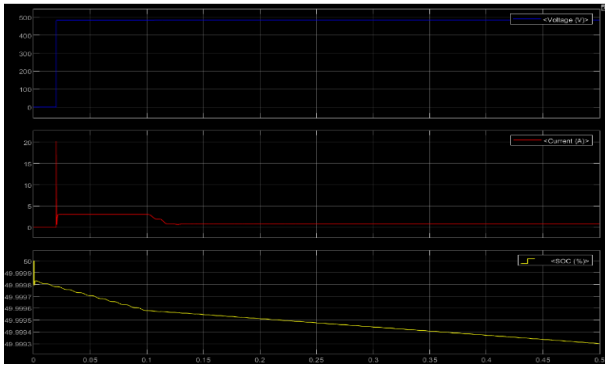
Fig. 7(b) External short circuit of BESS during discharging.
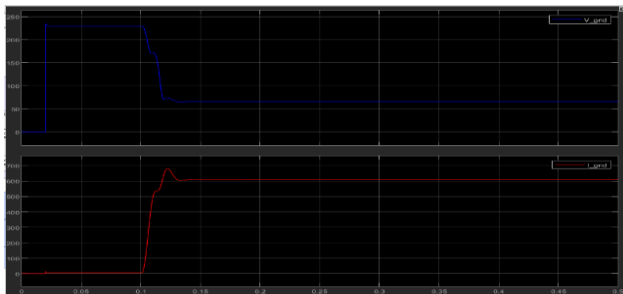
Fig. 8 Grid stability during BESS external short circuits under charging and discharging of BESS

As seen from Fig 8, grid voltage level drops to approximate 60V. Moreover, the grid current increases rapidly to approximate 600A. The gradual decrease and increase in voltage and current are due to grid impedance. The short circuit occurs on the AC bus which is also common point of coupling to the grid. Therefore, AC grid is directly affected by the short circuit.

4. Conclusion

A grid-connected BESS model applicable to a typical LV network in Malaysia has been developed in MATLAB/Simulink environment. The BESS can charge or discharge energy by adjusting the current settings in the converter block to achieve AC/DC bidirectional conversion. Short circuit analysis has also been conducted on the internal and external circuit of BESS. In the internal circuit, faults are conducted across the terminals of battery. Whereas in the external circuit, fault is conducted on the AC bus which is point of common couplings for BESS terminals and AC grid. Based on the fault analysis, internal short circuit causes high current discharged from BESS alongside with voltage drop and rapid decrease of SoC. These behaviors are damaging to the battery lifespan of BESS. On the other hand, external

short circuit does not cause BESS to discharge high current but operating at lower power efficiency instead due to current blocking effect of the converter. The protection scheme using a combination of ACCB and DCCB is designed for BESS power protection. The trip characteristics curve models referenced from manufacturer datasheet are developed to optimize the circuit breakers in compliance with standards outlined in IEEE Std 1375-1998, IEC/EN 60898-2, IEC/EN 60947-2, UL 1077, and CSA 22.2. No.235 according to manufacturer datasheet. AC breaker is also adjusted for DC application in the modelling. Then, fault cases are conducted son BESS to investigate the BESS system stability with power protection. The simulation results show that BESS operation is successfully interrupted in 0.02s by DCCB during an internal short circuit event. In addition, ACCB trips with longer time delay to internal fault when DCCB fails to trip. Therefore, ACCB acts as secondary protection in case of DCCB failure. During an external short circuit event, ACCB will trip to protect BESS from AC bus fault by means of BESS isolation. DCCB will never trip because the converter diodes and AC LCL filter have limited the fault current passing through converter.

Acknowledgement

References

1. Integrated ESS application and economic analysis, Grid-scale Energy Storage Systems and Applications, pp. 153–201, Jan. 2019.
2. Energy Storage: The Key to a Reliable, Clean Electricity Supply | Department of Energy. 2022.
3. A. Stoppato and A. Benato, The Importance of Energy Storage, pp. 1–26, Aug. 2017.
4. H. Abdi, B. Mohammadi-ivatloo, S. Javadi, A. R. Khodaei, and E. Dehnavi, Energy Storage Systems, Distributed Generation Systems: Design, Operation and Grid Integration, pp. 333–368, Jan. 2017.
5. C. D. Parker, APPLICATIONS – STATIONARY | Energy Storage Systems: Batteries, Encyclopedia of Electrochemical Power Sources, pp. 53–64, Jan. 2009.
6. R. Hedding and P. Hayes, Protection of battery energy storage systems, 2011 64th Annual Conference for Protective Relay Engineers, pp. 155–159, 2011.
7. X. Xu, M. Bishop, O. Donna G, and H. Chen, "Application and modeling of battery energy storage in power systems," CSEE Journal of Power and Energy Systems, vol. 2, no. 3, pp. 82–90, Sep. 2016.
8. R. H. G. Tan and G. K. Tinakaran, Development of battery energy storage system model in MATLAB/Simulink,

International Journal of Smart Grid and Clean Energy, pp. 180–188, 2020.

9. M. Monadi, C. Gavriluta, A. Luna, J. I. Candela, and P. Rodriguez, Centralized protection strategy for medium voltage DC microgrids, IEEE Transactions on Power Delivery, vol. 32, no. 1, pp. 430–440, Feb. 2017.

10. D. Salomonsson, L. Söder, and A. Sannino, Protection of low-voltage DC microgrids, IEEE Transactions on Power Delivery, vol. 24, no. 3, pp. 1045–1053, 2009.

11. D. M. Yehia and D. E. A. Mansour, Modeling and Analysis of Superconducting Fault Current Limiter for System Integration of Battery Banks, IEEE Transactions on Applied Superconductivity, vol. 28, no. 4, Jun. 2018.

12. E. W. Nahas, D. E. A. Mansour, H. A. Abd El-Ghany, and M. M. Eissa, Accurate Fault Analysis and Proposed Protection Scheme for Battery Energy Storage System Integrated with DC Microgrids, 2018 20th International Middle East Power Systems Conference, MEPCON 2018 - Proceedings, pp. 911–917, Feb. 2019.

13. M. Farhadi and O. A. Mohammed, A New Protection Scheme for Multi-Bus DC Power Systems Using an Event Classification Approach, IEEE Trans Ind Appl, vol. 52, no. 4, pp. 2834–2842, Jul. 2016.

14. Government of Western Australia, Department of Mines, Industry Regulation and Safety, Battery Energy Storage Systems A guide for electrical contractors, 2019.

15. R. Zhao, J. Liu, and J. Gu, Simulation and experimental study on lithium ion battery short circuit, Appl Energy, vol. 173, pp. 29–39, Jul. 2016.

16. 1375-1998 - IEEE Guide for the Protection of Stationary Battery Systems | IEEE Standard | IEEE Xplore. 1998.

Dr Rodney H.G. Tan

He currently is an Associate Professor at UCSI University, Malaysia. He received his PhD degree in Electrical Engineering from UNITEN in 2013. His research interests include Renewable Energy, Energy Storage System and Power Electronic Converter.

Authors Introduction

Mr. Kong De Kang

He received his Bachelor's degree from the Department of Electrical & Electronics Engineering, UCSI University in 2023. He is currently working as Electrical Engineer in one of Mechanical & Electrical Consulting Firm in Kuala Lumpur, Malaysia.

Dr Farah Adilah Jamaludin

She is an Assistant Professor of Faculty of Engineering Technology & Built Environment, UCSI University. She graduated from the Department of Electrical & Electronics Engineering, Universiti Putra Malaysia, in 2008. She received her PhD degree in Electrical Power Engineering from Universiti Putra Malaysia in 2020. Her research interest is Power System and Energy Management.

An Intelligent Cargo/Warehouse Management System

Zhongheng Sun[1], Yue Zhou[1], Xun Sun[1], Wenzhuo Fan[1], Wenxuan Zhou[1]

[1]Maynooth International Engineering College, Fuzhou University, Fujian, China
E-mail: ZHONGHENG.SUN.2022@MUMAIL.IE, YUE.ZHOU.2022@MUMAIL.IE, XUN.SUN.2022@MUMAIL.IE,
WENZHUO.FAN.2022@MUMAIL.IE, WENXUAN.ZHOU.2022@MUMAIL.IE

Abstract

This paper presents a low-cost and user-friendly warehouse management system developed using Arduino and ESP8266 hardware. The system accurately monitors temperature, humidity, and harmful gas concentration, and acts as an intrusion detector in a warehouse environment. It also includes functions for managing and detecting goods using RFID chips and for automatic delivery using unmanned intelligent vehicles. Data is uploaded to a cloud-based relational database in real-time. The system provides a cost-effective and efficient solution for warehouse *management.*

Keywords: Arduino, ESP8266, Warehouse management system, RFID chips, Unmanned intelligent vehicles

1. Introduction

In today's fast-paced business environment, efficient warehouse management plays a crucial role in maintaining a competitive edge [1], [2]. The ability to monitor and control warehouse conditions accurately has a significant impact on the quality of stored goods and overall operational efficiency [3]. To address these needs, the development team has developed a highly extensible, comprehensive, user-friendly, quick-to-deploy, and cost-effective warehouse management system [4].

The system offers real-time monitoring of temperature, humidity, harmful gas concentration, and intrusion detection, ensuring optimal storage conditions and enhanced safety measures [5]. The collected data is seamlessly uploaded to the cloud, enabling easy access to the warehouse conditions from anywhere.

Besides monitoring capabilities, the system also uses ESP8266 and RFID chips to enable efficient goods management and detection. By scanning the RFID chips [6] attached to goods boxes, the system can obtain detailed information about the goods which includes quantity and type. This information is then uploaded to a cloud-based relational database [7], providing real-time visibility of inventory levels and facilitating streamlined inventory management.

Moreover, an experimental proportionally scaled small car model is constructed. The small car can automatically perform point-to-point transportation of goods. Through the integration of routing algorithms, the system optimizes the routing of small cars, ensuring efficient and timely transportation of goods from one location to another without the need for human intervention. This automation reduces manual labor and enhances operational efficiency.

The warehouse management system offers a cost-effective and efficient solution that aligns with the demands of modern business operations. By leveraging cloud computing, RFID, and unmanned intelligent vehicles, a comprehensive platform that enhances warehouse productivity, accuracy, and adaptability is provided.

Figure 1 shows the overall structure of the system. The monitoring system and management system return the monitored data to the cloud platform, which then transfers the data to the control section for cargo management.
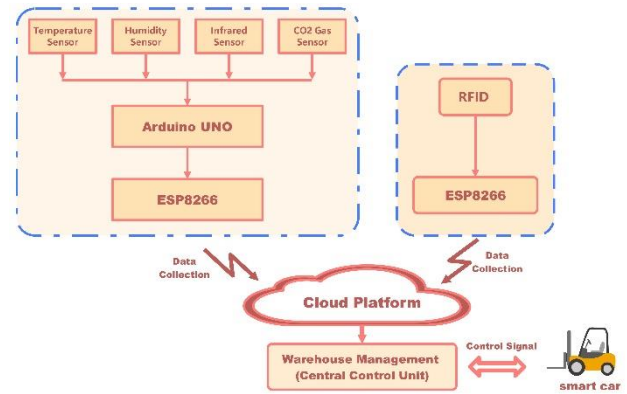
Figure 1 Overall system architecture diagram

2. Technical Background

The IoT-based warehouse monitoring and management system, utilizing Arduino and ESP8266 hardware, offers a comprehensive solution for real-time monitoring of crucial warehouse parameters. By continuously tracking temperature, humidity, harmful gas concentration, and intrusion detection, the system ensures optimal storage conditions and enhanced safety measures.

To enable seamless data integration and management, the system leverages the Alibaba Cloud IoT platform. This cloud-based infrastructure allows for the efficient uploading and storage of data, providing a centralized hub for monitoring and analysis. The platform also facilitates easy integration with third-party applications, thanks to Alibaba Cloud's well-documented APIs, enabling rapid expansion and deployment of new features.

In addition to real-time monitoring, the system incorporates a cloud inventory management system based on RFID recognition. By utilizing RFID technology, the system enables efficient and accurate goods tracking and management. Each item is equipped with an RFID tag, which allows for quick and reliable identification, inventory updates, and location tracking. This streamlines the

Zhongheng Sun, Yue Zhou, Xun Sun, Wenzhuo Fan, Wenxuan Zhou

inventory management process, reduces manual labor, and minimizes errors.

To facilitate the movement of goods within the warehouse, the system included smart delivery cars. In our experimental system, a proportional scaled small car model is implemented. This automatic car operates based on predefined algorithms and navigates through the warehouse to transport goods to their designated locations. The cars are also able to operate manually using control from Android devices. By automating this process, the system minimizes human intervention, reduces handling costs, and improves overall operational efficiency.

The integration of IoT in warehouse management systems has enabled users to manage their data in real time and proceed with accurate inventory tracking, demand forecasting, and other crucial operations.

To develop the software for the system, Arduino IDE is utilized and Arduino UNO and ESP8266 development chip are used and programmed using C language and Arduino language, creating a robust and reliable software solution that connects seamlessly with the Alibaba Cloud IoT platform using the WebSocket protocol. This protocol ensures low latency, long-lasting connections, and power-efficient environment monitoring.

3. Literature Review

Zhang, B. et al. developed an intelligent warehouse management system utilizing the L298N chip, RC522 for near-field communication, and W5500 module for network communication [8]. The MSP430F5529 microcontroller acts as the central control unit, overseeing tasks such as motor control, data reading, and network communication. The system facilitates smart cargo management, including card activation, encrypted data transmission, and remote monitoring. This innovation streamlines warehouse operations, saving resources, optimizing space, and ensuring reliable security. Particularly beneficial for small to medium enterprises and courier companies, it enhances efficiency and reduces costs.

Christos Spandonidis et al. developed a smart container monitoring system based on the Internet of Things [9]. Key challenges such as weight reduction, fire smoke detection, hardware safety, and logistics maintenance are addressed. By integrating low-cost, low-power sensors, the system is able to track container status and detect critical events such as fire/smoke, impacts, and accidental misuse. Provides better and safer control over cargo loading/unloading operations and flight processes. The experimental results show that the system can make a breakthrough in the current technical level of container technology and aircraft cargo operation.

Xu Zhijie et al. took account of clear and real-time video streams simultaneously [10]. They put forward hybrid multichannel video stream transmission strategies for a remote surveillance system, in which high-quality panoramic video and real-time interactive video are combined together to satisfy surveillance system requirements of high data stream quality and interactive operation.

4. Monitoring System Design

For the monitoring system, Arduino Uno and ESP8266 in conjunction are chosen with the Arduino IDE to implement its functionality. The Arduino Uno is a microcontroller board based on the ATmega328P, while the ESP8266 is a low-cost Wi-Fi microchip with a full TCP/IP stack and microcontroller capability. These devices were chosen for their ease of use, low cost, and versatility compared to other options. Figure 2 and Figure 3 show the monitoring system.
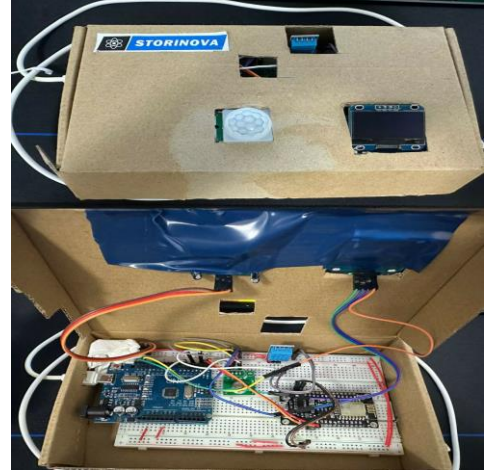
Figure 2 Picture of the monitoring system

In terms of specific functionality, a DHT11 temperature and humidity sensor, an infrared human body sensor, and a harmful gas concentration sensor is connected to Arduino Uno. The DHT11 sensor measures temperature and humidity using a capacitive humidity sensor and a thermistor. The harmful gas sensor detects the presence of harmful gases such as carbon monoxide and methane. Both sensors are connected using serial ports by the software serial port function of Arduino. The infrared sensor detects the presence of humans by measuring changes in infrared radiation. It uses an analog input to transmit data to the Arduino.
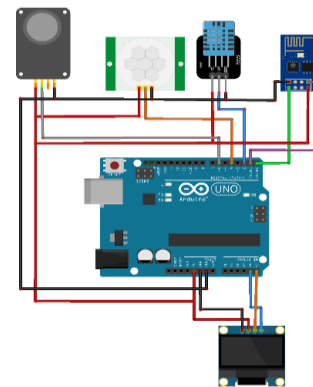
Figure 3 Hardware implementation map of the monitoring system

After reading data from the sensors, the Arduino UNO board hardware serial port transmits the data in real-time to the ESP8266 for processing and transmission. The ESP8266 is responsible for processing the data and transmitting it to a cloud server for storage and analysis.

For the logic design of the ESP8266, MQTT and WebSocket protocols are used for transmitting data. MQTT is a lightweight messaging protocol for small sensors and mobile devices, while WebSocket is a protocol for full-duplex communication over a single TCP connection. ESP8266 chip will receive data from the Arduino serial port in bit storage format and perform simple frame header and tail check and bit processing. Then the ESP8266's 2.4G Wi-Fi module is used for establishing a connection to a provided wireless network and to establish a connection with the Alibaba Cloud server using the MQTT protocol over WebSocket. The data is uploaded and displayed in real-time on a front-end page. Figure 4 shows the front-end web page and Figure 5 shows the Hardware implementation diagram of management module.

Figure 4 Picture of the front-end web page

Figure 5 Hardware implementation diagram of management module

5. Management System Design

RFID technology is applied to scan and track goods within the warehouse. The scanned information is then uploaded to a cloud platform, where it is organized into visualized tables, enabling real-time management. Additionally, a feature on the platform that allows us to send instructions to automated cars is implemented by the team, enabling them to transport goods efficiently. Figure 6 shows the management module and small car are shown below.

Figure 6 Photos of the RFID system and car

The implementation of RFID technology in the warehouse management system has significant positive impacts on management efficiency. By scanning RFID tags attached to each item, users are able to accurately track the quantity, name and other relevant information of the cargo. This data can then be transmitted to the cloud-based back-end server, where it is processed and transformed into user-friendly visualizations. This allows warehouse managers and personnel to have a clear overview of inventory levels, stock movement, and overall warehouse performance at any given time.

Furthermore, users can send instructions to the cars, directing them to specific locations within the warehouse to pick up or deliver goods. The system optimizes the routes of the car and ensures that goods of the same category are stored together, maximizing efficiency and minimizing the time required for manual handling.

With the integration of real-time data analytics and automated car operations, the smart warehouse solution provides a comprehensive and efficient approach to inventory management. The combination of RFID technology, cloud-based computing, and automation streamlines processes, reduces errors and enhances overall productivity. As a result, warehouse operations become more agile, cost-effective, and adaptable to changing business demands.

In summary, the warehouse management module leverages RFID scanning, cloud-based data processing, and intelligent delivery car systems to enable real-time inventory tracking, visualization, and streamlined operations. This intelligent solution empowers businesses to achieve higher levels of efficiency, accuracy, and scalability in their warehousing processes.

6. Effectiveness Validation

To verify the effectiveness of the system, the team conducted an on-site verification. The average number of backlogged items in the warehouse for incoming/outgoing shipments is recorded over two weeks, as well as the number of instances where the warehouse's humidity and temperature exceeded the standards (the temperature and humidity test data is collected every hour in a day). The system was not installed in the first week and then installed in the second week. Data changes in these two weeks were monitored and analyzed, and the results are shown in Table 1. The graph of average data for using the warehouse management system as shown in Figure 7.

Table 1 The effect of using our warehouse management system

Zhongheng Sun, Yue Zhou, Xun Sun, Wenzhuo Fan, Wenxuan Zhou

Items Time	Number of backlogged items		Temperature and Humidity Exceedance Incidents Times	
Time	Before	After	Before	After
Day1	10	5	7	5
Day2	8	3	7	4
Day3	9	4	8	4
Day4	6	2	11	1
Day5	11	4	9	1
Day6	8	1	9	3
Day7	8	1	7	2
Total	60	20	58	20
Average	8.57	2.86	8.29	2.86
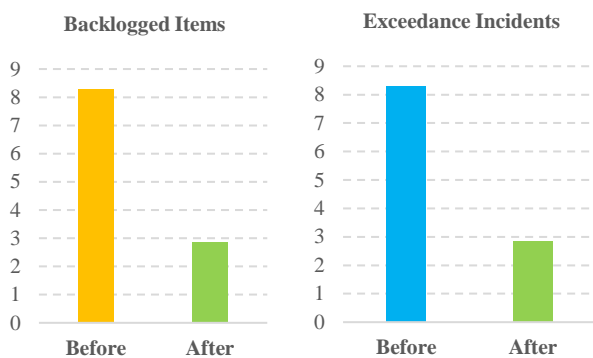
Figure 7 Graph of average data for using the warehouse management system

7. Backend Cloud Service

The backend cloud service offers a comprehensive set of features designed to streamline and optimize users' operations. With a wide range of functionalities, the service serves as an effective tool to enhance the efficiency and effectiveness of management and monitoring processes.

The main function of the cloud service is storing and displaying data. It provides a secure and reliable storage infrastructure that ensures all the data are safely preserved. The intuitive user interface allows for easy access to stored data, enabling quick retrieval and analysis. By centralizing the data of cargo to the cloud service, users can eliminate the hassle of managing multiple data sources and improve data consistency.

Moreover, the cloud service is designed to handle connections between different terminals and data sources, which facilitates seamless integration with various data stream outputs, enabling real-time data synchronization and updating. This functionality ensures that users have instant access to real-time information, enabling timely decision-making and accurate analysis.

One of the most prominent features of the cloud service is its algorithmic path calculation capability. This feature enables users to exercise semi-automatic control over robots or unmanned cars within their operations. The service can calculate optimized paths for the robots, guiding them to transport goods to designated locations with minimal human intervention. By automating this process, handling costs can be significantly reduced, errors minimized, and overall operational efficiency improved.

In addition, the cloud service acts as a transition layer, allowing for seamless integration with third-party secondary data analysis and prediction services. By leveraging these external resources, users can gain deeper insights into their data, identify patterns, and make more informed decisions. This integration enhances the analytical capabilities of the system, enabling users to extract valuable insights and drive continuous improvement in their operations.

8. Conclusion

The study successfully developed a cost-effective, user-friendly and expandable warehouse management system using Arduino and ESP8266 hardware, which has unique advances compared with similar systems [11]. It accurately monitors temperature, humidity, gas concentration, and intrusions. The system utilizes RFID chips for goods management and detection and unmanned intelligent vehicles for automated delivery. Real-time data is uploaded to a cloud-based database, providing an efficient and cost-effective solution for warehouse management [12], [13], and also being effective for monitoring and making decisions for controlling the carbon footprint of the venue under monitoring.

Overall, this system enhances operational efficiency, safety, and reliability. Further research can expand its functionality and applicability in warehouse management.

Acknowledgement

References

[1] S. L. Ullo and G. R. Sinha, "Advances in Smart Environment Monitoring Systems Using IoT and Sensors," Sensors (Basel), vol. 20, no. 11, p. 3113, May 2020. doi: 10.3390/s20113113.

[2] A. Čolaković, S. Čaušević, A. Kosovac and E. Muharemović, "A Review of Enabling Technologies and Solutions for IoT Based Smart Warehouse Monitoring System," in I. Karabegović (Ed.), New Technologies, Development and Application III. NT 2020. Lecture Notes in Networks and Systems, vol 128. Cham: Springer, 2020. doi: 10.1007/978-3-030-46817-0_73.

[3] Afzal, M. "IoT Based Real Time Warehouse Monitoring using Sparkfun ESP8266 Thing Dev and Cayenne MyDevices," Conference Paper, April 2018.

[4] D. Parida, A. Behera, J. K. Naik, S. Pattanaik and R. S. Nanda, "Real-time Environment Monitoring System using ESP8266 and Thing Speak on Internet of Things Platform," 2019 International Conference on Intelligent Computing and Control Systems (ICCS), Madurai, India, 2019, pp. 225-229, doi: 10.1109/ICCS45141.2019.9065451.

[5] Zafar, S., Miraj, G., Baloch, R., Murtaza, D., & Arshad, K. (2018). An IoT Based Real-Time Environmental Monitoring System Using Arduino and Cloud Service. Engineering, Technology & Applied Science Research, 8(4), 3238-3242.

[6] M. Kamrul Hasan et al., "IoT-Based Warehouse Management System," ResearchGate, Oct. 2022.

[7] M. K. Hasan, M. Junjie, A. K. M. Ahasan Habib, A. Al Mamun, T. M. Ghazal and R. A. Saeed, "IoT-Based Warehouse Management System," 2022 International Conference on Cyber Resilience (ICCR), Dubai, United Arab Emirates, 2022, pp. 1-6, doi: 10.1109/ICCR56254.2022.9995768.

[8] Xu Zhijie, Wang Geng. A Hybrid Real-time Video Streaming Remote Surveillance System Based on Mobile Robot[C]//International Conference on Computer Networks and Communication Technology (CNCT 2016). Advances in Computer Science Research (ACRS). Paris: Atlantis Press, 2017, 54:458-466.

[9] Spandonidis, C., Sedikos, E., Giannopoulos, F., Petsa, A., Theodoropoulos, P., Chatzis, K., & Galiatsatos, N. (2022). A novel intelligent iot system for improving the safety and planning of air cargo operations. Signals, 3(1), 95-112.

[10] Xu Zhijie, Wang Geng. A Hybrid Real-time Video Streaming Remote Surveillance System Based on Mobile Robot[C]//International Conference on Computer Networks and Communication Technology (CNCT 2016). Advances in Computer Science Research (ACRS). Paris: Atlantis Press, 2017, 54:458-466.

[11] J. Zhang et al., "Design and application of Internet of things-based warehouse management system," International Journal of Production Research, vol. 56, no. 15, pp. 5217-5235, Aug. 2018.

[12] Jarašūnienė, A., Čižiūnienė, K., & Čereška, A. (2023). "Research on Impact of IoT on Warehouse Management". Sensors (Basel), 23(4), 2213.

[13] W. C. Tan and M. S. Sidhu, "Review of RFID and IoT integration in supply chain management," Operations Research Perspectives, vol. 9, p. 100229, 2022, doi: 10.1016/j.orp.2022.100229.

Mr. Wenzhou Fan

He is currently pursuing Bachelor of Electronic Engineering as 3rd year student in the Department of Maynooth International Engineering College, Fuzhou University, China. His research interests are integrated circuit development.

Authors Introduction

Mr. Xun Sun

He is currently pursuing Bachelor of Electronic Engineering as 3rd year student in the Department of Maynooth International Engineering College, Fuzhou University, China. His research interests are embedded system development.

Miss. Yue Zhou

She is currently pursuing a Bachelor of Electronic Engineering as 3rd year student in the Department of Maynooth International Engineering College, Fuzhou University, China. His research interests are AI vision.

Mr. Wenxuan Zhou

He is currently pursuing a Bachelor of Electronic Engineering as 3rd year student in the Department of Maynooth International Engineering College, Fuzhou University, China. His research interests are embedded systems and hardware technology.

Mr. Zhongheng Sun

He is currently pursuing Bachelor of Electronic Engineering as 3rd year student in the Department of Maynooth International Engineering College, Fuzhou University, China. His research interests are integrated circuit development.

A Comprehensive Approach to Design and Implement an IoT-Enabled Intelligent Shopping Cart System with Obstacle-Aware Navigation and Enhanced Customer Engagement for Elevated Consumer Experiences

Yao Chen[1], Jiacheng Du[1], Bo Peng[1], Ningfei Wang[1], Zehan Huang[1], Wei Hong Lim[2], Sew Sun Tiang[2], Mastaneh Mokayef[2], Chin Hong Wong[1*]

[1]Maynooth International Engineering College, Fuzhou University, Fujian, China
[2]Faculty of Engineering, Technology and Built Environment, UCSI University, 1, Jalan Puncak Menara Gading, UCSI Heights, 56000 Cheras, Kuala Lumpur, Malaysia.
E-mail: yao.chen.2022@mumail.ie, jia.du.2022@mumail.ie, bo.peng.2022@mumail.ie, ning.wang.2022@mumail.ie, zehan.huang.2022@mumail.ie, limwh@ucsiuniversity.edu.my, tiangss@ucsiuniversity.edu.my, mastaneh@ucsiuniversity.edu.my, chinhong.wong@mu.ie

Abstract

Supermarket shopping is an experience that everyone has in life. This project aims to design an intelligent and user-friendly shopping cart and interactive web page, to elevate the overall customer shopping experience. The shopping cart designed in this project integrates multiple functionalities that allow users to access information about the items within the cart via the interactive web page and exert control over the movement of the cart. While moving, the shopping cart can autonomously detect obstacles in its path and navigate around them. Simultaneously, users can check the total cost of their selected items by reviewing the cart's contents, enabling them to decide whether the cart should continue following them. The project primarily utilizes an ultrasonic module to determine the cart's location and trail a specific customer. The shopping cart employs automatic barcode scanning to identify various products, and the WiFi module facilitates communication with the server via the MQTT protocol, enabling seamless interaction. The fruition of this project serves as a tangible representation of the Internet of Things application, demonstrating how connectivity through the internet enhances people's lives with greater convenience.

Keywords: Internet of Things, ultrasonic ranging, STM32singlechip, smart shopping cart

1. Introduction

The widespread of information, and communication technology, coupled with the advancements in Internet of Things technology, has brought our life into the stage of a smart city. The increase in electronic devices and applications has significantly improved convenience in our daily lives. With the increasing prevalence of mobile devices, the development of smart cities naturally falls toward this technological landscape. Therefore, it is necessary to develop intelligent equipment [1].

In the shopping scene, an intelligent shopping cart can greatly enhance the shopping experience. Essential and fundamental functions of an intelligent shopping cart include automatic following the user as well as automatic obstacle identification and avoidance. Another convenient feature is the automated calculation of the total price of items in the shopping basket. This enables customers to view the items and their respective prices on the web page from their mobile phones, eliminating the need for manual price calculations.

2. Literature Review

In the past, numerous researchers have conducted diverse studies on smart shopping carts. Supermarket owners benefit from these innovations in terms of time savings, reduced manpower, and space efficiency, leading to decreased investments. The smart shopping trolley can be used in all retail shopping malls, supermarkets, hypermarkets, and clothing showrooms [2].

Several innovative concepts have emerged to revolutionize shopping cart design, aiming to enhance the overall shopping experience and reduce human efforts. Sadia *et al.* [3] proposed the integration of RFID technology to bolster the durability of product identification tags. They suggest the implementation of RFID in shopping carts, creating an intelligent shopping system where all supermarket carts are equipped with RFID tags to reduce wait times during the checkout process and improve the customer purchasing experience.

Gunawan et al. [4] installed an IOIO microcontroller and an Android smartphone together with sensors and a controller in a shopping cart. The basket is configured in the shape of a two-wheeled mobile robot.

Yao Chen, Jiacheng Du, Bo Peng, Ningfei Wang, Zehan Huang, Wei Hong Lim, Sew Sun Tiang, Mastaneh Mokayef, Chin Hong Wong

Tai *et al.* [5] also utilized the ultrasonic sensor because of the advantages include a wide detection area, reduced sensitivity to light, the ability to detect glass and shiny surfaces, compact size, lightweight construction, minimal memory usage, cost-effectiveness compared to Laser Range Finder (LRF) or cameras, and lower power.

These investigations into shopping cart functionality and application technology have significantly expanded the research landscape of intelligent shopping carts. As a result, tasks have become more focused and precise, yielding results that align smart shopping carts more closely with people's evolving needs.

3. Methodology

Figure 1Figure 1 shows a flow chart for the smart shopping cart system from the user side whereas Figure 2 illustrates a flow chart for the operation of the shopping cart.
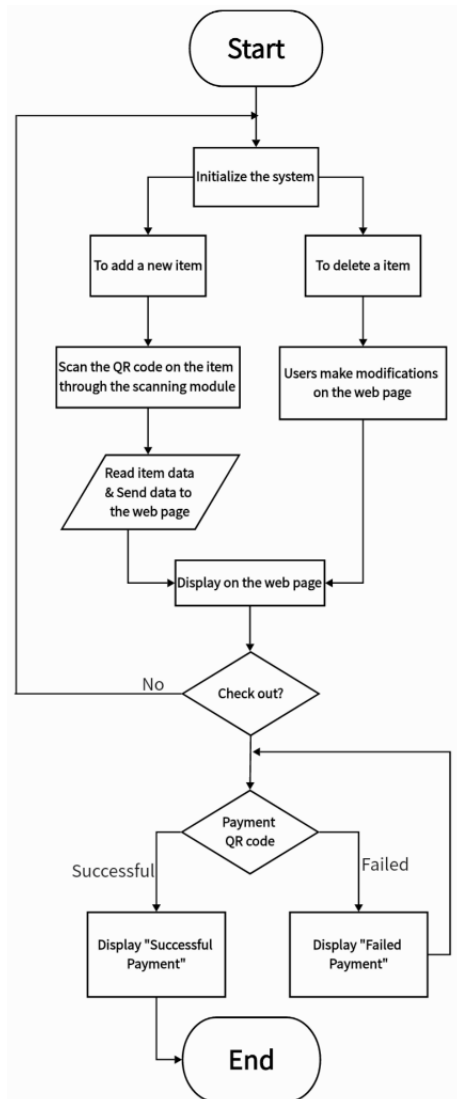
Figure 1 Flow chart for the smart shopping cart system from the user side
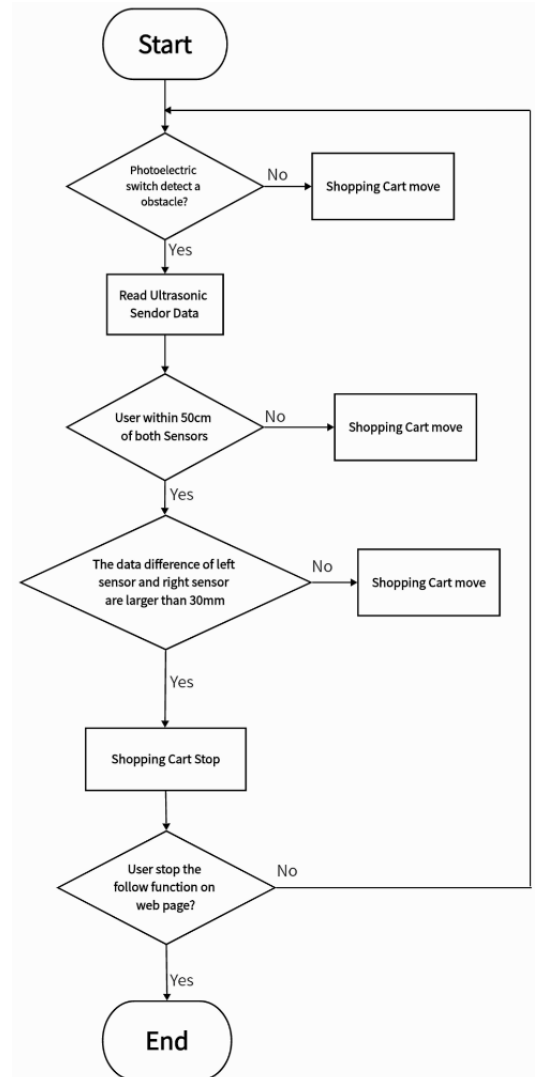
Figure 2 Flow chart for the operation of the shopping cart

When the shopping cart is started, it will detect obstacles by photoelectric switches and avoid them. The ultrasonic sensor will provide real-time feedback on the distance between the user and the shopping cart. If the distance between the shopping cart and the user, as well as the orientation of the shopping cart, exceeds the set range, the shopping cart will move and adjust the distance and direction. The mobile function of the shopping cart will continue until the user stops the function on the webpage. In addition, users can use the QR code scanning module on the shopping cart to scan the QR code of the product and add it to the shopping list on the webpage, to manage the list and settle the payment on the webpage.

3.1. Microcontroller

STM32 F407 ZET6 is used as the main controller of the shopping cart. By handling the information obtained from sensors and servers, the controller is able to control the movement of the shopping cart and communicate with the user.

3.2. PCB Module

The PCB circuit board consists of STM32 F407ZET6 main control board, DC-DC buck module, A4950 dual motor drive module, ultrasonic sensor, photoelectric switch sensor, two-dimensional code module and motor encoder as shown in Figure 3 and Figure 4. The circuit board is highly integrated and the modular design makes it easy to update and repair.
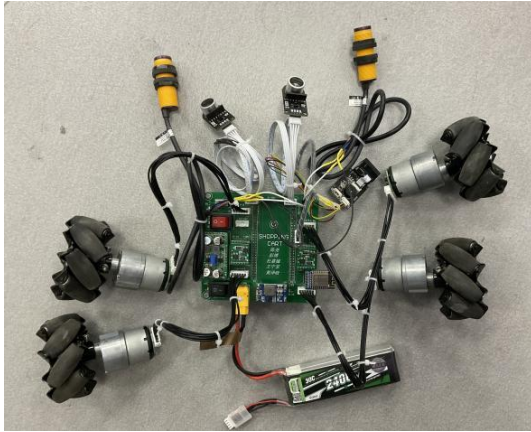
Figure 3 Connection between the main board, sensors, and wheels
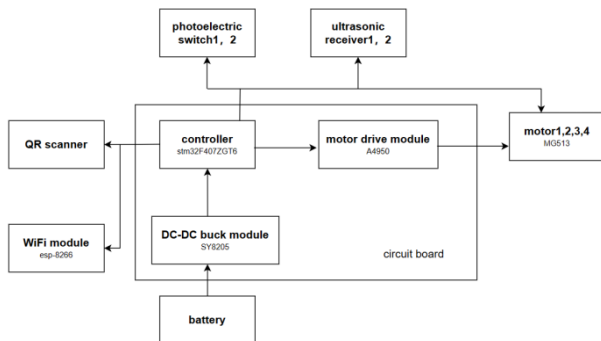
Figure 4 Block diagram for the circuit connection

3.3. DC-DC Antihypertensive Module

The DC-DC antihypertensive module employs the SY8205 chip, a high-efficiency synchronous stepdown DC-DC converter capable of delivering a 5 A output current. The SY8205 operates over a wide input voltage range from 4.5 V to 30V and integrates the main switch and synchronous switch with very low RDS(ON) to minimize the conduction loss. The module consists of two voltage outputs, one 5V voltage and another output to meet the diverse voltage requirements of various modules.

The chassis for the shopping cart is crafted from a 6 mm thick POM plate through laser cutting as shown in Figure 5(a). The chassis is used to hold the shopping cart bracket, circuit board and various sensors. Furthermore, 3D

printing technology is used to print the clamp to hold the ultrasonic sensors as shown in Figure 5(b).
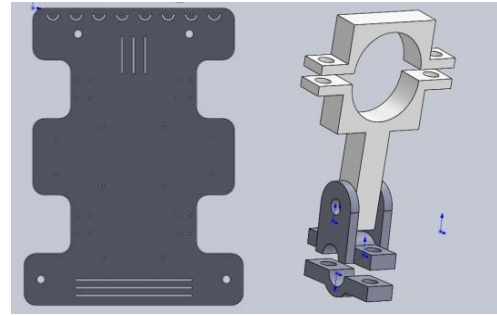
Figure 5 (a) The chassis and (b) the ultrasonic clamp

3.4. Mecanum Wheel

The outer ring of the Mecanum wheel as shown in Figure 6 is equipped with rollers arranged a 45° to the axle and contact with the ground. During rotation, the friction generates a 45° reverse thrust force along the axle. This oblique thrust force can be divided into two vectors which are longitudinal and transverse. The entire vehicle is propelled by two pairs of Mecanum wheels, each featuring a mirrored arrangement of rollers. Each wheel generates its own vector, and the combined force of these vectors determines the final motion of the car [6].
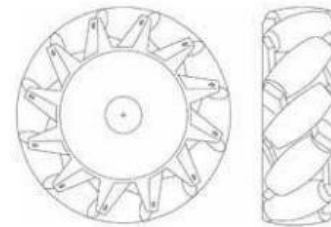
Figure 6 Mecanum wheel

Omnidirectional motion is achieved by installing four Mecanum wheels at the corners of the chassis. The overall velocity, V, can be obtained through Eq (1), where V_x, and V_y indicate velocity in the x-direction and y-direction respectively and ω is the angular velocity, as shown in Figure 7.
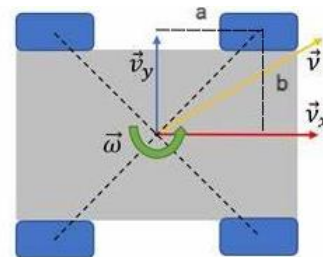
Figure 7: The direction of velocity.

$$\rightarrow \quad \rightarrow \quad \rightarrow \quad \rightarrow \tag{1}$$

Yao Chen, Jiacheng Du, Bo Peng, Ningfei Wang, Zehan Huang, Wei Hong Lim, Sew Sun Tiang, Mastaneh Mokayef, Chin Hong Wong

$$\vec{V_1} = \left|\vec{V_X}\right| - \left|\vec{V_Y}\right| - \left|\vec{\omega} \times (a+b)\right|$$
$$\vec{V_2} = \left|\vec{V_X}\right| + \left|\vec{V_Y}\right| - \left|\vec{\omega} \times (a+b)\right|$$
$$\vec{V_3} = \left|\vec{V_X}\right| + \left|\vec{V_Y}\right| + \left|\vec{\omega} \times (a+b)\right|$$
$$\vec{V_4} = \left|\vec{V_X}\right| - \left|\vec{V_Y}\right| + \left|\vec{\omega} \times (a+b)\right|$$

3.5. Ultrasonic Ranging

To measure the distance between the user and the shopping cart, two ultrasonic sensors were positioned at the front end of the shopping cart as illustrated in Figure 8. These sensors were positioned roughly parallel to the user's waist.
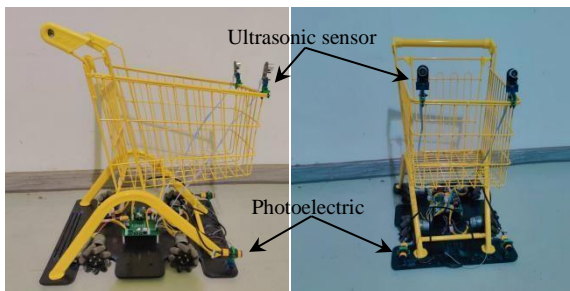
Figure 8: Overall view of the shopping cart.

By taking an ultrasonic transmitter in users' hands, the receivers can receive the ultrasonic waves and calculate the distance between the transmitters and receivers. These distances are denoted as X_1 and X_2 respectively [7]. The concept of the transmitters and the receivers is illustrated in Figure 9.
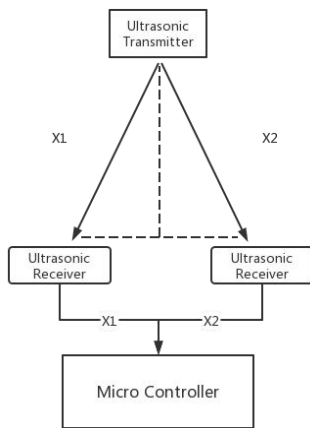
Figure 9: The concept of the transmitters and the receivers

3.6. Cascade PID

Cascade control is one of the most successful methods for enhancing single-loop control performance, particularly when the disturbances are associated with the manipulated variable and the final control element exhibits nonlinear behavior [5].

In the controlling of the Mecanum wheelbase, two desired distances between the ultrasonic transmitter and receivers are denoted as X_3 and X_4 serve as the input of the cascade PID. To control the distance between the user and the shopping cart, the distance $(X_1 + X_2)/2$ is considered as the feedback and 0.5 m is set as the desired input for the Position PD control. To ensure the front of the cart always faces the user, $(X_1 - X_2)$ is considered as the feedback and zero difference as the desired input of the Angle Loop PD control. With position control and angle control, V_x, V_y and ω can be calculated. By utilizing the calculated V_x, V_y and ω as the input for the Velocity Loop PID, the final PWM signal is conveyed to the motor as shown in Figure 10.
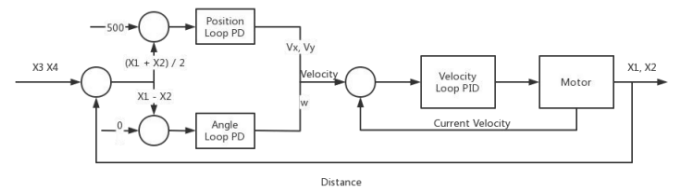
Figure 10: Block diagram of PID control.

3.7. Obstacle Avoidance

Two E18-D80NK photoelectric switches are placed on the front of the cart as illustrated in Figure 8. Upon any obstacle is detected, the cart will stop following the user and overtake the obstacle. As soon as the cart overtakes the obstacle and no other obstacles exist, the cart will continue the job.

3.8. Server Communication

Advancements in the Internet of Things (IoT) have enabled innovations in smart home and industrial automation, offering possibilities for remote monitoring and controlling devices. These solutions contribute to energy efficiency and cost saving, as appliances are monitored and controlled by small, resource-constrained embedded devices [8].

To establish communication with the server, an esp-8266 WiFi module is connected to the microcontroller board. By using the Message Queuing Telemetry Transport (MQTT) protocol, the controller is able to publish topics containing commodity information to the server and subscribe to topics conveying user input from the server.

In this system, the server initiates requests to the front-end controller through HTTP and utilizes the MQTT protocol for communication with the embedded device. The back-end is developed in Java, leveraging the SpringBoot framework to simplify the complexity of the

development, and incorporates MySQL for database connectivity and data persistence. In practical scenarios, the server needs to establish a long connection with a large number of embedded devices. The traditional blocking I/O involves threading, and the frequent switching between threads can lead to significant overhead. Therefore, domestic smart sockets are employed as the underlying I/O container to address these challenges.

A prototype incorporating these features has been successfully implemented and tested within a simple home automation network to validate its functionalities. The results demonstrate that the system responds rapidly, preventing overconsumption and electrocution hazards. This capability positions the technology to contribute to creating safer and smarter homes in the next generation [9].

The multiplexing of threads is achieved using thread pool and asynchronous callback technologies. This approach effectively reduces the overhead associated with thread creation and destruction, resulting in a substantial improvement in the system's load capacity.

3.9. QR Code Scanner

QR codes are a cost-effective technique because they are easily created and printed on a surface for distribution, often being incorporated into existing print materials. This is why the QR code was selected as a fundamental component in the design of this shopping cart [10]. The packaging of the goods is equipped with a corresponding QR code. To identify commodity information, a QR code scanner is employed, and the gathered data is transmitted to the microcontroller. The QR scanner is attached to the back of the shopping basket as shown in Figure 11.
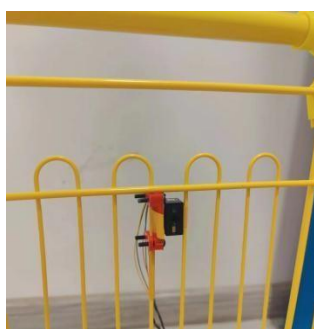
Figure 11: QR code scanner

3.10. Interactive Webpage

Users can communicate with the shopping cart and decide whether to enable the shopping cart to follow. Alternatively, even when the shopping cart is not set to follow during the shopping process, it can still be utilized for remote communication, allowing the shopping cart to

resume following the user as needed, achieving the effect of remote communication.

There is also a shopping cart commodity settlement interface where users can view the product information in the shopping cart through the web page. Additionally, each time a product is added, users can monitor the total cost and the quantity of items in the shopping cart through the settlement function. This process allows users to easily track if they exceed their estimated budget, minimizing the need for users to independently calculate commodity prices. Users can also modify the quantity of goods on the webpage or remove items from the shopping cart, contributing to a smoother shopping experience as illustrated in Figure 12. In short, this system enhances the user's shopping experience by streamlining the checkout process and providing greater control over the shopping cart contents.

Figure 12: Example of web interface.

4. Results and Discussions

To assess the stability and accuracy of the shopping cart in actual operation, conducting relevant tests becomes imperative.

4.1. Maximum Start Distance

To evaluate the maximum distance capability of the shopping cart's following function, a test involves initiating the following function in front of the shopping cart while it is in a static state. This method allows the observation of the shopping cart's following stability under varying distances between individuals and the cart.

Under unchanged conditions, the start distance between the user and the shopping cart was set to 10 cm, 50 cm, 100 cm, and 200 cm with the desired follow distance maintained at 40 cm. Figure 13, Figure 14, Figure 15 and Figure 16 show the results of the distance between the cart and the user against time. In the legend, 'Left'

Yao Chen, Jiacheng Du, Bo Peng, Ningfei Wang, Zehan Huang, Wei Hong Lim, Sew Sun Tiang, Mastaneh Mokayef, Chin Hong Wong

represents the data from the left ultrasonic receiver, and 'Right' indicates the data from the right ultrasonic receiver.
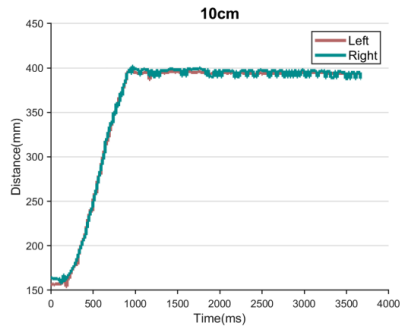
Figure 13: 10 cm starting distance between the user and the shopping cart against time.
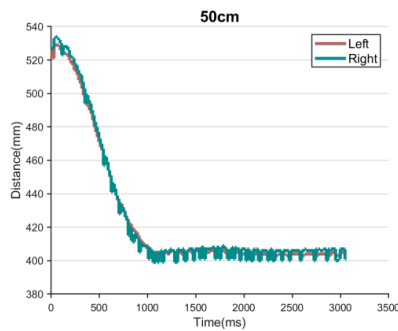
Figure 14: 50 cm starting distance between the user and the shopping cart against time.
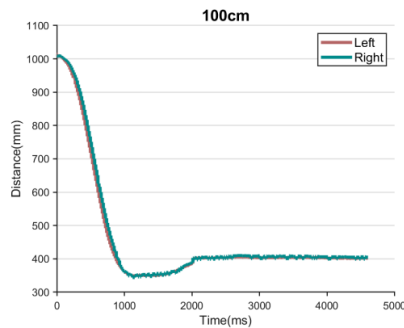
Figure 15: 100 cm starting distance between the user and the shopping cart against time.
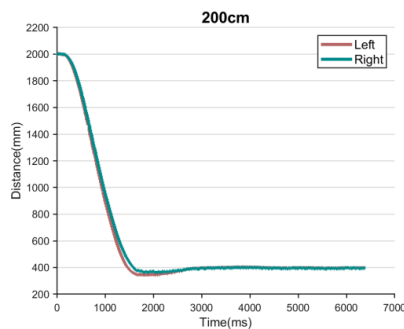
Figure 16: 200 cm starting distance between the user and the shopping cart against time.

From the test results, when the distance between the user and the shopping cart is between 10 cm to 200 cm, the shopping cart can quickly adjust the distance to the user after the start of the follow function. When the distance between the user and the shopping cart exceeds 200 cm, the ultrasonic module experiences unstable data reception. This instability prevents the shopping cart from promptly adjusting itself after activating the follow function, resulting in a failure to follow accurately. Therefore, the effective start distance for the shopping cart is set between 20 cm and 200 cm.

4.2. Turning Test

To test the follow function of the shopping cart during the user's turning process, two circular experiments were conducted. The tester walked along circles with radii of 1m and 2m at normal speeds, while the shopping cart followed. Figure 17 and Figure 18 show the turning test results.
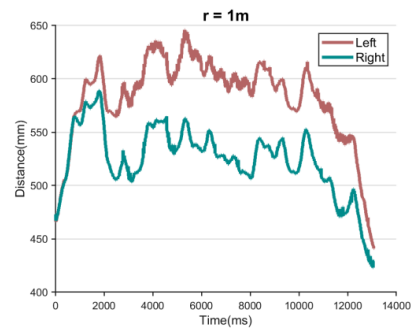
Figure 17: Distance between the user and the shopping cart against time for the tester walking along a circle with a radius of 1 m.
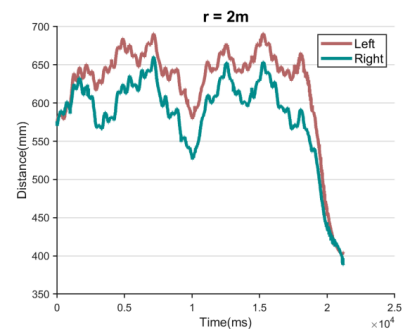
Figure 18: Distance between the user and the shopping cart against time for the tester walking along a circle with a radius of 2 m.

According to the test results, it is observed that when the turning radius is greater than 1 m, the shopping cart promptly adjusts its distance to the user.

4.3. Walking Test

To evaluate the follow function in a shopping scenario, two circular experiments were conducted. The tester

walked at speeds of 1.5 m/s and 3 m/s, representing normal walking and jogging speeds, while the shopping cart followed. The test results are illustrated in Figure 19 and Figure 20.
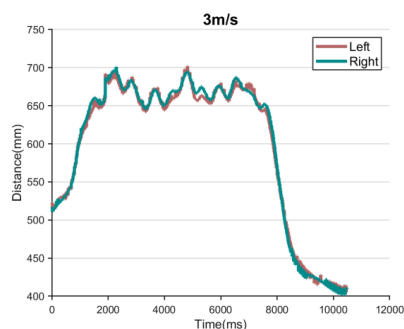
Figure 19: Distance between the user and the shopping cart against time with 3 m/s walking speed.
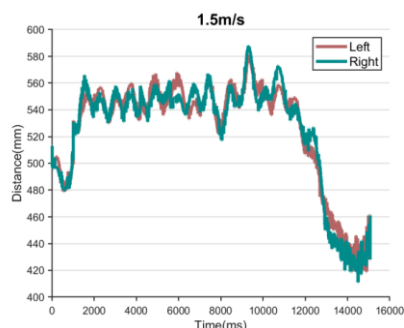
Figure 20: Distance between the user and the shopping cart against time with 1.5 m/s walking speed.

Based on the test results, when walking at a normal speed and activating the follow function, the shopping cart can successfully follow the user throughout the shopping process.

5. Conclusion

This paper highlights the accomplishments in developing a smart shopping cart system with automatic following, obstacle avoidance, and web interface-generated shopping lists. The acknowledgement of potential improvements, such as performance optimization and enhanced user interfaces, adds a realistic and forward-looking perspective. The mention of future research directions, including adaptability to different shopping scenarios, demonstrates an awareness of potential advancements.

References

1. M. Batty *et al.*, "Smart cities of the future," *European Physical Journal: Special Topics*, vol. 214, no. 1, pp. 481–518, Dec. 2012, doi: 10.1140/epjst/e2012-01703-3.
2. Viswanadha V, Pavan Kumar P, and Chiranjeevi Reddy S, "Smart Shopping Cart," 2018.
3. H. Sadia, S. Jee, K. Pal, S. Singh, and M. Marbaniang, "Iot Application Based Advanced Shopping Trolley," *Int J Eng Adv Technol*, 2019.
4. A. A. S. Gunawan *et al.*, "Development of smart trolley system based on android smartphone sensors," in *Procedia Computer Science*, Elsevier B.V., 2019, pp. 629–637. doi: 10.1016/j.procs.2019.08.225.
5. W. W. Tai, B. Ilias, S. A. Abdul Shukor, N. Abdul Rahim, and M. A. Markom, "A Study of Ultrasonic Sensor Capability in Human Following Robot System," in *IOP Conference Series: Materials Science and Engineering*, IOP Publishing Ltd, Dec. 2019. doi: 10.1088/1757-899X/705/1/012045.
6. H. Taheri, B. Qiao, and N. Ghaeminezhad, "Kinematic Model of a Four Mecanum Wheeled Mobile Robot," *Int J Comput Appl*, vol. 113, no. 3, pp. 6–9, Mar. 2015, doi: 10.5120/19804-1586.
7. A. Carullo and M. Parvis, "An Ultrasonic Sensor for Distance Measurement in Automotive Applications," 2001.
8. P. MacHeso, T. D. Manda, S. Chisale, N. Dzupire, J. Mlatho, and D. Mukanyiligira, "Design of ESP8266 Smart Home Using MQTT and Node-RED," in *Proceedings - International Conference on Artificial Intelligence and Smart Systems, ICAIS 2021*, Institute of Electrical and Electronics Engineers Inc., Mar. 2021, pp. 502–505. doi: 10.1109/ICAIS50930.2021.9396027.
9. A. Singaravelan and M. Kowsalya, "Design and Implementation of Standby Power Saving Smart Socket with Wireless Sensor Network," in *Procedia Computer Science*, Elsevier B.V., 2016, pp. 305–310. doi: 10.1016/j.procs.2016.07.360.
10. S. Statler, "Barcodes, QR Codes, NFC, and RFID," in *Beacon Technologies*, Apress, 2016, pp. 317–331. doi: 10.1007/978-1-4842-1889-1_18.

Authors Introduction

Mr. Yao Chen

He is currently pursuing Bachelor of Robotics and Intelligent Devices as 3rd year student in the Department of Maynooth International Engineering College, Fuzhou University, China. His research interests are embedded systems and machine learning.

Mr. Jiacheng Du

He is currently pursuing Bachelor of Robotics and Intelligent Devices as 3rd year student in the Department of Maynooth International Engineering College, Fuzhou University, China. His research interests are embedded systems and reinforcement learning.

Yao Chen, Jiacheng Du, Bo Peng, Ningfei Wang, Zehan Huang, Wei Hong Lim, Sew Sun Tiang, Mastaneh Mokayef, Chin Hong Wong

Mr. Bo Peng

He is currently pursuing Bachelor of Robotics and Intelligent Devices as 3rd year student in the Department of Maynooth International Engineering College, Fuzhou University, China. His research interests are data science and software development.

Ms. Ningfei Wang

She is currently pursuing Bachelor of Robotics and Intelligent Devices as 3rd year student in the Department of Maynooth International Engineering College, Fuzhou University, China. Her research interests are data science, data mining and data protection.

Mr. Zehan Huang

He is currently pursuing Bachelor of Robotics and Intelligent Devices as 3rd year student in the Department of Maynooth International Engineering College, Fuzhou University, China. His research interests are electronic information and machine learning.

Dr. Wei Hong Lim

He is an Associate Professor in Faculty of Engineering at UCSI University in Malaysia. He received his PhD in Computational Intelligence from Universiti Sains Malaysia in 2014. His research interests are optimization and artificial intelligence.

Dr. Sew Sun Tiang

She is an Assistant Professor in Faculty of Engineering at UCSI University in Malaysia. She received her PhD in Electrical and Electronic Engineering from Universiti Sains Malaysia in 2014. Her research interests are optimization and antenna design.

Dr.Mastaneh Mokayef

She is an Assistant Professor in the Faculty of Engineering at UCSI University in Malaysia. She received her PhD from the Wireless Communication Centre Faculty of Electrical Engineering at University Technology Malaysia in 2014. Her research interests include Wireless communications, spectrum sharing method, spectrum management, cellular communication systems and Antenna design.

Dr. Chin Hong Wong

He is a Lecturer at Maynooth International Engineering College at Fuzhou University, China. He received his PhD in Electrical and Electronic Engineering from Universiti Sains Malaysia in 2017. His research interests are Energy harvesting, signals and systems, and machine learning.

A Design of Four-Port Flexible UWB-MIMO Antenna for Wearable and IoT Applications

Jia Wei Tan[1], Sew Sun Tiang[1, *], Kim Soon Choong[1], Mohammad Arif Bin Ilyas[1], Mastaneh Mokayef[1], Zhi Ying Yeoh[1], Wei Kang Lai[1], Wei Hong Lim[1]

E-mail: 1001954282@ucsiuniversity.edu.my, tiangss@ucsiuniversity.edu.my, ChongKS@ucsiuniversity.edu.my, mohammadarif@ucsiuniversity.edu.my, mastaneh@ucsiuniversity.edu.my, 1001953809@ucsiuniversity.edu.my, 100953677@ucsiuniversity.edu.my, limwh@ucsiuniversity.edu.my

([1]UCSI University, Malaysia)

* Corresponding author

Abstract

This paper introduces a compact four-port MIMO antenna for ultrawideband (UWB) applications measuring 60×60 mm^2. Printed on a single layer flame resistant (FR-4) substrate (permittivity of 4.3, thickness 1.6mm), the antenna features four microstrip cells, each orthogonal to its neighbor for improved isolation. It includes a rectangular patch with staircase slits and a stepped feed line. Square stubs at the top center enhance isolation. The antenna boasts a significant return loss (-43.75dB), wide impedance bandwidth (1.967-12GHz), and an isolation below -15dB. Its envelope correction coefficient (ECC) is under 0.02 with a moderate 4.4dBi gain. Although its 63% radiation efficiency could be enhanced, the antenna's ultrawide bandwidth and compactness make it suitable for UWB wearable IOT applications.

Keywords: UWB antenna, MIMO antenna, MIMO performance.

1. Introduction

The frequency of ultra-wideband (UWB) technology was set to 3.2-10.6GHz by the Federal Communication Commission (FCC) which garnered great attention of researchers due to its promising solution to scarcity of spectrum resources. Compared to traditional narrowband technology, UWB offers more advantages and several capabilities including short range, high transmission, high bandwidth, low energy usage, low complexity, and good radiation. With this, UWB positions itself to being useful in applications such as short-range communications, sensor networks, tracking, and positioning systems [1], [2].

In a multi-path environment, UWB systems are prone to signal fading which degrades the performance of the entire system. To address this issue, Multi-Input-Multi-Output (MIMO) technology has been used to provide a multi-element patch antenna for signal transmission and reception. This implementation significantly increases spectral efficiency, reliability, and channel capacity without utilizing additional power and spectrum. The enhanced features of MIMO UWB antennas such as high data rate and exclusion from interference make it a perfect candidate for wearable and IoT applications especially in healthcare industries. Radiation exposure from the antenna is heavily monitored to be as small as possible to ensure that it does not harm human tissues since it is in close contact with human skin [2].

However, the usage of more than one radiating element, closely arranged to one another in MIMO may lead to high signal correlation and reduce the overall antenna efficiency. Therefore, appropriate decoupling needs to be done between antenna elements for better isolation in modern communication. The challenging part is to design UWB-MIMO with low isolation by keeping minimum distance between antennas. In [1], a UWB-MIMO antenna with square shaped decoupler and rectangular metallic stubs was designed to reduce mutual coupling and improve isolation. The antenna offers stable gain with peak gain at 6.8dBi, low ECC value of 0.001 and high diversity gain of 9.9. Besides, flexible substrates such as polymer, paper, and Kapton are important in the designing of wearable UWB. In [3], a wristband MIMO antenna is designed that covers frequency range of 2.75-12GHz with silicone rubber as substrate material that is used to monitor children and patients, ensuring user safety. In [4], the author proposed the use of liquid crystal polymer (LCP) that operates between 2.9-10.86GHz which results in a flexible MIMO antenna that performs well in bending and on-body conditions, suitable in wearable fields. Another result obtained from [5], uses bendable substrate FR-4 that operates up to 17GHz from 3.89GHz with stable gain.

Jia Wei Tan, Sew Sun Tiang, Kim Soon Choong, Mohammad Arif bin Ilyas, Mastaneh Mokayef, Zhi Ying Yeoh, Wei Kang Lai, Wei Hong Lim

Moreover, the design of UWB is very challenging to achieve optimized frequency and time, stable polarization, as well as low dispersion. In [6], the author proposed a methodology that introduces slotted ground structure which is a widely used technique to improve the impedance bandwidth and minimize the ground plane effects in UWB. In [7], 4 ports rectangular monopole antennas with step etching on ground plane and arrow-shaped slot on radiating patch using FR-4 substrate operates at frequency 3.1-10.6GHz, has high isolation of -17.4dB, ECC less than 0.001 which in turn decreases mutual coupling. In the medical field, wearable devices are achieving profound recognition when designed using MIMO antenna due to their quality signal and high capacity without increasing transmitted power.

This paper proposes the design of four port UWB-MIMO antenna suitable for wearable and IoT applications which is done by utilizing the CST software that involves careful selection of flexible substrate properties, structure of microstrip patch, and feed network. The design uses copper conductor for patch and FR-4 as substrate material that covers wider bandwidth and higher efficiency characteristic with low dielectric loss and high thermal endurance. The width of the microstrip and gap between the patches are adjusted to obtained impedance matching of 50 Ohm. The finalized antenna is measured at different parameters such as reflection coefficients, gain, radiation pattern, current distribution, ECC, DG, radiation and total efficiency for further analysis and improvements. Table1 shows summary of comparison between existing UWB-MIMO antennas and proposed work.

Table 1. Summary Of Comparison Between Existing UWB-MIMO Antennas and Proposed Work

References	Operating Frequency (GHz)	Peak Gain (dBi)	Dimension (mm^3)	Envelope Coefficient Correlation (ECC)	Diversity Gain (DG)	Isolation (dB)	Substrate Type
[3]	2.75-12.0	3.41	24 × 30 × 0.13	< 0.18	> 9.5	> 25	Silicon Rubber
[4]	2.90-10.86	4	65× 65 × 0.1	< 0.01	> 9.99	> 22	Liquid Crystal Polymer (LCP)
[5]	3.89-17.09	5.87	75 × 91 × 0.26	< 0.02	-	> 22	FR-4
[6]	2.9-12.4	3.09	30 ×30 × 1.6	-	-	-	FR-4
[7]	2.6-11	3.38	80 × 80 × 0.781	< 0.001	-	-17.4	FR-4
[8]	3.6-16.00	7	58× 58 × 0.8	< 0.07	-	> 18	FR4
[9]	1.5-3.6	2.4	26× 26 × 0.8	< 0.02	> 9.9	~ 25	FR4
[10]	3.3-13.6	5.7	42 × 32.5 × 1.0	< 0.02	> 9.96	> 18	Textile
[11]	3.10-12	6.2	30× 30 × 1.6	< 0.001	> 9.9	~ 17	FR-4
[12]	3.1-13.1	4	45 × 45 × 1.6	< 0.02	> 9.9985	< 17	FR-4
[13]	3.5-11	6	24 × 30 × 0.13	< 0.03	> 9.9	< 20	FR-4
Proposed work	1.926-12	4.409	60 × 30 × 1.187	<0.02	>9.93	>15	FR-4

2. Material and Methods

2.1 Single Element Antenna Design

The configuration of the proposed UWB single element microstrip antenna shown in Fig. 1. The fundamental structure of the antenna consists of three layers which are metal layer on top and bottom with substrate material, in between. The design begins with a rectangular radiating patch that uses a 50Ohm transmission line feeding technique to feed signal to the resonating patch. The ground plane size is decreased for low profile and optimization for attaining enhanced impedance bandwidth. Two symmetrical staircase-shaped slits are introduced on both sides of the radiating patch to further enhance the impedance bandwidth.

a. Front view of antenna design

b. Magnified area for staircase-like slits
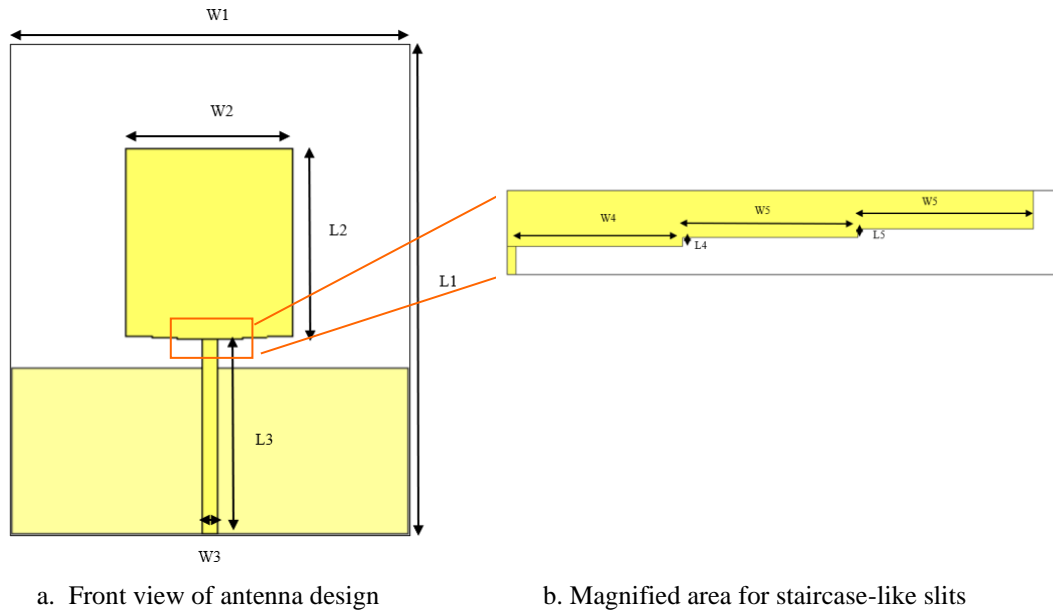
Fig. 1 Geometry of proposed antenna design.

Table 2. Optimized Parameters for the Proposed Four Port MIMO Antenna.

Parameters	W1	W2	W3	W4	W5
Values (mm)	31	13	3	1.65	2
Parameters	L1	L2	L3	L4	L5
Values (mm)	42	15.6	15.3	0.1	13

Table 2 shows the optimized parameters for the proposed four port MIMO antenna. The material chosen for the antenna patch is copper (annealed) because it is much softer and malleable which makes it easier to shape during the fabrication stage. On the other hand, the flexible dielectric substance used is flame-resistant (FR4) with dielectric constant of 4.3 and tangent loss of 0.019. Some key features of this substrate material include low cost, easily available and relative mechanical and electrical stability. It is mostly found in microstrip and stripline circuits, as well as point-to-point digital radio antennas. The thickness of the UWB microstrip antenna is 1.8mm with copper (annealed) 0.1mm and substrate as 1.6mm respectively. Overall, the FR4 substrate based UWB antenna, and its resonating patch has sizes of $31 \times 42 \times 1.7$ mm^3 and fed by a $3 \times 17 \times 0.1$ mm^3 transmission line.

The operating range of antenna generated is 2.74-10.59GHz shown in Fig. 2 which fulfills the requirement of UWB but is lesser compared to the values from the article [14] which obtain 2.9-11GHz.
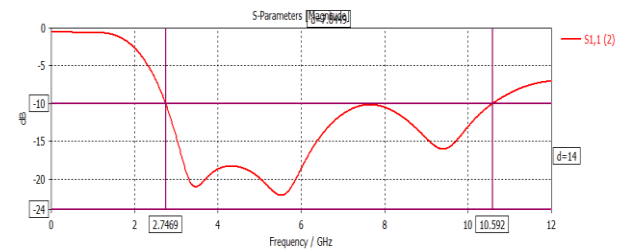
Fig. 2 Simulation results of return loss (S11) for single-slot antenna

2.2 Four-Element Flexible UWB Antenna

From the single element in Fig. 1, a four-port flexible MIMO antenna array was designed as shown in Fig. 3 (b) with a square-shaped stub inserted between orthogonal element antennas for better isolation performance.

a. Without isolator

b. With isolator branch
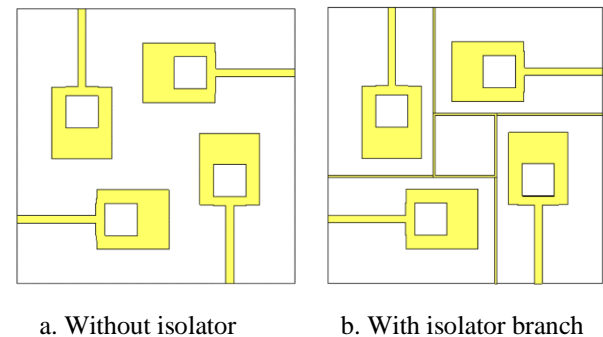
Fig 3. Four Port UWB MIMO antenna

3. Result and Discussion

The layouts were designed and simulated using CST software to determine the antenna performance in different aspects.

The S11 parameter graph is a measure of how well devices or lines are matched. A high return loss is desirable for lower insertion loss. The simulated result for return loss from the proposed UWB four-port antenna is displayed in Fig. 4 (a). The first antenna result without branch isolator (green line) shows much narrow impedance bandwidth with frequency range of 3.925-12GHz for S11 < -10dB. Besides, the return loss value is acceptable, which is -22.49dB at resonant frequency of 4.42GHz but can be improved to reduce power reflected to the system. This leads to the addition of a branch isolator (red line) to improve both return loss and bandwidth values. The S11 value almost increased twofold with value highest return loss value of -43.94dB at resonant frequency of 8.124GHz, larger bandwidth, and wider frequency range of 1.926-12GHz. The rest of the resonant frequencies for S11 below -10dB are recorded at 2.292GHz, 4.14GHz, 6.6GHz and 11.076GHz.

The remaining parameters are also compared such as S12, S13 and S14 to determine the isolation as shown in Fig 4 (b), (c), (d). The measured operating frequency covers 2.0-12GHz. The implementation of branch isolators affects S12, S13, and S14 whereby most of them are below -15dB (red line) and this shows good isolation properties. With an almost similar isolation design, the result from A. A. Ibrahim, et al [13] shows better isolation which is greater than -20dB when the isolation ports are added.
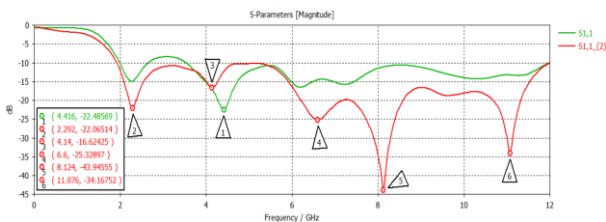
(a) S11
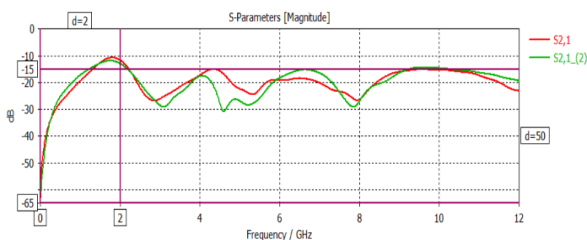
(b) S12
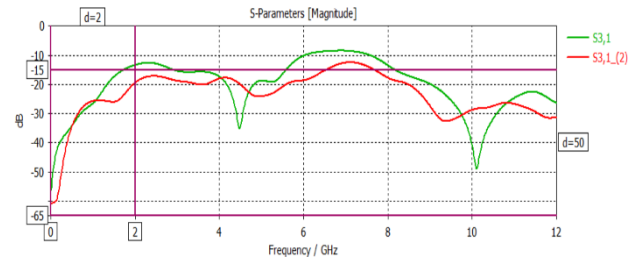
(c) S13
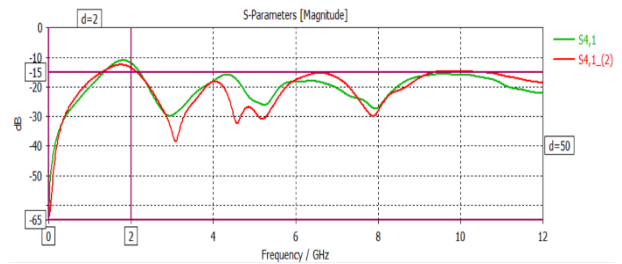
(d) S14

Fig 4. Results of return loss and isolation loss

The gain and directivity for the final proposed antenna on different resonant frequencies is tabulated in Table 3 while the far field pattern for resonant frequency, 8.124GHz is shown in Fig. 5 (a) and (b).

Table 3. Gain and Directivity Values for Different Resonant Frequencies

Resonant Frequency (GHz)	Gain (dBi)	Directivity (dBi)
2.292		
4.14		
6.6	4.407	6.36
8.124		
11.076		
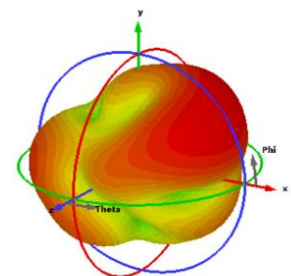
(a) Gain at resonant frequency of 8.12GHz.
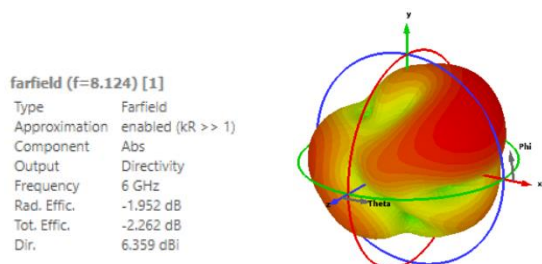
(b) Directivity at resonant frequency of 8.124GHz.

Fig 5. Gain and directivity at resonant frequency of 8.124GHz.

Fig. 6 (a), Fig. 6(b), Fig. 6(c), Fig. 6(d), and Fig. 6(e) shows the radiation pattern of the proposed four-port flexible antenna on H-plane and E-plane at different resonant frequencies. The simulation and test results were obtained when port 1 was used as excitation. The H-field and E-field exhibit a directional pattern whereby the energy is concentrated and focused on specific directions at degrees of 44° and 178° for all resonant frequencies. It allows far-distance communication due to higher gain and directivity.
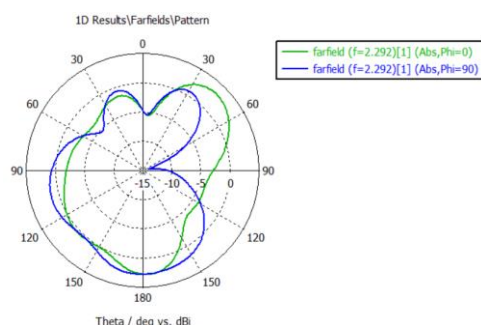
(a) Radiation pattern at 2.292GHz.
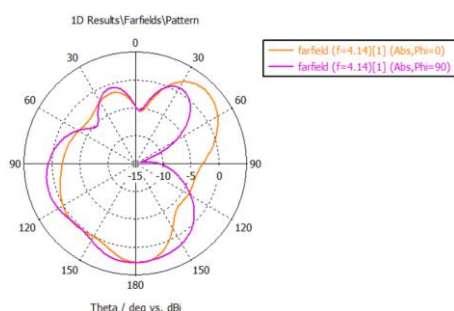
(b) Radiation pattern at 4.414GHz.
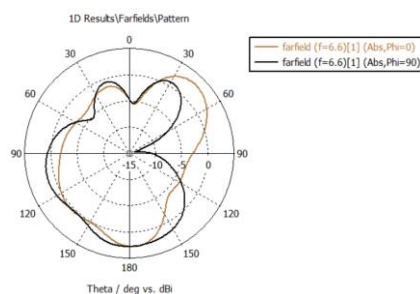
(c) Radiation pattern at 6.6GHz.
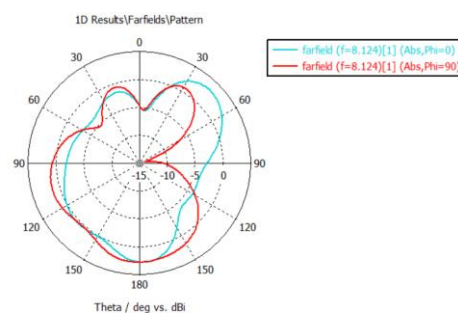
(d) Radiation pattern at 8.124GHz.
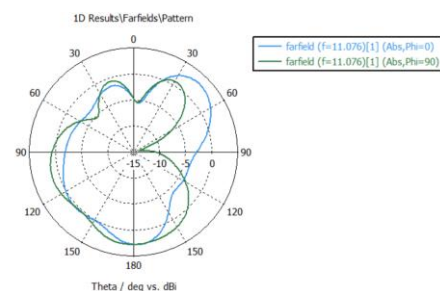
(e) Radiation pattern at 11.076GHz.

Fig 6. Radiation pattern on H-plane and E-plane of four port MIMO UWB antenna at different resonant frequencies

The surface current distribution is measured at port 1, 2, 3 and 4 with resonant frequencies at 2.292GHz, 4.14GHz, 6.6GHz, 8.124GHz and 11.076GHz and tabulated in Table 4. Fig. 7 shows the pattern of the selected resonant frequency, 8.124GHz, where the rest have similar current diagrams. The current density is most evident at the antenna feedline denoted by warm colors of red and orange color map which contributes to stronger radiation and better performance but can result in higher potential for losses and prone to heating easily due to high dissipation of electromagnetic fields and energy. The rest of the areas that are mostly displayed in blue or green colors experience low current density whereby it has

Jia Wei Tan, Sew Sun Tiang, Kim Soon Choong, Mohammad Arif bin Ilyas, Mastaneh Mokayef, Zhi Ying Yeoh, Wei Kang Lai, Wei Hong Lim

weaker radiation and signal strength but less susceptible to losses.

Table 4. Surface Current Value at Port 1 to 4 at Different Resonant Frequencies.

Resonant Frequency (GHz)	Port 1 (A/m)	Port 2 (A/m)	Port 3 (A/m)	Port 4 (A/m)
2.292				
4.14	37.0092	37.9358	38.3408	37.0869
6.6				
8.124	29.5007	33.4613	30.3317	34.7738
11.076	37.0092	37.9358	38.3408	37.0869
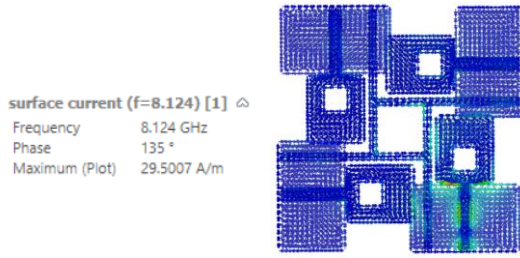
Fig. 7 Surface current distribution of four-port UWB MIMO antenna at resonant frequency, 8.124GHz.

The average radiation and total efficiency of the antenna is only recorded between 6GHz-12GHz due to mesh limitation in frequency domain. Fig. 8 shows the value in dB (magnitude) which is approximately 55% and 54% after converted to percentage form. These values are considered marginally acceptable because good antenna design can reach up to 80%. Hence, the design can still be further improved by careful calculation of geometry and dimensions, matching networks to achieve maximum performance whereby it provides better signal strength and reduces interference which are important factors for improving the reliability and operability of wearable devices in IoT applications.
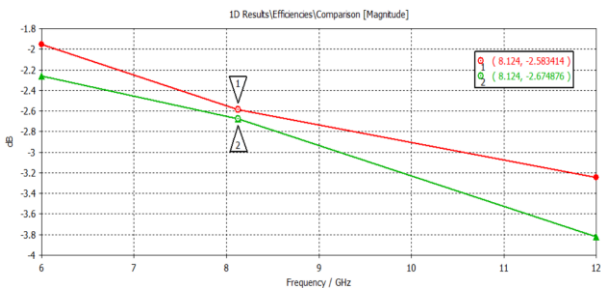
Fig. 8 Radiation efficiency and total efficiency of four-port UWB MIMO antenna.

The diversity performance of the UWB-MIMO antenna was evaluated by Envelope Coefficient Correlation (ECC) and diversity gain (DG). ECC parameters reflect the degree of correlation between adjacent elements of MIMO antennas and are expected to be less than 0.5. The ECC parameters between the antenna's ports are shown in Fig. 9 (a) which are less than 0.02 in the operating frequency band and indicate good isolation between ports.

DG evaluates the quantified improvement in signal-to-noise ratio when antenna in the MIMO system receive the RF signal. The DG parameters between different ports are given in Fig. 9 (b) It shows that satisfying DG values are achieved with values above 9.938 for the whole operating frequency band (1.926-12GHz) which gives better diversity characteristics.
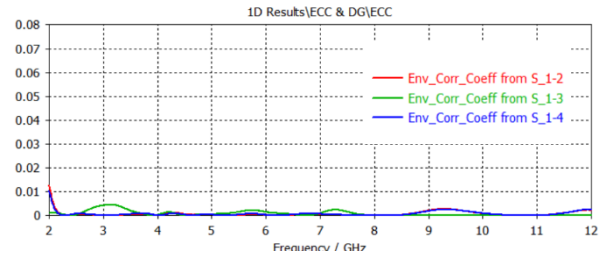
(a) ECC
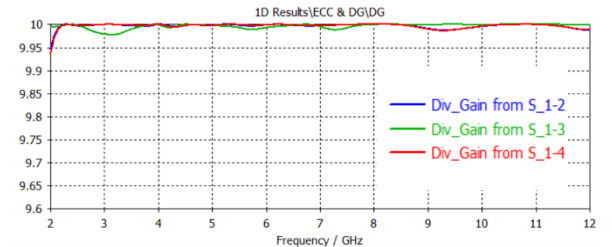
(b) DG

Fig 9. ECC and DG values for four-port UWB MIMO antenna.

Conclusion

This paper proposed a small, compact, and flexible four-port UWB MIMO antenna designed for wearable and IoT applications with overall dimensions of $60 \times 60 \times 1.8$ mm^3. The proposed antenna was arranged orthogonally to each other to form a four-port MIMO antenna structure and the final design shows that the antenna can operate in the whole UWB band covering from 1.926 to 12GHz with moderate isolation of -15dB. The isolation was achieved by adding a square-shaped stub at the top center with elongated strips extending to the edge of the substrate. It

has a high return loss value of -43.65dB, total efficiency of approximately 60%, gain of 4.4dBi, and good diversity performance, with ECC less than 0.02 and DG larger than 9.93dB. In summary, the proposed antenna can be considered as a viable candidate for UWB-MIMO wireless applications.

Acknowledgements

References

1. Abbas, N. Hussain, M. A. Sufian, J. Jung, S. M. Park, and N. Kim, "Isolation and gain improvement of a rectangular notch uwb-mimo antenna," *Sensors*, vol. 22, no. 4, Feb. 2022.
2. P. Kumar, P. Kumar, T. Ali, P. Kumar, and S. Vincent, "Ultrawideband Antennas: Growth and Evolution," *Micromachines*, vol. 13, no. 1. MDPI, Jan. 01, 2022.
3. T. Govindan *et al.*, "On the design and performance analysis of wristband MIMO/diversity antenna for smart wearable communication applications," *Sci Rep*, vol. 11, no. 1, Dec. 2021.
4. J. Zhang, C. Du, and R. Wang, "Design of a Four-Port Flexible UWB-MIMO Antenna with High Isolation for Wearable and IoT Applications," *Micromachines (Basel)*, vol. 13, no. 12, Dec. 2022.
5. Desai, J. Kulkarni, M. M. Kamruzzaman, S. Hubalovsky, H. T. Hsu, and A. A. Ibrahim, "Interconnected CPW Fed Flexible 4-Port MIMO Antenna for UWB, X, and Ku Band Applications," *IEEE Access*, vol. 10, pp. 57641–57654, 2022.
6. T. Ali and R. C. Biradar, "A Miniaturized Volkswagen Logo UWB Antenna with Slotted Ground Structure and Metamaterial for GPS, WiMAX and WLAN Applications," 2017.
7. W. Naktong and A. Ruengwaree, "Four-Port Rectangular Monopole Antenna for UWB-MIMO Applications," 2020.
8. P. Kumar, S. Urooj, and F. Alrowais, "Design and implementation of quad-port MIMO antenna with dual-band elimination characteristics for ultra-wideband applications," *Applied Sciences (Switzerland)*, vol. 10, no. 5, Mar. 2020.
9. I. Elfergani *et al.*, "Low-profile and closely spaced four-element mimo antenna for wireless body area networks," *Electronics (Switzerland)*, vol. 9, no. 2, Feb. 2020.
10. S. Kumar *et al.*, "Wideband Circularly Polarized Textile MIMO Antenna for Wearable Applications," *IEEE Access*, vol. 9, pp. 108601–108613, 2021.
11. A. H. Jabire, S. Sani, S. Saminu, M. J. Adamu, and M. I. Hussein, "A crossed-polarized four port MIMO antenna for UWB communication," *Heliyon*, vol. 9, no. 1, Jan. 2023.
12. A. Wu, M. Zhao, P. Zhang, and Z. Zhang, "A Compact Four-Port MIMO Antenna for UWB Applications," *Sensors*, vol. 22, no. 15, Aug. 2022.
13. A. A. Ibrahim, M. I. Ahmed, and M. F. Ahmed, "A systematic investigation of four ports MIMO antenna depending on flexible material for UWB networks," *Sci Rep*, vol. 12, no. 1, Dec. 2022.
14. M. Mustaqim, B. A. Khawaja, H. T. Chattha, K. Shafique, M. J. Zafar, and M. Jamil, "Ultra-wideband antenna for wearable Internet of Things devices and wireless body area network applications," *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields,* vol 32, no.6 Nov. 2019.

Authors Introduction

Ms. Jia Wei Tan

She is currently pursuing Bachelor of Engineering Communication and Electronics with Honors as final year student in Faculty of Engineering, Technology and Built Environment, UCSI University, Malaysia. Her research interests are communication and antenna design.

Dr. Sew Sun Tiang

She is an Assistant Professor in Faculty of Engineering at UCSI University in Malaysia. She received her PhD in Electrical and Electronic Engineering from Universiti Sains Malaysia in 2014. Her research interests are optimization and antenna design.

Dr. Kim Soon Choong

He obtained his B.Eng., M.S., and Ph.D. degrees in electrical and electronic engineering from Universiti Kebangsaan Malaysia in 2012, 2015, and 2019, respectively. He currently serves as an Assistant Professor in the Department of Computer Engineering (Artificial Intelligence) within the Faculty of Engineering Technology and Built Environment at UCSI Kuala Lumpur. His research focuses on areas such as medical signal processing and instrumentation.

Dr. Mohammad Arif Bin Ilyas

He (M'32) was born in Malaysia in 1991. He received his PhD in Electrical Engineering from Universiti Tun Hussein Onn Malaysia in 2020. He also earned his bachelor's degree in electronic engineering at the same university. Currently, he is a lecturer in the Faculty of Engineering, Technology, and Built Environment at UCSI University. His main areas of research interest are machine learning, artificial intelligence, the Internet of Things, embedded systems, and optical wireless communication, especially visible light communication.

Dr. Mastaneh Mokayef

She has received her PhD from Wireless Communication Centre Faculty of Electrical Engineering in University Technology Malaysia (UTM) in 2014. She has also obtained her master's degree from the faculty of engineering in 2009 from the University Technology Malaysia. Her research interests include Wireless communications, spectrum sharing method, spectrum management, cellular communication systems and Antenna design.

Ms. Zhi Ying Yeoh

She is currently pursuing Bachelor of Engineering Electrical and Electronics with Honours as final year student in Faculty of Engineering, Technology and Built Environment, UCSI University, Malaysia. Her research interests are power electronics and simulation.

Mr. Wei Kang Lai

He is currently pursuing Bachelor of Electrical and Electronics Engineering as final year student in Faculty of Engineering, Technology and Build Environment, UCSI University, Malaysia. His research interests are electrical and solar system design.

Dr. Wei Hong Lim

He is an Assistant Professor in Faculty of Engineering at UCSI University in Malaysia. He received his PhD in Computational Intelligence from Universiti Sains Malaysia in 2014. His research interests are optimization and artificial intelligence.

Investigate Power Efficiency in PLECS and MATLAB Software by Designing USB 5-Watt Charger

Zhi Ying Yeoh[1], Kim Soon Chong[1], Sew Sun Tiang[1], Mohammad Arif Ilyas[1*], Jia Wei Tan[1], Wei Kang Lai[1], Wei Hong Lim[1]

Email: 1001953809@ucsiuniversity.edu.my, ChongKS@ucsiuniversity.edu.my, tiangss@ucsiuniversity.edu.my, MohammadArif@ucsiuniversity.edu.my, 1001954348@ucsiuniversity.edu.my, 1001953677@ucsiuniversity.edu.my, limwh@ucsiuniversity.edu.my

([1]UCSI University, Malaysia)

Abstract

In power electronics, accurately assessing simulation tools is key for precise, reliable electronic system designs. This study compares MATLAB Simulink and PLECS in modeling a 5W USB charger's power characteristic. The charger, using an AC-DC full bridge rectifier and DC-DC flyback topology, delivers a stable 5VDC at 1A. The analysis focuses on power efficiency and thermal characteristics, incorporating real-life components for detailed insights. Results show PLECS, specialized in power electronics, surpasses MATLAB in accuracy and consistency. This research aids in understanding simulation tools' effectiveness, guiding engineers and researchers in power efficiency evaluations for electronic systems.

Keywords: Power electronics conversion, MATLAB Simulink, PLECS, USB charger, simulation, power efficiency, thermal analysis.

1. Introduction

Power electronics conversion is an essential technology utilized in the landscape of modern technology, permeating various aspects of daily life applications. Power electronics converters serve as a transformative agent in the regulation and control of electric energy flow.

A static power converter is a converter made up with the architecture of electrical components such as transformers, diodes, capacitors, and inductors. Electrical components work together transforming electrical energy from one form to another. From Fig. 1, notable ideal power converter can be observed, controlling the flow of power between two sources aiming to enhance the power efficiency [1]. Power conversion can delve into four categories which are DC to DC converter, AC to AC converter, DC to AC converter and AC to DC converter.
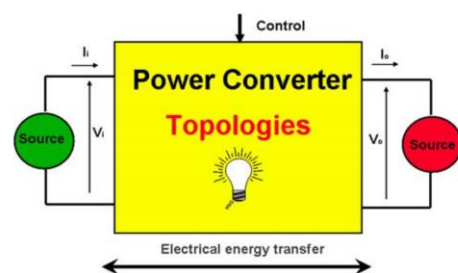
Fig. 1 Power converter topologies.

In recent years, significant improvements in power electronics technology. These advances include breakthroughs in converter topologies, control strategies, and power semiconductor devices, which together have contributed to the continuous improvement and widespread application of modern power conversion systems [2].

USB chargers are one of the illustrations of daily life, where the fundamental principles of power conversion. USB chargers are ubiquitous for charging electronics devices such as laptops and smartphones. In contemporary times, USB chargers are designed with emphasis on fast charging of electronics devices [3]. However, this design of faster charging rates can lead to increased energy

Zhi Ying Yeoh, Kim Soon Chong, Sew Sun Tiang, Mohammad Arif Ilyas, Jia Wei Tan, Wei Kang Lai, Wei Hong Lim

loss and heat generated, potentially diminishing the performance of power efficiency.

This paper focuses on using a USB 5-Watt charger as an evaluation method to evaluate the performance of MATLAB Simulink and PLECS software in accurately modeling and predicting the power characteristics of the charger. By conducting a detailed comparative analysis, this study aims to elucidate the disparities in results obtained from each software, shedding light on the strengths and limitations of MATLAB Simulink and PLECS in the context of power efficiency simulations.

The outcomes of this research will not only contribute to a better understanding of the capabilities of these simulation tools but will also assist engineers and researchers in making informed decisions when assessing power efficiency in various electronic systems.

2. Related Work

In the pursuit of advancing understanding of power electronics and simulation tools, a comprehensive review of related work lays the groundwork for our exploration. This section begins by exploring the AC-DC bridge converters, and DC-DC Flyback converter, essential components in designing the 5-Watt USB charger. By continuously adjusting the Proportional-Integral value to fine-tune and achieve the desired output voltage in the 5-Watt USB charger. Understanding converters is crucial for operating and evaluating their role in enhancing power efficiency, and thermal analysis.

2.1. *AC-DC Bridge Converter*

The AC-DC bridge converter, also known as the dual active bridge (DAB) converter, is a power electronic device used for converting alternating current (AC) to direct current (DC) and vice versa and the process known as rectification. The DAB converter, as outlined in Fig. 2 utilizes a bidirectional power flow and high-frequency isolation with a bridge topology featuring two sets of switches and a transformer [4].
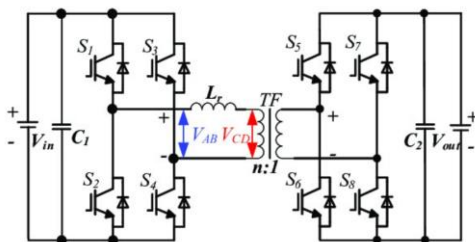
Fig. 2 DAB Converter Topology.

The DAB converter operates by modulating the phase-shift of the switches to control the power flow and achieve AC-DC or DC-AC conversion. It implements both frequency and time-domain analysis to capture the transient behavior of the converter, offering the advantages of steady-state frequency-domain analysis [5]. The converter can be configured to achieve full soft-switching operation, high efficiency, minimize distortion in input current, and provide output short-circuit protection [6]. One notably advantage of DAB is seamless four-quadrant operation [6], [7], [8], [9].

2.2. *DC- DC Flyback Converter*

A Flyback converter, specifically DC-DC converter is a type of power converter that uses a transformer to transfer energy from the input to the output.

The working principles of flyback converter is storing energy in the transformer during on time switching transistor, and then releasing it to the output during off time of switching transistor, as outlined in Fig. 3. This allows for voltage conversion and isolation between the input and output. The flyback converter can be used in various applications, such as AC-DC power supplies, power factor correction circuits, and active-clamp flyback converters [10].
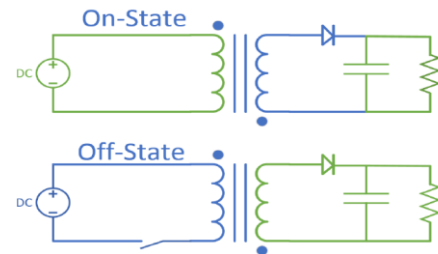
Fig. 3 The Operation of Two States of Flyback Converter.

2.3. *Proportional-Integral*

A Proportional-Integral (PI) controller is a type of feedback control system commonly used in power electronics to regulate the output of the power converters. It combines proportional (P) and integral (I) control actions to achieve the desired output.

The P value generates an output proportional to the current error, while the I value aids eliminating any residual steady-state error by integrating the error over time. In this project, the PI controller is implemented in the DC-DC Flyback converter in regulation to achieve desired output of USB charger.

2.4. *Thermal Analysis*

Thermal analysis is an important aspect of power USB charger configuration, due to it bringing effect on the reliability and safety of the system. The proper working of the charger's thermal considerations is a basic part of power electronics design. Several studies have been conducted on the thermal design of wireless chargers for electric vehicles [11], miniaturization and thermal design of battery chargers [12], and thermal analysis of transformers in on-board chargers for electric vehicles [13]. These studies use simulation tools to predict losses and temperature and establish thermal networks of heating components to calculate hotspot temperatures.

2.5. *Power Efficiency*

Power efficiency in USB chargers can be described as the charger's ability to convert electrical power from the source to the device, allowing it to be charged with minimal loss. It is a critical factor in determining the charging time and energy conservation.

Achieving high conversion efficiency is essential for optimal operation and robustness under varying conditions, as seen in the context of mobile battery charger ICs, where techniques like zero current switching (ZCS) are employed [14]. Therefore, power efficiency in USB chargers is essential for improving device charging, energy conservation, and the development of innovative charging technologies.

3. Materials and Methods

Overview of 5-Watt USB charger described overall input AC voltage to output DC voltage, as outlined in Fig. 4. The process begins with an input of 141/325 VAC with a frequency of 50/60Hz. This input voltage is directed through the AC-DC bridge converter. In this stage, the AC input voltage undergoes rectification, transforming it into a pulsating DC voltage. Subsequently, the pulsating DC voltage undergoes filtering and smoothing to ensure a stable waveform. The smoothed DC voltage then progresses through the DC-DC flyback converter, where it is stepped down to the desired level, providing isolation and regulation in the transformation process. To fine-tune and optimize the output results, a PI controller is applied, ensuring that the output voltage remains regulated at 5V DC with a current capacity of 1A.
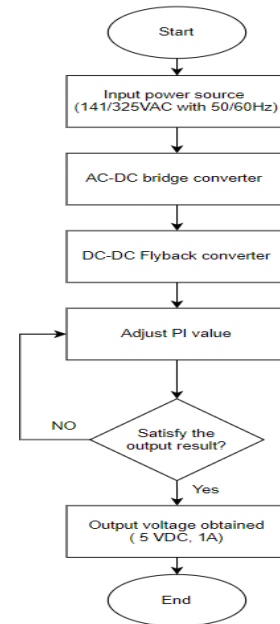
Fig. 4 Overview of 5-Watt USB Charger.

3.1. *The Design Procedure of 5-Watt USB Charger*

The design and implementation of a 5-Watt USB power charger involves the integration of an AC-DC full bridge rectifier employing a DC-DC flyback topology.

The primary objective of this charger is to deliver a maximum output current of 1A with a consistent output voltage of 5VDC. Achieving this goal requires the inclusion of a 5-ohm load resistor in the circuit.

To ensure the charger's adaptability for global use, it accepts two AC input voltages (V_{in}), as outlined in Table 1, providing corresponding output DC voltage, (V_{out}). The switching frequency, (f_{sw}) is set at 40kHz to optimize the performance of the charger.

Table 1. The Calculation of AC Voltage Input with Frequency.

AC Voltage	Input Voltage	Frequency
100V	100 x $\sqrt{2}$ = 141.42V	50/60Hz
230V	230 x $\sqrt{2}$ = 325.26V	50/60Hz

The Flyback equation was employed to calculate the winding ratio (n), critical inductance (Lc), and critical capacitance (Cc) for component selection in the high-frequency transformer and Flyback circuit capacitor. This calculation involves utilising known variables such as resistance (R), duty cycle (D) and output ripple voltage (ΔVo).

Number of turns ratio of transformer in flyback converter to calculate the value of Lc and Cc.

$$n = \frac{N2}{N1} = \frac{V_{out}(1-D)}{V_{in} * D} \qquad (1)$$

Critical Inductance and critical capacitor for Flyback circuit.

$$L_c = \frac{(1-D)^2 * R}{2 * f_{sw}} (\frac{N2}{N1})^2 \qquad (2)$$

$$C_c = \frac{D * V_{out}}{R * \Delta V_o * f_{sw}} \qquad (3)$$

Following the calculation of Lc and Cc using the Flyback equation, the appropriate high-frequency transformer, and inductor was selected.

To mitigate ripple voltage and ensure converter stability, a higher capacitor value was subsequently chosen in accordance with the formula. The selection of a higher capacitor value is essential for reducing ripple voltage, even though it may result in a higher peak current at the input. These components specifications are outlined in Table 2.

Table 2. The Specification of the Converter.

Converter Type	AC-DC Bridge Rectifier, DC-DC Flyback Topology
Switching Device	MOSFET (C3M0120065J by CREE was used)
Switching Frequency	20kHz -100kHz
HF Transformer	EE Series (EE16 74091 Flyback Transformer by MYRRA was used)
AC to DC Diode	IN5819
Capacitor	250uF
Inductor	6000uH
Controller Type	PWM PI Feedback Control (No overshoot and 0,1 Second Settling)
Input Voltage	Region Independent (230/100 VAC at 50/60Hz)
Output Voltage	5 VDC
Output Current	1A
Load Type	Resistor (5-ohm)

The proportional-integral (PI) value was tuned for a settling time within 0.1 seconds without overshooting.

For diode selection, the circuit's maximum current was observed on an oscilloscope, guiding the choice of a diode capable of withstanding the surge forward current. Finally, measurements were conducted to assess metal-oxide-semiconductor field-effect transistor (MOSFET) temperature and power efficiency.

3.2. *The Simulation on MATLAB Simulink and PLECS*

MATLAB Simulink and PLECS were utilized for running simulations, which play a significant role in the selection of components for the power converter. The design overview of the AC-DC 5-Watt USB charger on MATLAB Simulink and PLECS software are shown on Fig. 6 and Fig. 7. Additionally, the PWM PI feedback control was implemented to address steady-state error, requiring lower forward gain.

The PI controller ensures a stable 5V output under varying input voltages by generating a corrective signal to address system faults. Manual adjustments to the proportional and integral values of the PI controller were made until the desired output was stabilized without overshooting.

The flyback capacitor is adjusted to 250 microfarads during PI value tweaking to prevent excessive ripple. The PI value should be at P=0.06 and I=16, it was subsequently found.

The MATLAB Simulink environment is configured with the solver ode23t (modified Stiff/Trapezoidal), as outlined in Fig. 5 to facilitate analysis before conducting performance tests.

Fig. 5 The Configuration Parameters of the Simulink Solver.
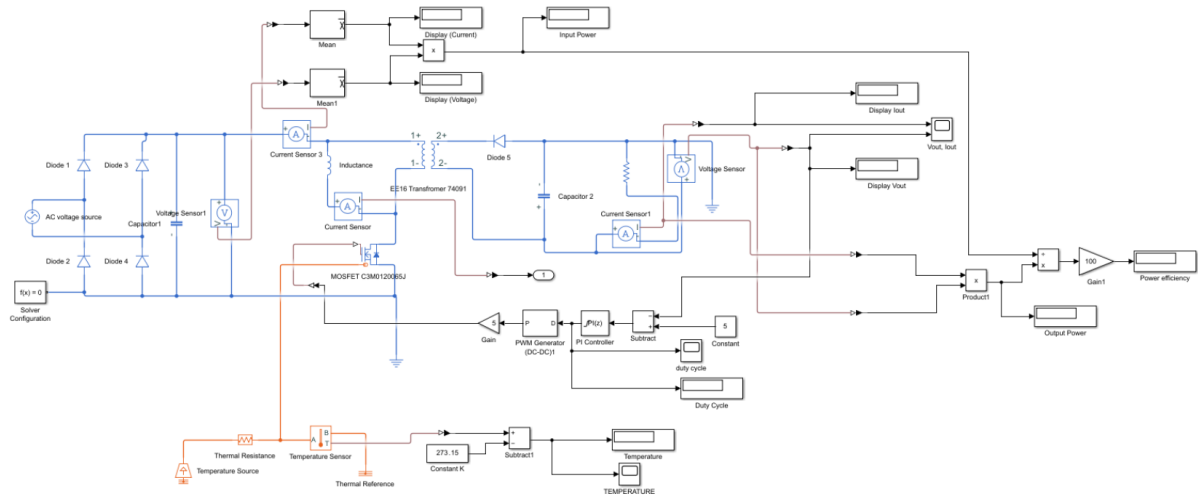
Fig. 6 The MATLAB Model of the AC – DC 5-Watt USB Charger.
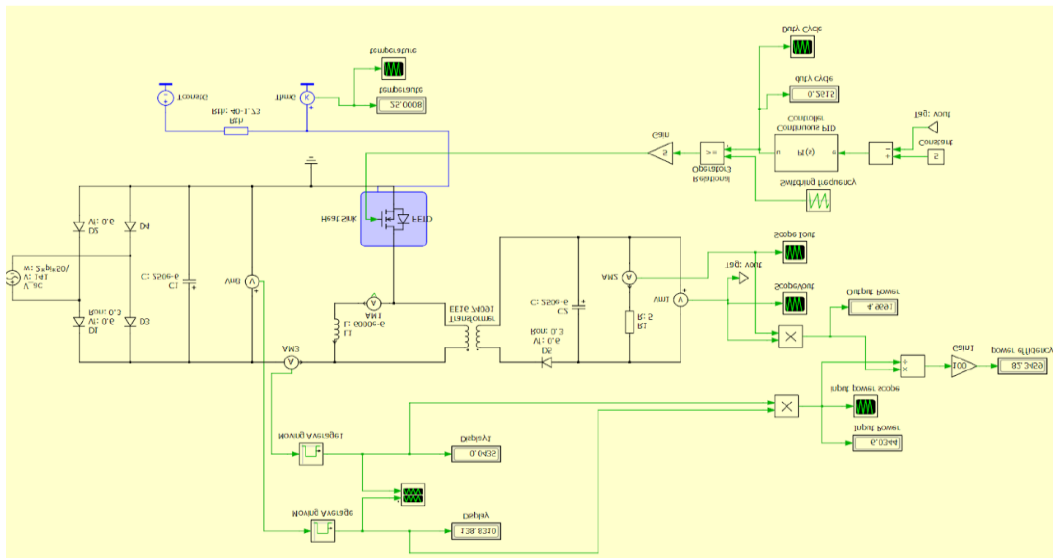
Fig. 7 The PLECS Model of the AC – DC 5-Watt USB Charger.

4. Results and Discussion

The 5-Watt USB Charger simulation results, as detailed in Table 3, are presented utilizing both MATLAB Simulink and PLECS. The simulations were conducted using a f_{sw} value of 40kHz and a PI value of P=0.06 and I=16. The results from both simulations are tabulated and compared to each other.

Zhi Ying Yeoh, Kim Soon Chong, Sew Sun Tiang, Mohammad Arif Ilyas, Jia Wei Tan, Wei Kang Lai, Wei Hong Lim

Table 3. Comparison of the Simulation Results of the AC – DC 5-Watt USB Charger.

Voltage Region	Frequency	Input Voltage Input Current Input Power		Output Voltage Output Current Output Power	
		MATLAB	PLECS	MATLAB	PLECS
141V	50Hz	138.9V 45mA 6.31W	138.83V 43.5mA 6.03W	5.002V 1.00A 5.003W	4.985V 997mA 4.969W
	60Hz	138.9V 45mA 6.307W	138.9V 43.4mA 6.034W	5.002V 1.00A 5.004W	4.985V 997mA 4.970W
325V	50Hz	323.1V 22.6mA 7.307W	323.23V 18.5mA 5.989W	4.996V 999mA 4.992W	4.987V 997mA 4.973W
	60Hz	323.2V 22.5mA 7.302W	323.25V 18.5mA 5.989W	4.996V 999mA 4.992W	4.987V 997mA 4.974W

4.1. *The Output Voltage, Current and Power*

The result indicates that the USB charger produces an output voltage of 5VDC and an output current of 1A and shown in Fig. 8.
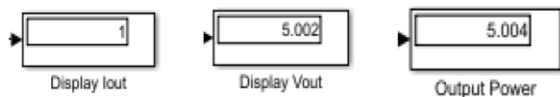
Fig. 8 The Scope of Voltage, Current, and Power Output at 325/100V AC with 50/60Hz.
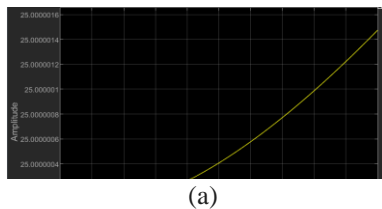
(a)
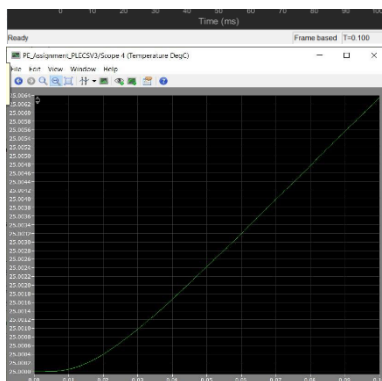
(b)

Fig. 9 The scope thermal analysis at 230/100VAC at MATLAB Simulink (a) and PLECS (b).

4.2. *Thermal Analysis*

Fig. 9 performs the scope thermal analysis of the input voltage at 230/100 VAC in MATLAB Simulink (a) and PLECS (b). The result of MOSFET temperature reaches a steady-state value of 25 °C in MATLAB Simulink and 25.06 °C in PLECS, after a simulation time of 0.1 seconds.

Table 4 performs detail comparison on the MOSFET temperature after simulation at different voltage regions and frequencies.

The MOSFET temperature is slightly higher in PLECS than in MATLAB Simulink. This is consistent with the results of the simulation setup, which showed that PLECS software utilized a more detailed thermal model of the MOSFET. As can be seen from Fig. 9 (a) and (b), the MOSFET temperature is slightly higher in PLECS than in MATLAB Simulink.

Table 4. The MOSFET Temperature After Simulation

Voltage Region	Frequency	MOSFET Temperature	
		MATLAB	PLECS
141V	50Hz	25 °C	25.06 °C
	60Hz	25 °C	25.06 °C
325V	50Hz	25 °C	25.01 °C
	60Hz	25 °C	25.01 °C

4.3. *Power Efficiency*

Power efficiency results can be observed on differences input voltages and frequencies in Table 5. The table showed power efficiency was also compared with using MATLAB and PLECS software. The results in the MATLAB are between 68% and 79%. On the other hand, the power efficiency in PLECS is more consistent at 82.35% to 83.04%. It is notably underscoring the difference in the power efficiency in both software.

Table 5. Power Efficiency After Simulation.

Voltage Region	Frequency	Power Efficiency	
		MATLAB	PLECS
141V	50Hz	79.29%	82.35%
	60Hz	79.35%	82.37%
325V	50Hz	68.31%	83.04%
	60Hz	68.37%	83.04%

Comparing to MATLAB Simulink, the performance of PLECS on simulation power efficiency USB charger exhibits a 3.06% improvement at 141 VAC with 50/60Hz. Moreover, PLECS further outperforms, demonstrating a significant 14.73% increase in power efficiency with a 325V AC input voltage at 50/60Hz frequency.

MATLAB is a general-purpose software used in engineering fields for numerical analysis, data visualization, and simulation of complex systems, using a discrete-time simulation method. MATLAB's general approach to simulation may not be as accurate in modelling power electronics components, resulting in less consistent power efficiency results.

PLECS is designed specifically for simulating power electronics systems, using a continuous-time simulation method with specialized models for simulating power electronic components. The different simulation methods and models used by these two software tools could result in variations in simulation outcomes, including power efficiency.

PLECS specialized models accurately represent the behaviour of power electronic components leading to more precise simulations and consistent power efficiency results.

While the same datasheet was implemented, this does not demonstrate that the MATLAB Simulink software is superior for simulating the circuit because the MOSFET's specific parameter differs from that of the PLECS Software. This is because the MOSFET in the PLECS Software is based on the characteristics of the actual MOSFET that have been implemented by the manufacturer, CREE. On the other hand, the parameters for the MOSFET in the MATLAB/Simulink Software are manually entered based on the datasheet.

When simulated in 0.1 seconds, the MOSFET temperature of each software is slightly comparable. Nevertheless, the temperature of the simulation in PLECS Software is slightly higher than that of the simulation in MATLAB Simulink Software. This is because MOSFET's specific parameter varies depending on the software.

In MATLAB Simulink software, the MOSFET is manually inserted based on the datasheet, which may result in less precision than the model used in the PLECS software. Whereas in PLECS software, MOSFET is based on the manufacturer, CREE. The parameters and thermal model of MOSFET are needed to add into PLECS thermal library.

The ways to solve the numerical methods involved to solve the system of equations also can impact the outcome of simulations in both MATLAB Simulink and PLECS Software, which might result in some degree of error throughout the simulation results.

Step size, solver parameters, and the accuracy of the used models are only a few of the variables that affect how accurate the simulation is. It is advised to decrease the step size or raise the solver accuracy in both MATLAB and PLECS software and evaluate if the simulation results converge to improve the accuracy of the simulation results.

5. Conclusion

This paper successfully constructs a 5-Watt USB power charger, exploring AC to DC full bridge rectifier and DC-to-DC Flyback converters. Each circuit component's parameter adjustments produce the 5VDC, 1A output, necessitating a PI controller for stability. This controller manages P = 0.06 and I=16, minimizing the difference between reference and actual voltage. Two simulation software were compared for their performance and outcomes. PLECS excels, designed specifically for power electronics, offering greater accuracy and faster, memory-efficient, showing a 3.06% advantage at 141V and 50/60Hz. Moreover, at 325V AC input voltage with 50/60Hz frequency, PLECS excels further, exhibiting 14.73% increase in power

Zhi Ying Yeoh, Kim Soon Chong, Sew Sun Tiang, Mohammad Arif Ilyas, Jia Wei Tan, Wei Kang Lai, Wei Hong Lim

efficiency compared to MATLAB Simulink. The simulation study provides a comprehensive understanding of converter interactions and highlights PLECS Software advantages in power electronics simulations over MATLAB Simulink Software.

Acknowledgement

References

1. F. Bordry and D. Aguglia, "Definition of Power Converters," CAS - CERN Accelerator School: Power Converters, Baden, Switzerland, pp. 15–43, Jun. 2015.

2. M. G. Simões and F. A. Farret, Modeling Power Electronics and Interfacing Energy Conversion Systems. Wiley, 2016.

3. C.-J. Chen, C.-H. Cheng, P.-S. Wu, and S.-S. Wang, "Unified Small-Signal Model and Compensator Design of Flyback Converter With Peak-Current Control at Variable Frequency for USB Power Delivery," IEEE Trans Power Electron, vol. 34, no. 1, pp. 783–793, Jan. 2019.

4. J. Zhai, F. Wang, L. Ye, and R. Li, "Optimal Strategy of Triple Phase Shift Transient DC Bias for Dual Active Bridge DC-DC Converter," E3S Web of Conferences, vol. 185, p. 01088, Sep. 2020.

5. A. Singh, A. K. Yadav, and A. Khaligh, "Steady-State Modeling of a Dual-Active Bridge AC–DC Converter Considering Circuit Nonidealities and Intracycle Transient Effects," IEEE Trans Power Electron, vol. 36, no. 10, pp. 11276–11287, Oct. 2021.

6. C. S. Wong, K. H. Loo, and L. Cao, "A Single-stage Three-phase Bidirectional AC-DC IPT Converter based on SWISS-Rectifier for EV Charging Applications," in IECON 2022 – 48th Annual Conference of the IEEE Industrial Electronics Society, IEEE, Oct. 2022, pp. 1–6.

7. Filho, "Predictive Control Applied to a Single-Stage, Single-Phase Bidirectional AC-DC Converter," IEEE Access, vol. 10, pp. 34984–34995, 2022.

8. S. Bagawade, I. Askarian, M. Pahlevani, and P. Jain, "A New Discrete Four-Quadrant Control Technique for Grid-Connected Full-Bridge AC–DC Converters," IEEE J Emerg Sel Top Power Electron, vol. 10, no. 2, pp. 2362–2379, Apr. 2022.

9. J. Liu, Y. Cheng, Y. Jia, and H. Wu, "Optimal Design of a Constant Frequency Series-Resonant DC/DC Converter with Wide Voltage Gain Range for Single-Stage Isolated AC/DC Power Conversion," in 2022 International Power Electronics Conference (IPEC-Himeji 2022- ECCE Asia), IEEE, May 2022, pp. 1036–1043

10. C. Wang, D. Sun, X. Zhang, W. Gu, and S. Gui, "A Constant Current Digital Control Method for Primary-Side Regulation Active-Clamp Flyback Converter in CCM Mode," in 2021 IEEE Energy Conversion Congress and Exposition (ECCE), IEEE, Oct. 2021, pp. 3501–3506.

11. B. Zhang, J. Deng, L. Li, Z. Wang, S. Wang, and G. Guidi, "Thermal Analysis and Design of a 30kW EV Wireless Charger with Liquid-Cooled Shell for Magnetic Coupler and Integrated Power Converter," in 2021 IEEE Applied Power Electronics Conference and Exposition (APEC), IEEE, Jun. 2021, pp. 426–431.

12. J. Weimer, D. Koch, M. Nitzsche, J. Haarer, J. Roth-Stielow, and I. Kallfass, "Miniaturization and Thermal Design of a 170 W AC/DC Battery Charger Utilizing GaN Power Devices," IEEE Open Journal of Power Electronics, vol. 3, pp. 13–25, 2022.

13. B. K. Kushwaha, R. R., S. Jadhav, and S. T. S., "Thermal Analysis of 2.2kW Transformer of On-Board Charger for Electric Vehicles," in 2023 Second International Conference on Electronics and Renewable Systems (ICEARS), IEEE, Mar. 2023, pp. 311–314.

14. Y. D. Kuan, S. R. Lee, S. M. Lee, M. F. Sung, and H. S. Chiou, "The Performance Analysis of a Portable Charger System Using Direct Methanol Fuel Cells," in Proceedings of the 35th International MATADOR Conference, London: Springer London, 2007, pp. 129–132.

Authors Introduction

Ms. Zhi Ying Yeoh

She is currently pursuing Bachelor of Engineering Electrical and Electronics with Honors as final year student in Faculty of Engineering, Technology and Built Environment, UCSI University, Malaysia. Her research interests are electrical and power electronics.

Ms. Jia Wei Tan

She is currently pursuing Bachelor of Engineering Communication and Electronics with Honors as final year student in Faculty of Engineering, Technology and Built Environment, UCSI University, Malaysia. Her research interests are communication and antenna design.

Mr. Wei Kang Lai

He is currently pursuing Bachelor of Electrical and Electronics Engineering as final year student in Faculty of Engineering, Technology and Build Environment, UCSI University, Malaysia. His research interests are electrical and solar system design.

Dr. Sew Sun Tiang

She is an Assistant Professor in Faculty of Engineering at UCSI University in Malaysia. She received her PhD in Electrical and Electronic Engineering from Universiti Sains Malaysia in 2014. Her research interests are optimization and antenna design.

Dr. Kim Soon Choong

He obtained his B.Eng., M.S., and Ph.D. degrees in electrical and electronic engineering from Universiti Kebangsaan Malaysia in 2012, 2015, and 2019, respectively. He currently serves as an Assistant Professor in the Department of Computer Engineering (Artificial Intelligence) within the Faculty of Engineering Technology and Built Environment at UCSI Kuala Lumpur. His research focuses on areas such as medical signal processing, etc..

Dr. Mohammad Arif Bin Ilyas

He received his PhD in Electrical Engineering from Universiti Tun Hussein Onn Malaysia in 2020. Currently, he is a lecturer in the Faculty of Engineering, Technology, and Built Environment at UCSI University. His main areas of research interest are machine learning, artificial intelligence, the Internet of Things, embedded systems, etc..

Design and Simulation and Performance of Grid Connected Photovoltaic System for Small, Tall Building in Malaysia

Wei Kang Lai[1], Kim Soon Chong[1,*], Mohammad Arif Bin Ilyas[1], Sew Sun Tiang[1], Zhi Ying Yeoh [1], Jia Wei Tan[1], Wei Hong Lim[1]

[1]Department of Electrical and Electronics Engineering, Faculty of Engineering, Technology and Built Environment, UCSI University, 56000 Cheras, Kuala Lumpur, Malaysia
E-mail: 1001953677@ucsiuniversity.edu.my, chongks@ucsiuniversity.edu.my, MohammadArif@ucsiuniversity.edu.my, tiangss@ucsiuniversity.edu.my, 1001954348@ucsiuniversity.edu.my, 1001953809@ucsiuniversity.edu.my, limwh@ucsiuniversity.edu.my
([1]UCSI University, Malaysia)

Abstract

In Malaysia's rapidly urbanizing landscape, sustainable energy for small, tall buildings is increasingly vital. This study addresses this need through the design, simulation, and performance analysis of grid-connected photovoltaic systems tailored for these unique structures. Utilizing AutoCAD and PVSyst for design and simulation, the research details rooftop array dimensions, PV panel wiring, and system components aligned with MS 1837-2018 standards. Results indicate the designed system can save 526.70MWh annually, costing RM 339,306.98, with a 9.8% return on investment (ROI) and a 13-year breakeven point, emphasizing its sustainability and economic viability.

Keywords: PVSyst, AutoCAD, Cost saving, ROI, Breakeven point

1. INTRODUCTION

In recent years, as non-renewable energy sources such as coal, oil, natural gas, chemical energy, and nuclear fuel have begun to run out and sustainable development has begun to be threatened, the use of renewable energy has become increasingly important. All these energy sources are also harmful to the environment and can lead to environmental problems such as greenhouse gas (GHGs) emissions and hazardous chemical waste [1], [2], [3]. Therefore, research and development of renewable energy sources is crucial.

Among various renewable energy options, rooftop solar photovoltaic (PV) systems are a popular and effective method for collect solar energy [1]. A grid-connected rooftop solar photovoltaic system is a solar power generation system that is connected to the grid and can transmit excess energy back to the grid. In addition, Malaysia is one of the countries with the greatest potential for solar energy utilization due to its strategic location near the equator. Through the research, the monthly solar irradiance for Malaysia is estimated at 400-600 MJ/m^2 as well as its hot and sunny weather throughout the year, thus the potential for solar power is very large [4]. However, the design and implementation of a grid-connected rooftop solar PV system requires careful consideration of several factors, including the selection of appropriate solar panels, system sizing, inverter selection, and system integration with existing electrical infrastructure.

In Malaysia, the rapid growth of urbanization has led to an increased demand for sustainable energy solutions. Despite the potential benefits of grid-connected photovoltaic (PV) systems in meeting this demand, there is a lack of comprehensive research addressing the design, simulation, and performance analysis tailored specifically for the unique challenges posed by small, tall buildings in the Malaysian context. The main objective of this paper is to bridge this gap by investigating the optimal design parameters, conducting detailed simulations, and performing a thorough performance analysis of grid-connected PV systems. The outcomes of this study are expected to provide valuable insights into the feasibility and efficiency of implementing grid-connected PV systems in this specific urban setting, contributing to the advancement of sustainable energy solutions for the region.

2. METHODOLOGY
2.1. *Calculation and Selection for Solar Panel*

In this study, one of the buildings in the UCSI university was selected for installing the PV panels. The selected building, named Block C, located at No. 1, UCSI Heights, Jalan Puncak Menara Gading, Taman Connaught, 56000 Cheras, Federal Territory of Kuala Lumpur. As shown in Fig. 1, the satellite view shows that there are 12 faces on the rooftop

Wei Kang Lai, Kim Soon Chong, Mohammad Arif Bin Ilyas, Sew Sun Tiang, Zhi Ying Yeoh, Jia Wei Tan, Wei Hong Lim

with the tilts of 30 degrees in each azimuth. The area of the rooftop was calculated to be $3194m^2$. However, only 8 faces are decided to install the PV panels which is $2129m^2$ and the other 4 faces are for cleaning and technical maintenance. These data were then used to create a 3D model of the rooftop in 3D AutoCAD, and the appropriate PV array size was decided using 1120 modules of SunPower X22-360 solar panels based on calculated by using expected power for each orientation which is 50kW and spaces on the rooftop. Below is the calculation:

$$Area\ for\ placing\ PV\ panels = \frac{3,194m^2}{12} \times 8$$
$$= 2,129m^2$$

$$No.\ of\ PV\ panels = \frac{50,000W}{360W}$$
$$= 140\ modules$$

$$Total\ no.\ of\ PV\ panels = 140\ module\ \times 8$$
$$= 2,129m^2$$

$$Power\ generated = 1,120\ modules\ \times 8$$
$$= 403,200W$$

Apart from that, to convert the DC power generated by the solar panels into AC power, the model of Goodwe GW50KLV-MT inverter had been selected.

The SunPower X22-360 Solar panel is a high-performance solar panel designed for residential and commercial applications. In a six-story building, SunPower's X22-360 solar panels can be used to generate renewable energy to offset the building's energy consumption. The electricity generated by the solar panels can be used to power lighting, heating, ventilation, and air conditioning (HVAC) systems and other electrical equipment in the building, while excess energy can be fed back to the grid to generate revenue for the university through a net metering program. Additionally, the X Series panels have a 25-year warranty and are made to be long-lasting and robust. This long-term durability improves X-Series panels a cost-effective, sustained producing electricity for many years [5].

The Goodwe GW50KLV-MT is a three-phase grid-tied inverter intended to transform the direct current (DC) power produced by the solar panel arrays into alternating current (AC) power that can be supplied into the utility system. Goodwe GW50KLV-MT has also built-in safety features like ground fault protection (GFP) and surge protection devices (SPDs) for both DC and AC inputs, which provide both the safety of the solar PV system and the building's occupants [6].

The inverters of the system were installed at exterior placement to create a more spacious indoor environment and address safety concerns related to preventing electrical hazards. Placing the inverters at exterior placement not only minimizing energy waste from cable resistance but also enhance efficiency by reducing the length of the DC wire between the solar panel and the inverter.

Figure 1: Satellite view of a building's rooftop

2.2. *Solar panel layout design*

The floor plan of the roof PV panels was construct based on the Google Maps satellite images as shown in Fig. 1. The dimensions of the SunPower X22-360 were obtained from the data sheet in order to arrange the panels in the most effective and aesthetically acceptable way possible. The route of the DC wires from the PV panels to the inverter and the AC cables from the inverter to the electrical panel was also determined at this point. Thorough the analysis, a 3D model of PV system design on the rooftop was showed in 3D AutoCAD shown in Fig. 2.
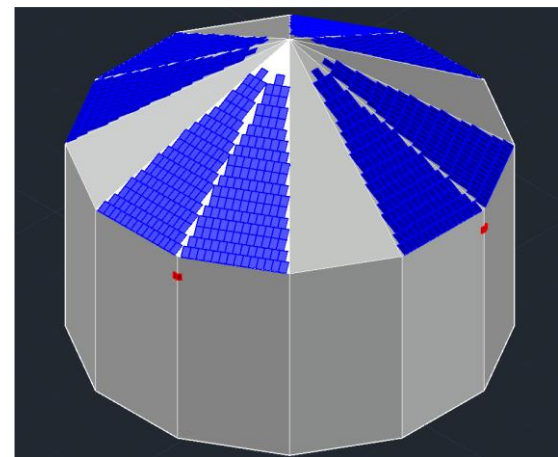
Figure 2: 3D Modelling rooftop design

2.3. Wire Connection

Leapfrog chain wiring, as outlined in Fig. 3 is the best and suitable wiring design for string panels that are not organized in a straight line due to its simplicity, ease of use, cost saving and electromagnetic loop reduction which represent less voltage drop [7], [8].
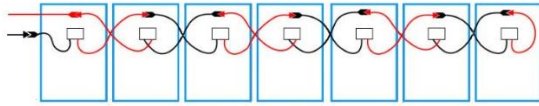
Figure 3: Leapfrog Wiring

According to the MS 1837-2018 standard, single-line diagram of the system shown in Fig. 4 was designed and outlined the minimum

requirements for the installation of grid-connected PV systems. This diagram provides a clear and concise representation of the system components, including the arrangement of PV array, inverters, DC and AC disconnects, and electrical panels. It also includes details on the system grounding and protection devices, such as surge protectors and overcurrent devices [9].

Figure 4: Single-line diagram

In Fig. 4, the direct current (DC) electricity produced by the 5 strings of 7 solar panels was routed through DC surge protective devices (SPDs) positioned before and after the PV DC switch. The DC SPDs safeguard the system against surges and overvoltage. The DC electricity was then routed through the PV DC switch to the eight generators. Each inverter has four MPPT (maximum power point tracking) sources to optimize electricity production from the panels.

The wires before and after the PV DC switch and the wires before and after the PV AC switch were selected as MC4 compatible for the head and tail of the wire and through calculation. Due to the design of 7 module of solar panels in series connection and each module can generate 60.6V at maximum power point, the output of the voltage will be 424.2V.

Through the calculation, the voltage for the wires required to withstand is at least 424.2V. Thus, due to the critical safety aspect, THHN/THWN-2 wire which can withstand 600V was selected to prevent the wiring from overheating and overload that may cause fire hazard and short circuit. While the other reason for choosing THHN/THWN-2 wire was the temperature rating for this wire is 90 degrees Celsius and could use in wet conditions [10].

2.4. Selection in PVsyst software

In Fig. 5, it depicts the desired tilts and azimuths. Azimuths of 0°, 30°, 90°, 120°, 180°, -120°, -90°, -30° have been selected for placing the solar panels on the rooftop.

Figure 5: Orientations (Tilts and Azimuths)

In the PVSyst software, the selection and application of dimensions, photovoltaic modules and inverters are shown in Fig. 6. The SunPower X22-360 model for the solar panel and the Goodwe GW50KLV-MT model for the inverter were chosen

for each orientation in PVSyst software. The inverters will conduct 4 MPPTs each, as the operating DC voltage needs to fall within the MPPT voltage range of the inverter. Therefore, 5 strings that are connected in parallel will be connected to one of the MPPT for the inverter. A total of 20 strings, each consisting of 7 solar panels per string, will be subject to application based on the expected power determined in PVSyst software, which is 50 kW.

After all the values and the selection was applied, the voltage maximum power point obtained is 367V and 418V at 60 degrees and 20 degrees, while AC voltage is 531V. Moreover, the value of array nominal power at Standard Test Conditions (STC) also determine as 50.4kWp.

Besides, to account for the various variables that may influence the system's real performance as opposed to the idealized circumstances anticipated in laboratory testing, the Thermal Losses, Ohmic Losses, Power Loss at Maximum Power Point(MPP), Soling Loss, and Aging Losses has been applied in the PVsyst software. These variables may include shading, weather, soiling, inefficient modules, and inverters, among other practical and environmental problems that could eventually result in a drop in power production. It is possible to get a more accurate estimate of the system's real energy yield and financial success by including these losses in the simulation. It is also possible to find prospective areas for optimisation and development.

Figure 6: Sizing and system applied in PVSys

2.5. *Arrangement of Solar Panels*

The spacing between the solar panels is an essential factor in the design of solar photovoltaic systems, as shown in Fig. 7. The rationale behind providing adequate space was to ensure that there is proper ventilation and air flow around the panels. This aids to keep the panel cool, subsequently enhances performance.

Minimizing shading is imperative, it can significantly reduce the output of photovoltaic systems. Therefore, by leaving space between panels, can mitigate the risk of shading, which occurs when one panel casts a shadow on another panel.

The accumulation of dust or debris can also degrade the panel's performance [11]. Therefore, leaving space between panels is essential, providing convenient for maintenance and cleaning. Overall, allocating sufficient space for solar panels is vital, ensuring optimal performance, power efficiency, and longevity of solar photovoltaic systems.

Figure 7: Spacing between Solar panels

2.6. *Budget of the project*

An expenditure of 911,192 dollars, equivalent to RM 4,309,026.97, as shown in Table 1, was anticipated for the acquisition of PV panels and solar inverters utilized in this project.

Table 1: Total cost of material

Product	Model	Total unit	Single unit price (dollar)	Total price of all units (dollar)
PV panel	SunPower X22-360	1,120	800	896,000
Solar inverter	Goodwe GW50KLV-MT	8	1,899	15,192
Total				911,192

3. RESULT AND DISCUSSIONS
3.1. *PVsyst Analysis*

The array loss analysis simulated by PVSyst software as outlined in Fig. 8. The series diode typically exhibits a default voltage drop of 0.7 V. Concurrently, ohmic losses emerge from the resistance in wires and connectors, contributing to a potential decline in the solar system's output power.

To quantify these losses within the subfield, the loss fraction was specified as STC which is 2%. In this scenario, the solar system achieves a total power output of 400 kW at STC and ohmic losses are stipulated at 2%, the resultant power loss due to ohmic factors amounts to 2 kW.

Apart from that, the yearly soiling loss factor indicates the decrease in solar panel energy production caused by dust and dirt accumulation on the surface. It varies depending on the location and environmental conditions of the solar panel installation such as rainfall frequency, wind patterns, and proximity to dust sources can affect the rate of soiling. The annual soiling loss factor was found to be 3% analysed by PVSyst software, reducing solar panel's performance. This outcome can be used to determine the cleaning frequency needed to ensure the solar panel performing optimally.

The PVSyst module analysis revealed an average degradation of 0.4% per year with a 10-year loss factor. This indicates that the module's efficiency is expected to decrease by 0.4% each year over the course of 10 years.

The analysis also revealed a mismatch due to degradation and resulting in Root Mean Square of Current (Imp RMS) and Root Mean Square of Voltage (Vmp RMS) dispersions of 0.4% per year. Owing to the degradation of the panels, the current and voltage output from the solar panels will experience slight deviations over time. These findings highlighted the importance of regular maintenance and monitoring of the solar panels to ensure optimal performance and efficiency over the long-term.

Figure 8: Array Loss analysis

Within the total PV power characteristic, the total number of utilized PV modules is 1,120, generating a nominal (STC) output of 403 kWp as shown in Fig. 9, which is consistent with the calculated value. The covered module area and unit area are 1,826m^2 and 1,648 m^2 respectively.

The total power of the inverters that have converted was 400kWac and the number of the converter used is 8 units. The ratio of PV to inverter power is measured as the DC/AC ratio whereby ideal value is 1.25. Nevertheless, the simulation analysis indicates the power nominal (Pnom) ratio is 1.01, falling short of the recommended healthy ratio for the PV system.

Total PV power

Nominal (STC)	403 kWp
Total	1120 modules
Module area	1826 m²
Cell area	1648 m²

Total inverter power

Total power	400 kWac
Number of inverters	8 units
Pnom ratio	1.01
No power sharing	

Figure 9: Total PV power and inverter power

In Fig. 10, it shows the total global horizontal irradiation (GlobHor) for the given area was found to be 1,783.4 kWh/m^2, whereas the total horizontal diffuse irradiation (DiffHor) was calculated to be 963.42 kWh/m^2. During this period, the average ambient temperature (T_Amb) was 27.72℃. The total global incidence in coll. plane (Globinc) was calculated to be 1,641.8 kWh/m^2, indicating a reduction in the total energy incident on the module due to the tilt angle and orientation.

According to the PVSyst analysis the obtained results affirm that the PV system was performing well in terms of energy production. The high specific production or high output of 1,306 kWh/kWp/year indicates that the system was well-designed and well-maintained, which was a positive sign. In addition, the total energy injected into the 526,695kWh grid indicated a critical commitment to the overall electricity supply on site.

The analysis of the results indicates that, in the given location, the PV module tilt point and the degree of inclination are not optimally configured. This was reflected in the comparatively lower value of Globinc in comparison to GlobHor.

By focusing on the DiffHor value of 963.42 kWh/m2, the location of the PV system is not highly shaded, which was a positive factor for energy production. The average ambient temperature during the study period of 27.72℃ was considered within the cycle and is also within the ideal temperature range for photovoltaic system performance.

Apart from that, the average performance ratio (PR) that indicates the efficiency of the system in converting the solar energy into electrical energy was calculated to be 0.796, equivalent to 79%, which is close to 80%. Nevertheless, for the year 2010 is found to be between ~70 and ~90% and shows a median PR of 84%. However, the analysis shows that it does not getting 84% the PV system due to the imperfection of tilts. To solve this problem, it is recommended to add an adjustable mounting system for optimal tilt and performance [12].

Main results

System Production

Produced Energy	526695 kWh/year		Specific production	1306 kWh/kWp/year
			Performance Ratio PR	79.57 %

Balances and main results

	GlobHor	DiffHor	T_Amb	GlobInc	GlobEff	EArray	E_Grid	PR
	kWh/m²	kWh/m²	°C	kWh/m²	kWh/m²	kWh	kWh	ratio
January	154.1	81.19	27.23	145.0	139.1	47350	46599	0.797
February	159.5	75.66	27.83	148.3	142.4	48208	47478	0.794
March	170.3	83.83	28.17	156.9	150.6	50936	50127	0.792
April	164.1	76.77	27.77	149.6	143.6	48652	47858	0.793
May	156.9	85.34	28.61	142.6	136.7	46442	45669	0.794
June	136.0	73.75	28.09	123.0	117.8	40156	39457	0.795
July	144.9	75.50	28.10	130.9	125.5	42688	41949	0.795
August	147.3	94.46	28.09	134.4	128.8	44036	43307	0.799
September	145.1	85.35	27.28	133.5	128.0	43740	43013	0.799
October	141.5	81.72	27.56	131.0	125.6	42757	42021	0.796
November	130.5	75.15	26.75	121.6	116.5	39819	39118	0.798
December	133.2	74.70	27.17	124.9	119.7	40823	40099	0.796
Year	1783.4	963.42	27.72	1641.8	1574.4	535605	526695	0.796

Figure 10: Balance and main results

In this study, the efficiency of the PV system has comprehended by the system output power distribution graph as shown in Fig. 11. The graph shows that the system operates at optimum efficiency, with a peak energy injection of 9 MWh/class of 2.5 kW, corresponding to a power injection of 240 kW. Notably, the majority of the energy injected into the grid falls within a stable range of 5 to 7 MWh/class of 2.5 kW, suggesting efficient and consistent operation. The distribution graph provides useful insights into the system's performance that can be used to find potential areas for development and optimize the functioning of the PV system. By knowing the power levels at which the system produces the most energy, the system can be operated to maximize energy output at these power levels.
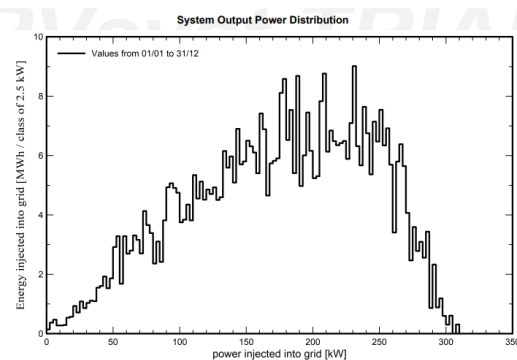
Figure 11: Power distribution graph

3.2. *Estimate Cost Saving, Breakeven point and Return of Investment (ROI)*

According to the analysis, 526.70 MWH can be generated throughout the year, with an average of 43.891 MWH per month. The electricity bill is calculated using the formula provided by Tenaga National Berhad (TNB), as shown below:

Electricity Bill = (Consumed electricity x First 200 kWh rate) + (Consumed electricity x Next 100 kWh rate) + (Consumed electricity x Next 300 kWh rate) + (Consumed electricity x Next 300 kWh rate) + (Consumed electricity x Next 901 kWh onwards per month) *(6% tax)

Wei Kang Lai, Kim Soon Chong, Mohammad Arif Bin Ilyas, Sew Sun Tiang, Zhi Ying Yeoh, Jia Wei Tan, Wei Hong Lim

Table 2: Year of the ROI

	Electricity Bill
Per Month	RM 28,146.24
Per Year	RM 339,306.98

Through the calculation, user can save minimum up to RM 28,146.24 per month and RM 339,306.98 per year as shown in Table 2. Besides, the use of electricity from the solar panels are prioritised instead of using electricity from the TNB grid whereby surplus electricity will not be wasted and exported back to TNB.

$$ROI = \frac{Net\ profit}{Cost\ of\ Investment} \times 100$$

$$Year\ 1 = \frac{RM339,307 - RM4,350k}{RM4,350k} \times 100$$
$$= -92\%$$

Table 3: Year of the ROI

Year	ROI = (Net profit / cost of investment) x 100
1	-92%
2	-84.4%
3	-76.6%
4	-68.8%
5	- 61%
6	-53.2%
7	-45.4%
8	-37.6%
9	-29.8%
10	-22%
11	-14.2%
12	-6.4%
13	1.4%
14	9.2%

Breakeven point (in year)
= Fixed Costs / Gross Profit Margin
= RM 4,309,026.97/ RM 339,306.98
= 12.69 year

At the beginning of the investment, it resulted in a net loss and no returns due to the high upfront cost of the system which is RM 4,309,026.97 [13]. In Year 13, the ROI has become positive after 12 years as shown in Table 3. Beyond the 13th year, it is expected to be in a profit-making state, where the saving from the electrical bill exceeds the total cost of panels ordered. Regarding the breakeven point analysis, it shows this project needs to take 12.69 year which is almost 13 years to reach the breakeven point.

4. CONCLUSION

This study examines the potential of installing a solar photovoltaic (PV) system at a small, tall building in Malaysia. The analysis considers factors such as the building's energy demand, available space, and the financial feasibility of the proposal. The findings of the analysis reveal that a 403-kW solar PV system has the capacity to generate approximately 526695 kWh of electricity and helps to save up to RM 339,306.98 per year which can satisfy the building's power requirements.

Apart from cost saving and power supply, the performance of PV systems is not optimal. In order to take full advantage of the potential capabilities of the system and reduce performance differences, key recommendations for the future include the adoption of adjustable mounting systems. This technology facilitates dynamic tilt regulation, ensuring optimal alignment of photovoltaic panels with incoming sunlight, thereby improving overall energy conversion efficiency. However, the calculation of the ROI and the breakeven point shows it takes 13 years to reach profit-making state with no expenses. This is because of the upfront cost of the system is high to begin with.

5. ACKNOWLEDGE

REFERENCES

1. M. Alsulaiman and N. Mohammadi, "Optimal Pitch Distance and Tilt Angle of PV Power Plant for Different Climate," 2020.
2. S. I. Mustapa, L. Y. Peng, and A. H. Hashim, "Issues and challenges of renewable energy development: A Malaysian experience," in Proceedings of the International Conference on Energy and Sustainable Development: Issues and Strategies, ESD 2010, IEEE Computer Society, 2010.
3. F. Chien, L. Huang, and W. Zhao, "The influence of sustainable energy demands on energy efficiency: Evidence from China," *Journal of Innovation and Knowledge*, vol. 8, no. 1, Jan. 2023.
4. W. S. W. Abdullah, M. Osman, M. Z. A. A. Kadir, and R. Verayiah, "The potential and status of renewable

energy development in Malaysia," Energies (Basel), vol. 12, no. 12, 2019.

5. F. Reil, I. Baumann, J. Althaus, and S. Gebhard, "Evaluation of current standards and practices for the simulation of wind-blown sands and their applicability as accelerated ageing tests for PV modules," in Conference Record of the IEEE Photovoltaic Specialists Conference, Institute of Electrical and Electronics Engineers Inc., 2013, pp. 1537–1541.

6. "MT G2 Series User Manual Grid-Tied PV Inverter." [Online]. Available: www.goodwe.com

7. T. Bravo Mahachi and A. Johan Rix, "Energy yield analysis and evaluation of solar irradiance models for a utility scale solar PV plant in South Africa," 2018.

8. A. Eric El Obeid, "BEST-PRACTICES GUIDELINES AND LESSONS LEARNT FOR ON-GRID AND PV-DIESEL HYBRID SYSTEMS." [Online]. Available: www.lb.undp.org/DREG

9. "MALAYSIAN STANDARD Installation of grid-connected photovoltaic (PV) system (Second revision) DEPARTMENT OF STANDARDS MALAYSIA DEVELOPMENT OF MALAYSIAN STANDARDS," 2018, [Online]. Available: http://www.sirim.my

10. C. Marketing, "THHN/THWN-2 600V CT RoHS CABLE." [Online]. Available: www.nexans.co

11. F. Saeed and A. Zohaib, "Quantification of Losses in a Photovoltaic System: A Review †," Engineering Proceedings, vol. 11, no. 1, 2021.

12. N. H. Reich, B. Mueller, A. Armbruster, W. G. J. H. M. Van Sark, K. Kiefer, and C. Reise, "Performance ratio revisited: Is PR > 90% realistic?," in Progress in Photovoltaics: Research and Applications, Sep. 2012, pp. 717–726.

13. C. Houston, S. Gooberman-Hill, R. Mathie, A. Kennedy, Y. Li, and P. Baiz, "Case study for the return on investment of internet of things using agent-based modelling and data science," Systems, vol. 5, no. 1, Mar. 2017.

Authors Introduction

Mr. Wei Kang Lai

He is currently pursuing Bachelor of Electrical and Electronics Engineering as final year student in Faculty of Engineering, Technology and Build Environment, UCSI University, Malaysia. His research interests are electrical and solar system design.

Ms. Zhi Ying Yeoh

She is currently pursuing Bachelor of Engineering Electrical and Electronics with Honors as final year student in Faculty of Engineering, Technology and Built Environment, UCSI University, Malaysia. Her research interests are electrical and power electronics.

Ms. Jia Wei Tan

She is currently pursuing Bachelor of Engineering Communication and Electronics with Honors as final year student in Faculty of Engineering, Technology and Built Environment, UCSI University, Malaysia. Her research interests are communication and antenna design.

Dr. Sew Sun Tiang

She is an Assistant Professor in Faculty of Engineering at UCSI University in Malaysia. She received her PhD in Electrical and Electronic Engineering from Universiti Sains Malaysia in 2014. Her research interests are optimization and antenna design.

Dr. Kim Soon Choong

He obtained his B.Eng., M.S., and Ph.D. degrees in electrical and electronic engineering from Universiti Kebangsaan Malaysia in 2012, 2015, and 2019, respectively. He currently serves as an Assistant Professor in the Department of Computer Engineering (Artificial Intelligence) within the Faculty of Engineering Technology and Built Environment at UCSI Kuala Lumpur. His research focuses on areas such as medical signal processing and instrumentation.

Dr. Mohammad Arif Bin Ilyas

He received his PhD in Electrical Engineering from Universiti Tun Hussein Onn Malaysia in 2020. He also earned his bachelor's degree in electronic engineering at the same university. Currently, he is a lecturer in the Faculty of Engineering, Technology, and Built Environment at UCSI University. His main areas of research interest are machine learning, artificial intelligence, etc.

A Design of Dual-Band Coplanar Waveguide (CPW) Printed Antenna for 1.9-3.6GHz Applications

Jia Wei Tan[1], Sew Sun Tiang[1, *], Kim Soon Choong[1], Mohammad Arif Bin Ilyas[1], Mastaneh Mokayef[1], Zhi Ying Yeoh[1], Wei Kang Lai[1], Wei Hong Lim[1]

E-mail: 1001954282@ucsiuniversity.edu.my, tiangss@ucsiuniversity.edu.my, ChongKS@ucsiuniversity.edu.my, mohammadarif@ucsiuniversity.edu.my, mastaneh@ucsiuniversity.edu.my, 1001953809@ucsiuniversity.edu.my, 100953677@ucsiuniversity.edu.my, limwh@ucsiuniversity.edu.my

([1]UCSI University, Malaysia)
* Corresponding author

Abstract

A dual-band coplanar waveguide (CPW) printed antenna for IoT application is proposed, mounted on a low-profile RT/Duroid 5880 substrate with dielectric constant of 2.2, loss tangent of 0.0009 and a standard height of 0.787mm. This design aims to cover the major frequency bands from LTE to Bluetooth/Wi-Fi band/WiMax/Zigbee, Extended IMT and 5G whereby the bandwidth range between 1.7 GHz to 3.6 GHz. The antenna is miniaturized through the CPW technique and has a rectangular size of $60 \times 30 \times 1.187$ mm^3. The design and simulation of the result records a return loss of -25.12dB, peak gain of 4.23dBi, voltage standing wave ratio (VSWR) close to 1, omnidirectional radiation, and current distribution. Radiation efficiency reaches approximately 94%, with a total efficiency of 89.4% between 1.9GHz-3.6GHz.

Keywords: Dual-band CPW, Internet of Things, impedance bandwidth, reflection coefficient, radiation pattern, current distribution

1. Introduction

With increasing world population and the incorporation of Internet of Things (IoT) in multiple computer networks, microelectronics and modern communication, there are higher demands for better, flexible, and advance antenna module to support the embedded systems [1]. Antennas play a primary role in supporting such systems because they enable long-distance communication and efficient use of frequency spectrum that allows maximization of data transmitted over a given frequency band. An example is the dual-band antenna which is designed to support two frequencies that can cover multiband applications and provide sufficient bandwidth for the operating frequencies as well as reduced interference by neighbor devices.

In recent years, coplanar waveguide (CPW) has garnered immense attention across the globe especially in modern wireless communication systems due to their simple integration with microwave integrated circuits, low radiation loss and reduced dispersion as compared to microstrip antenna [2]. This type of waveguide has an additional third conductor centered in the slot region and specifically useful in developing active circuitry. The presence of the additional strip can support even or odd quasi-TEM mode at low frequencies as well as TE mode at high frequencies [3]. The modification of the dimensions of the signal strip or ground plane such as thickness, length and width play a vital role in determining the characteristics of a CPW-fed antenna [2].

There are a few main challenges in designing a dual-band CPW antenna which include miniaturization, large gain, high bandwidth coverage, low return loss, impedance matching, power, and radiation efficiency. Some of the techniques introduced to satisfy the abovementioned criteria includes the usage of metamaterials and complementary split ring resonator (CSRR) to increase gain or implementation of a magneto electric (ME) to achieve wider impedance bandwidth, higher gain, and compatible radiation patterns on both electric and magnetic plane [3].

Hence, in this paper, the design of this waveguide is done by utilizing the CST software which involves careful selection of substrate properties, structure of CPW, and feed network. The simulation result allows optimization and evaluation of antenna performance over a desired frequency range. The proposed antenna presents a dual-band coplanar waveguide using copper

Jia Wei Tan, Sew Sun Tiang, Kim Soon Choong, Mohammad Arif bin Ilyas, Mastaneh Mokayef, Zhi Ying Yeoh, Wei Kang Lai, Wei Hong Lim

conductor and RT/Duroid 5880 as substrate material. The width of the microstrip and gap between the sides of the ground is adjusted to obtained impedance matching of 50 Ohm. The main consideration of the paper is to develop a design that operates on all major frequency bands from LTE (1900MHz) to Extended LTE (2100MHz), Bluetooth/Wi-Fi band/WiMAX/Zigbee (2.4GHz) and 5G (3.4-3.6GHz). Table 1 shows the summary of comparison between existing CPW-fed antennas and proposed work.

Table 1. Summary Of Comparison Between Existing CPW-Fed Antennas and Proposed Work.

References	Feeding Technique	Operating Frequency (GHz)	Gain (dBi)	Dimension (mm³)	Substrate Type
[2]	CPW	2.54-2.69/ 5.68-6.05	5.30 3.44	18.2 × 20 × 1.6	FR-4
[3]	CPW	2.9-21.0	-	23.0 ×10 ×0.8	FR-4
[4]	CPW	1.10-2.70/ 3.15-3.65	8.90	35.0 × 25 × 1.6	FR4
[5]	CPW	3.24-8.29/ 9.12-11.25	4.33 4.90	24.90 × 20 × 1.6	FR-4
[6]	CPW	1.28-4.50	3.5	73 × 65 × 1.52	Rogers RO4232
[7]	CPW	5.15-7.29	2.25	20 × 8.7 × 0.4	FR-4
[8]	CPW	2.45/ 4.0-6.0	0.38 0.765	30 × 30 × 2	Glass
[9]	CPW	2.19-2.51	> 4.2	46.5 × 29.0 × 0.76	Rogers RT6002
[10]	CPW	3.04-10.70/ 15.18-18.0	3.94	47 × 25 × 0.135	PET
Proposed work	CPW	1.90-3.60	4.23	60 × 30 × 1.187	RT/Duroid 5880

2. Related Works

The configuration of the proposed antenna shown in Fig. 1 is inspired by the basic design from S. Das, H. Islam, T. and Bose, N. Gupta [7]. The fundamental structure of the antenna consists of two layers which are antenna layer on top and dielectric substrate at the bottom. Our design assumes no ground plate is placed underneath the dielectric which means it is of free space or air.

a. Front View b. Bottom View
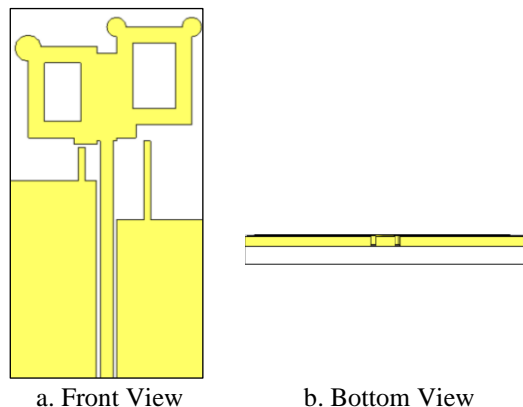
Fig 1. Structure of proposed antenna design.

The material chosen for the antenna patch is copper (annealed) because this form of copper is much softer and malleable which makes it easier to shape during the fabrication stage. On the other hand, the dielectric substance used is the RT/duroid 5880 with dielectric constant of 2.2 and tangent loss of 0.0009. Some key features according to the datasheet of this substrate material include low moisture absorption, resistance towards chemicals, isotropic and uniform chemical properties over frequency. It is mostly found in microstrip and stripline circuits, as well as point-to-point digital radio antennas. The thickness of the dual-band CPW is 1.187mm with copper (annealed) 0.4mm and substrate as 0.787mm respectively.

The detailed geometry of the proposed antenna is shown in Fig. 2. Some of the parameters such as length and width of the main rectangular patch were taken from the values by S. Das, H. Islam, T. and Bose, N. Gupta [7] and is progressively modified to fit the requirements of our antenna. The characteristic impedance of the feed line can be calculated with a certain formula, however, in this paper, we considered utilizing the coplanar waveguide calculator found online which requires value including substrate thickness, width of antenna strip, ground plane spacing, and dieletric substrate [11]. We ran a couple of trials and errors to achieve matching impedance of 50 Ohm and the result was to adjust

antenna width to 2mm while spacing between the edge of feed element and ground plane is 0.1933mm. The calculator equations are based on the 'Coplanar Waveguide Circuits, Components, and Systems' textbook written by R. N. Simons [12].

The equations applied in the online calculator are just an approximation of the CPW impedance as it does not consider the thickness of antenna strips, hence may not display the most accurate results. When creating critical design, the usage of 3D electro-magnetic analysis of

CPW needs to be considered. The gap between the ground plates and antenna feed line is calculated to be 0.193mm which is less than half of the substrate thickness to control leakage of electromagnetic energy in free space. The antenna design is gradually modified to fulfil the required frequency range and the evolution of the proposed antenna design is shown in Fig. 3. The detailed antenna parameters of the final design are depicted in Table 2.
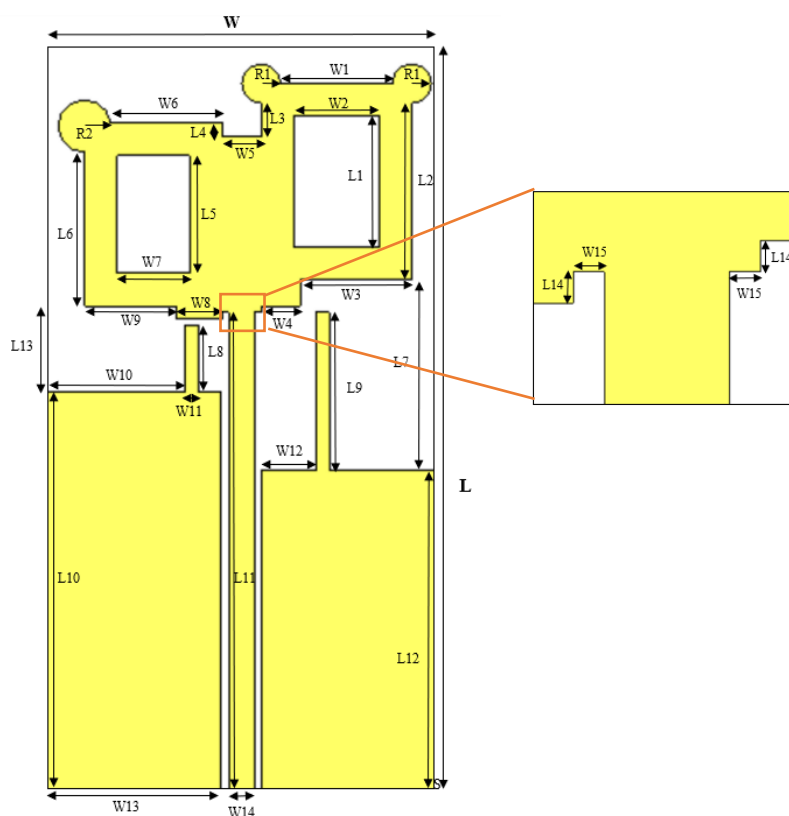
Fig 2. Geometry of proposed antenna.

Table 2. Proposed CPW Antenna Design Parameters Value.

Parameters	L	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10
Value(mm)	60	10	13.5	2.5	1.0	9.0	11.75	13.5	5.0	12	33.0
Parameters	L11	L12	L13	L14	R1	R2	W	W1	W2	W3	W4
Value(mm)	39.0	27.0	0.5	6.5	2.0	1.5	40	8.5	6.5	8.5	3.0
Parameters	W5	W6	W7	W8	W9	W10	W11	W12	W13	W14	W15
Value(mm)	3.0	8.52	5.5	3.5	7.0	10.7	1.0	4.11	13.41	2.0	0.5

Jia Wei Tan, Sew Sun Tiang, Kim Soon Choong, Mohammad Arif bin Ilyas, Mastaneh Mokayef, Zhi Ying Yeoh, Wei Kang Lai, Wei Hong Lim

| a. Antenna 1 | b. Antenna 2 | c. Antenna 3 | d. Antenna 4 |
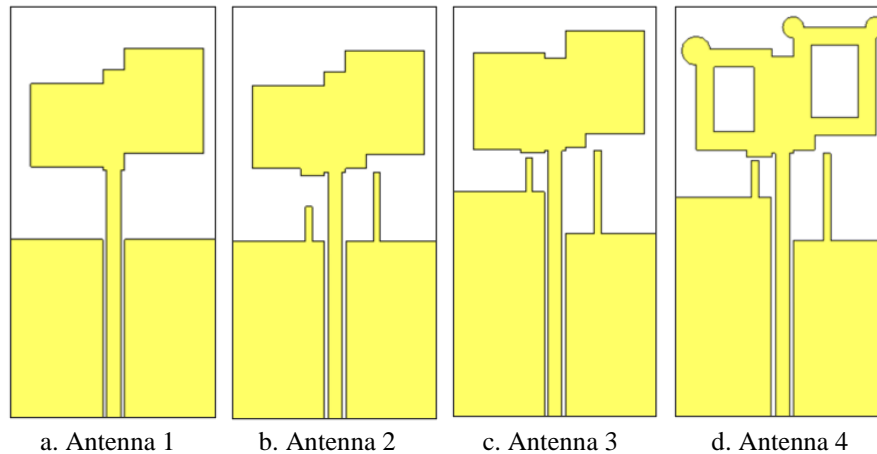
Fig 3. Evolution process of antenna.

3. Result and Discussion

The different layout of CPW antennas were designed and simulated to determine the antenna performance in different aspects including S11-parameter (return loss), reflection coefficients, gain, radiation pattern, current distribution, , and voltage standing wave ratio (VSWR).

The S11 parameter graph shows the return loss of the antenna design which is a measure of the reflection coefficient. The simulated result for return loss from the first to the final proposed antenna is displayed in Fig. 4. The first antenna result (red line) shows a narrow impedance bandwidth less than 30% (2.0-2.6GHz) for S11 < -10dB threshold. Besides, the return loss value is fairly small, which is -13.88dB and this may cause more power reflected back to the system which is not desirable.

This leads to our second antenna (green line) which is designed to increase the return loss as can be seen in Fig. 4. By adding two rectangular patches with 1mm width on each side of the ground plate as shown in Fig. 3 (b), the return loss improved to -21.47dB.

Moving on, the goal of antenna 3 (blue line) is to increase the impedance bandwidth to fulfill the criteria for IoT applications which covers all major frequency bands between 1.7GHz to 3.6GHz. By changing the length between the newly added rectangular strips with the ground plate, the bandwidth coverage improved from the initial antenna design which is 2.0GHz-2.6GHz to 1.97GHz-3.58GHz with more than 80% of the desired spectrum. This design also produces two resonant frequencies at 2.215GHz and 3.28GHz.

Lastly, the addition of circular slots on the corner of the rectangular patches further decreases the return loss to -25.12dB (orange line) and improved the impedance bandwidth to 1.88GHz-3.62GHz. However, one of the major frequency spectrums failed to be covered by this design, which is the LTE with frequency of 1.7GHz.

More detailed parametric studies need to be done such as modifying the major length and width of antenna and dieletric substrate, positions of rectangular strips, separation distance between feed line and ground plates to achieve wider frequency spectrum. Additional tuning and optimization are also needed to attain the appropriate frequency response.
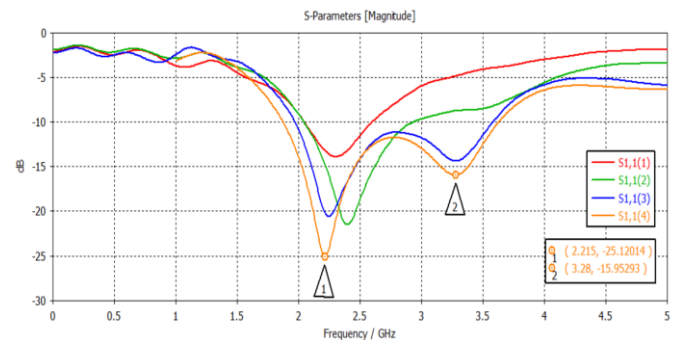
Fig 4. Simulation result of S11 parameter (return loss) for all four antennas.

The gain for final proposed antenna is shown in Fig 5 with the peak gain captured at 4.23dBi when the operating frequency is around 3.6GHz. At resonant frequencies of 2.215GHz and 3.28GHz, the gain observed is 2.59dBi and 3.92dBi respectively.
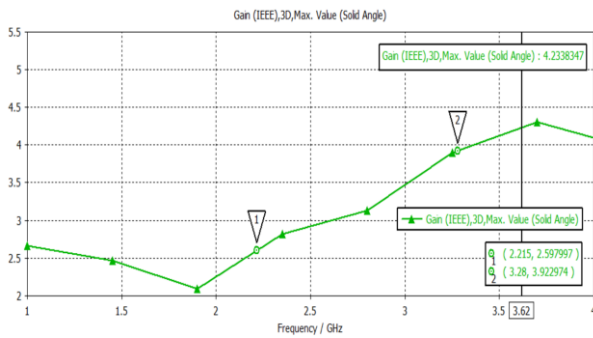
Fig 5. Simulation result of gain for final proposed antenna.

The radiation pattern of the final proposed antenna at 2.215GHz and 3.28GHz frequencies viewed in 3D is shown in Fig. 6 (a) and (b). Based on the results, the antenna design is omnidirectional, or commonly referred to as 'omni' whereby its shape is similar to a bagel. This type of antenna radiates and receives signal in all directions equally. Along its axis, this antenna evenly emits electromagnetic radiation in all directions. An omnidirectional antenna has the benefit of being able to send and receive signals from any direction without the requirement for exact targeting. The directivity recorded at 2.2GHz and 3.28GHz is 2.665dBi and 3.28dBi respectively. The radiation pattern in E-field and H-field can also be viewed in 2D for operating frequency at 2.215GHz and 3.28GHz as shown in Fig. 6 (c) and (d).

a. 3D radiation pattern at 2.215GHz.

b. 3D radiation pattern at 3.28GHz.
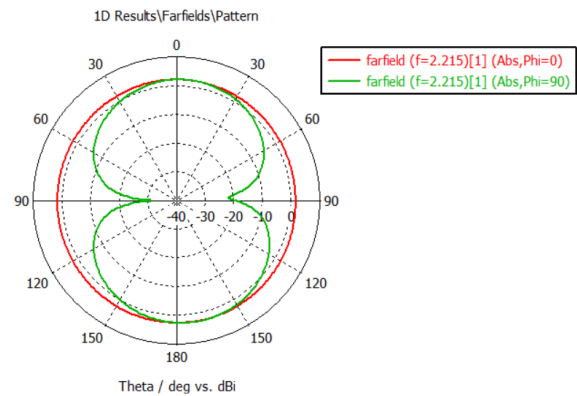
c. Far field radiation pattern at 2.215GHz.
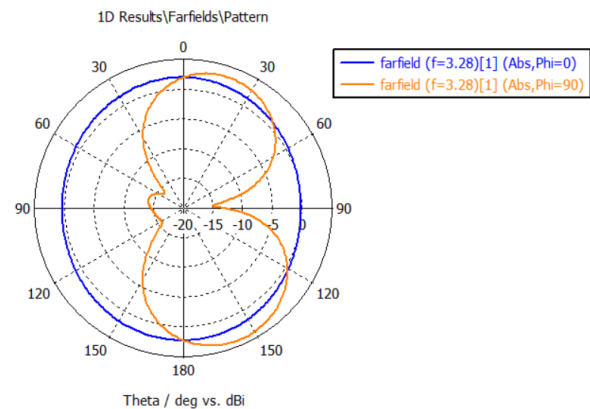
d. Far field radiation pattern at 3.28GHz.

Fig 6. Radiation pattern of final proposed antenna.
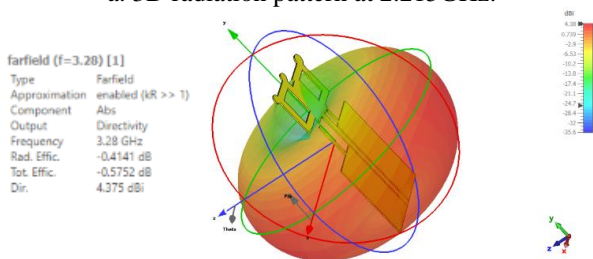
The surface current distribution is measured at both resonant frequencies of 2.215GHz and 3.28GHz as shown in Fig 7 (a) and (b). In both figures, the current density is most evident at the antenna strip denoted by red and orange color map which has higher potential for losses and prone to heating easily due to high dissipation of electromagnetic fields and energy. Some methods to reduce these losses include having proper grounding and optimizing the feed network as well as geometry design. The rest of the areas that are mostly drawn in blue or green area experience low current density whereby it indicates that the performance of the antenna still has room for improvement.

Jia Wei Tan, Sew Sun Tiang, Kim Soon Choong, Mohammad Arif bin Ilyas, Mastaneh Mokayef, Zhi Ying Yeoh, Wei Kang Lai, Wei Hong Lim
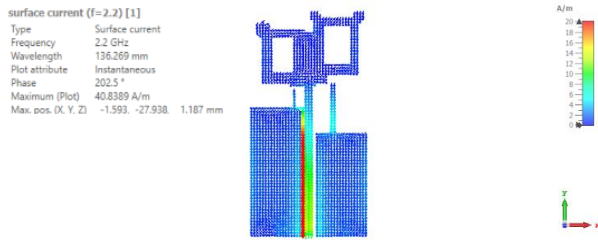
a. Current distribution at 2.215GHz.
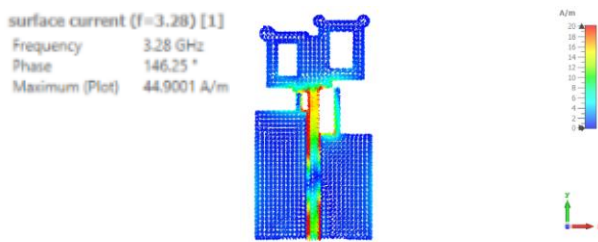
b. Current distribution at 3.28GHz.

Fig 7. Current distribution on final proposed antenna.
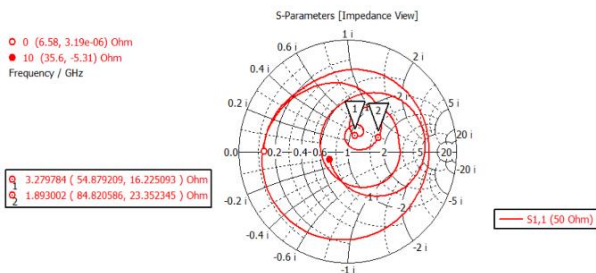
Fig 8. Smith chart of the final proposed antenna.

Smith chart is shown in Fig. 8. The VSWR value plotted at 2.215GHz is 1.11 and at 3.28GHz is 1.38 as shown in Fig. 9 whereby both values can be considered close to 1. This value of VSWR indicates a well-matched antenna with the transmission line and also reduces power loss due to impedance mismatch.
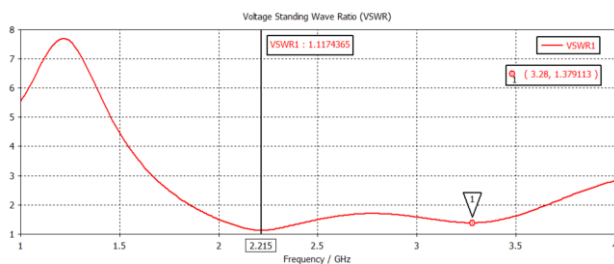
Fig 9. VSWR graph of the final proposed antenna.

Based on Fig. 10, efficiencies are recorded in terms of magnitude (dB). The highest radiation efficiency and total efficiency converted into percentage are 98% and 92.4% respectively at 2.35GHz. The radiation and total efficiency for resonant frequencies of 2.2GHz and 3.28GHz are mostly above 90% and values are tabulated in Table 3.

Table 3. Radiation efficiency and total efficiency of different resonant frequencies.

Resonant Frequency (GHz)	Radiation efficiency (%)	Total efficiency (%)
2.2	92.2%	86.6%
3.28	94.6%	90%
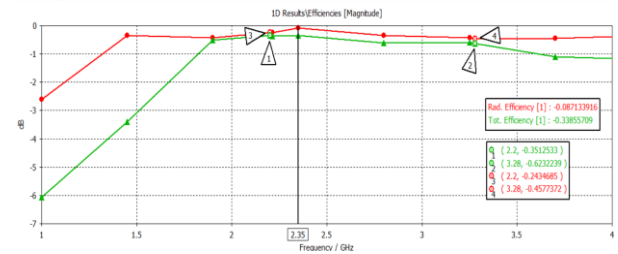
Fig 10. Radiation efficiency and total efficiency of dual-band CPW.

4. Conclusion

In this proposed paper, the rectangular dual-band CPW antenna structure with RT/duroid 5880 as substrate that has standard height of 0.787mm and dieletric constant of 2.2 is designed and simulated using CST software. The design can be classified for the usage on the applications of Internet of Things as it can satisfy most of the major frequency bands from LTE to Bluetooth/Wi-Fi band/WiMax/Zigbee, Extended IMT and 5G. However, it is also worth mentioning that the proposed design fails to cover a lower bandwidth spectrum of LTE which is 1.7GHz and this can be due to various factors such as dimensions, size, and layout of the antenna design as well as substrate material. The final antenna design has a wide impedance bandwidth of 1.9GHz-3.6GHz, improve in overall antenna gain with highest peak value at 4.23dBi in the two working bands and VSWR approximately close to 1. The final antenna design demonstrates monopole radiation patterns (omnidirectional) characteristics and high radiation efficiency up to 98% between the stated impedance bandwidths. The proposed antenna also exhibits moderate return loss of -29.54dB with an acceptable size of 1800m^2 and total efficiency of approximately 92.4% which makes it a comparable candidate for portable handheld communication devices.

Acknowledgements

References

[1] D.M Pozar, Microwave Engineering, 4th Edition, United States of America: John Wiley & Sons Inc, 2012.

[2] R. Vivek, S. Sreenath, P.V. Vinesh, and K. Vasudevan, "Coplanar waveguide (CPW)-fed compact dual band antenna for 2.5/5.7GHz applications," Progress In Electromagnetics Research M, vol. 74, pp. 51-59, 2018.

[3] P. Ranjan, A. Maurya, H. Gupta, S. Yadav, and A. Sharma, "Ultra-wideband CPW fed band-notched monopole antenna optimization using machine learning," Progress in Electromagnetics Research M, vol. 108, pp, 27-38, 2022.

[4] Q. Awais, H.T. Chattha, M. Jamil, Y. Jin, F.A. Tahir, and M.U. Rehman, "Novel dual ultrawideband CPW-fed printed antenna for interent of things (IoT) applications", Wireless Communications and Mobile Computing, vol. 2018.

[5] M. M. Alam, R. Azim, N. M. Sobahi, A. I. and Khan, and M. T. Islam, "A dual-band CPW-fed miniature planar antenna for S-, C-, WiMAX, WLAN, UWBm and X-band applications," Scientific Reports, vol. 12, 2022.

[6] J. Kulkarni, and C.Y.D. Sim, "Wideband CPW-fed oval-shaped monopole antenna for Wi-Fi 5 and Wi-Fi 6 applications", Progress in Electromagnetic Research C, vol. 107, pp. 173-182, 2021.

[7] S. Das, H. Islam, T. and Bose, N. Gupta, "Ultra wide band CPW-fed circularly polarized microstrip antenna for wearable applications", Wireless Personal Communications, vol. 108, pp. 87-106, 2019.

[8] N. Morales-Centla, R.Torrealba-Melendez, E. I. Tamariz-Flores, M. Lopez-Lopez, C.A Arriaga-Arriaga, J.M. Munoz-Pacheco, and V. R. Gonzalez-Diaz, "Dual-band CPW graphene antenna for smart cities and IoT applications," MDPI, 2022.

[9] S. Ahmad, N. Cherif, S. Naseer, U. Ijaz, Y.S. Faouri, A. Ghaffar, and M. Hussein, "A wideband circularly polarized CPW-fed substrate integrated waveguide based antenna array for ISM band applications," Elsevier Ltd, 2022.

[10] S.G.Kirtania,B.A. Younes, A. R. Hossain, T. Karacolak, and P.K. Sekhar, "CPW-fed flexible ultra-wideband antenna for IoT applications," Micromachines, vol. 12, pp. 453, 2021.

[11] Coplanar Waveguide Calculator. everythingRF. Microwave101.com. https://www.microwaves101.com/calculators/864-coplanar-waveguide-calculator

[12] R. N. Simons, Coplanar Waveguide Circuits, Components, and Systems. John Wiley & Sons Inc, pp. 15-21, 2001.

Authors Introduction

Ms. Jia Wei Tan

She is currently pursuing Bachelor of Engineering Communication and Electronics with Honors as final year student in Faculty of Engineering, Technology and Built Environment, UCSI University, Malaysia. Her research interests are communication and antenna design.

Dr. Sew Sun Tiang

She is an Assistant Professor in Faculty of Engineering at UCSI University in Malaysia. She received her PhD in Electrical and Electronic Engineering from Universiti Sains Malaysia in 2014. Her research interests are optimization and antenna design.

Dr. Kim Soon Choong

He obtained his B.Eng., M.S., and Ph.D. degrees in electrical and electronic engineering from Universiti Kebangsaan Malaysia in 2012, 2015, and 2019, respectively. He currently serves as an Assistant Professor in the Department of Computer Engineering (Artificial Intelligence) within the Faculty of Engineering Technology and Built Environment at UCSI Kuala Lumpur. His research focuses on areas such as medical signal processing and instrumentation.

Dr. Mohammad Arif Bin Ilyas

He received his PhD in Electrical Engineering from Universiti Tun Hussein Onn Malaysia in 2020. He also earned his bachelor's degree in electronic engineering at the same university. Currently, he is a lecturer in the Faculty of Engineering, Technology, and Built Environment at UCSI University. His main areas of research interest are machine learning, artificial intelligence, etc.

Dr. Mastaneh Mokayef

She has received her PhD from Wireless Communication Centre Faculty of Electrical Engineering in University Technology Malaysia (UTM) in 2014. She has also obtained her master's degree from the faculty of engineering in 2009 from the University Technology Malaysia. Her research interests include Wireless communications, spectrum sharing method, spectrum management, cellular communication systems and Antenna design.

Ms. Zhi Ying Yeoh

She is currently pursuing Bachelor of Engineering Electrical and Electronics with Honours as final year student in Faculty of Engineering, Technology and Built Environment, UCSI University, Malaysia. Her research interests are power electronics and simulation

Mr. Wei Kang Lai

He is currently pursuing Bachelor of Electrical and Electronics Engineering as final year student in Faculty of Engineering, Technology and Build Environment, UCSI University, Malaysia. His research interests are electrical and solar system design.

Dr. Wei Hong Lim

He is an Assistant Professor in Faculty of Engineering at UCSI University in Malaysia. He received his PhD in Computational Intelligence from Universiti Sains Malaysia in 2014. His research interests are optimization and artificial intelligence.

Optimized Convolutional Neural Network Towards Effective Wafer Defects Classification

Koon Hian Ang[1], Koon Meng Ang[1], Chin Hong Wong[2,3], Abhishek Sharma[4], Chun Kit Ang[1], Kim Soon Chong[1], Sew Sun Tiang[1]*, Wei Hong Lim[1]*

[1]*Faculty of Engineering, Technology and Built Environment, UCSI University, Kuala Lumpur 56000, Malaysia*
[2]*Maynooth International Engineering College, Maynooth University, Maynooth, Co Kildare, Ireland*
[3]*Maynooth International Engineering College, Fuzhou University, Fujian, 350116, China*
[4]*Department of Computer Science and Engineering, Graphic Era Deemed to be University, Dehradun 248002, India*
Email: 1001850063@ucsiuniversity.edu.my, 1001436889@ucsiuniversity.edu.my, chinhong.wong@mu.ie,
abhishek15491@gmail.com, angck@ucsiuniversity.edu.my, ChongKS@ucsiuniversity.edu.my,
tiangss@ucsiuniversity.edu.my, limwh@ucsiuniversity.edu.my

Abstract

Semiconductor defect inspection is crucial for yield improvement but is hindered by manual inspection's subjectivity and error. This paper employs Convolutional Neural Networks (CNNs) for automated wafer defect classification, addressing the challenges of time-intensive training and complex hyperparameter tuning. We propose the Arithmetic Optimization Algorithm (AOA) to efficiently optimize CNN hyperparameters like momentum, initial learning rate, maximum epochs, and L2 regularization. Our method reduces the trial-and-error in hyperparameter tuning. Using the AOA-optimized ResNet-18 model, our simulations show superior performance in defect classification compared to the unoptimized model, demonstrating its effectiveness and practical potential.

Keywords: Arithmetic optimization algorithm, Convolutional neural networks, Hyperparameter optimization, Wafer defect classification

1. Introduction

The advent of the fourth industrial revolution (IR4.0) has escalated the need for semiconductor industries to produce increasingly complex integrated circuit chips. This complexity arises from the need to pattern and etch more components on semiconductor wafers, catering to diverse chip specifications like lifespan, size, memory storage, and access speed. Consequently, this escalation in production demands heightens the likelihood of process-induced defects on wafer surfaces, adversely affecting the manufacturing yield. A key step in mitigating this yield reduction involves the identification and classification of wafer defect patterns, which often correlate with various manufacturing stages, including contamination, robot handoff, and flow leakages. Accurate identification of these defects enables engineers to pinpoint and rectify the underlying issues, thereby enhancing chip production yield [1].

Most semiconductor industries currently depend on manual visual inspection for detecting wafer defects, a method plagued by subjectivity and a high risk of erroneous classifications due to long-term fatigue. To address these limitations, there is a growing need for an automated machine vision system, integrated with an optimized deep learning model, for reliable wafer defect classification. Convolutional Neural Networks (CNNs) are a widely used deep learning technique [2], [3], [4], [5], [6], [7]. particularly effective in wafer defect

classification [8], [9], owing to their capability to learn the nonlinearity between inputs and outputs by automatically extracting relevant information from raw data. Pre-trained CNN architectures like GoogleNet [10]. AlexNet [11], VGG [12], and ResNet [13] have been successfully adapted using transfer learning to tackle new tasks, benefiting from reduced data and training time requirements. The efficacy of these pre-trained networks in addressing new tasks through transfer learning process is heavily influenced by the hyperparameter settings employed during their training phase.

In conventional practice, the fine-tuning of CNN hyperparameters relies on a labor-intensive and time-consuming trial-and-error approach. To streamline this process, metaheuristic search algorithms (MSAs), drawing inspiration from various natural phenomena [14]—including swarm intelligence, natural evolution, physics, mathematical principles, and human activities—have been developed. These algorithms are effective in addressing complex problems [15], [16], [17], [18], [19], [20], [21], [22] and are particularly adept at hyperparameter tuning in CNNs [23], thanks to their robust global search capabilities. One such algorithm is the Arithmetic Optimization Algorithm (AOA) [24], which derives its strategy from the distribution behaviors of basic arithmetic operations such as addition, subtraction, multiplication, and division.

In this study, we employ a pre-trained network, ResNet-18, retrained on new datasets using transfer learning for

the purpose of wafer defect classification. The AOA is utilized to optimize four critical hyperparameters of the CNN: momentum, initial learning rate, maximum epochs, and L2 regularization. The efficacy of this optimized CNN model in classifying wafer defects is then assessed and benchmarked against the performance of a corresponding unoptimized CNN model.

2. Related Works
2.1. *Conventional CNN and ResNet-18*

A typical CNN architecture comprises three fundamental components: convolutional layers (Conv) for feature extraction, pooling (Pool) layers for reducing the size of feature maps, and fully connected (FC) layers for classification. Table 1 illustrates an example of a standard CNN model.

Table 1. Architecture of a typical CNN.

Layer	Layer Type	#Feature Maps	Feature Map Size	Filter Size
1	Input	1	14×14	
2	Conv 1	6	7 × 7	5×5
3	Pool 1	6	4×4	2 ×2
4	Conv 2	16	4×4	5 × 5
5	Pool 2	16	2 × 2	2 × 2
6	FC 1	1	120	
7	FC 2	1	84	

He et al. [13] introduced ResNet models, known for their deep architectures, which have exhibited remarkable convergence and high accuracy. In 2015, ResNet models achieved first place in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) and the Common Objects in Context (COCO) classification challenge. ResNet is characterized by multiple stacked residual units and comes in various configurations, including 18, 34, 50, 101, 152, and 1202 layers, each with specific operations tailored to its architecture. ResNet-18, a model offering an optimal balance between depth and performance, comprises five convolutional layers, one average pooling layer, and a fully connected layer with SoftMax activation. ResNet-50 includes 49 convolutional layers, culminating in a fully connected layer. Considering the balance between computational efficiency and performance, ResNet-18 is chosen for this study.

2.2. *Arithmetic optimization algorithm (AOA)*

In 2021, Abualigah et al. [24] introduced the AOA, a MSA inspired by the characteristics of the four basic arithmetic operations: addition, subtraction, multiplication, and division. The AOA operates through three primary phases: initialization, exploration, and exploitation, strategically navigating the solution space to address optimization challenges effectively.

In the initialization phase of the AOA, a set of potential solutions is generated, each falling within predefined dimensional boundaries, to address specific optimization problems. Subsequently, the algorithm calculates the Math Optimizer Accelerated (MOA) value, which dictates the search strategy between exploration and exploitation phases. The MOA function is mathematically expressed as:

$$MOA(C_{Iter}) = Min + C_{Iter} \times \left(\frac{Max - Min}{M_{Iter}}\right) \quad (1)$$

where C_{Iter} denotes the current iteration number, M_{Iter} is the maximum number of iterations, and *Min* and *Max* are the lower and upper limits of the accelerated function, respectively. The AOA enters the exploration phase if the MOA value is lower than a randomly generated number between 0 and 1; otherwise, it proceeds to the exploitation phase.

During the exploration phase of the AOA, both Multiplication and Division operators are employed to enhance the search space coverage and introduce a diverse array of solutions. This approach leverages their potential for high dispersion and distributed values. The position of each *i*-th AOA solution in the *d*-th dimension is updated as follows:

$$x_{i,j}(C_{Iter} + 1)$$
$$= \begin{cases} x_j^{best} \div (MOP + \varepsilon) \times \left((UB_j - LB_j) \times \mu + LB_j\right), \\ \qquad\qquad r1 < 0.5 \\ x_j^{best} \times MOP \times \left((UB_j - LB_j) \times \mu + LB_j\right), \\ \qquad\qquad otherwise \end{cases}$$
$$(2)$$

where $x_{i,j}(C_{Iter} + 1)$ denotes the *j*-th dimension of the *i*-th solution in the next iteration ($C_{Iter} + 1$), with *j* ranging from 1 to *D*, *i* ranging from *1* to *I*, and *I* is the population size; C_{Iter} ranges from 1 to M_{Iter}, the latter is the algorithm's maximum iteration number; x_j^{best} is the *j*-th dimension of the best solution; UB_j and LB_j are the upper and lower boundaries in the *j*-th dimension; $r1$ is a random number between 0 and 1; μ adjusts the search range. Define the Math Optimizer Probability (MOP) as:

$$MOP(C_{Iter}) = 1 - \frac{C_{Iter}^{\frac{1}{\alpha}}}{M_{Iter}^{\frac{1}{\alpha}}} \quad (3)$$

where $MOP(C_{Iter})$ represents the function value at iteration C_{Iter}, and α controls the exploitation accuracy.

In the exploitation phase of the AOA, Subtraction and Addition operators are utilized to precisely target the search regions identified by the most promising AOA solution. This step leverages the operator's characteristic of being densely concentrated yet having low dispersion. The position of each *i*-th AOA solution in the *d*-th dimension is updated as follows during this phase:

$$x_{i,j}(C_{Iter} + 1)$$
$$= \begin{cases} x_j^{best} - MOP \times \left((UB_j - LB_j) \times \mu + LB_j\right), \\ \qquad\qquad r2 < 0.5 \\ x_j^{best} + MOP \times \left((UB_j - LB_j) \times \mu + LB_j\right), \\ \qquad\qquad \text{otherwise} \end{cases}$$

$$(4)$$

Here, $r2$ represents a randomly generated number between 0 and 1, following a uniform distribution.

The AOA algorithm iteratively executes these exploration and exploitation processes until predefined stopping criteria are met. The optimal solution generated by AOA at the end of the optimization process is adopted as the best set of hyperparameters for training the ResNet-18 network to classify wafer defects.

3. Proposed Optimized Deep Learning Model
3.1. *Preprocessing of dataset with wafer defects*

The optimized ResNet-18 architecture undergoes training and evaluation for wafer defect classification utilizing the WM-811K dataset [25]. This particular dataset is noted for its significantly imbalanced distribution across various classes. To mitigate potential biases and overfitting stemming from this imbalance, we formed a new dataset using an under-sampling method. This approach involved compiling 30,000 images classified as 'None defect' from the WM-811K dataset, along with all other images categorized as defects. The resultant dataset, balanced in terms of defect and non-defect images, is detailed in Table 2, providing a comprehensive breakdown of the image categories.

Table 2. Number of defect and non-defect wafer images after preprocessing stage

Layer	Defect Type	#Labeled Images
1	Center	4,294
2	Donut	555
3	Edge-Loc	5,189
4	Edge-Ring	9,680
5	Loc	3,593
6	Near-Full	149
7	Random	866
8	Scratch	1,193
9	None	30,000

3.2. *Hyperparameter tuning of ResNet-18 with AOA*

The revised WM-811K dataset, detailed in Table 2, is used to train the chosen deep learning model, ResNet-18, for wafer defect classification via transfer learning. To tailor the pre-trained network for this specific classification task, several modifications are made to both the input datasets and the configuration of ResNet-18. The size of each input image is increased from 48×48×1 to 224×224×3, the padding size of the initial convolution layer is adjusted to align with the new input size, and the output size of the fully connected layer is altered to 9 to reflect the nine defect types identified in this study.

Additionally, to enhance ResNet-18's performance in classifying wafer defects, the AOA is implemented to optimize the network's training hyperparameters during the transfer learning process. The AOA generates solution vectors representing four key decision variables: momentum, initial learning rate, maximum epochs, and L2 regularization. The ranges for these variables are specified in Table 3. The effectiveness of each AOA solution is assessed based on the classification accuracy achieved. Fig. 1 illustrates the AOA-optimized ResNet-18 framework developed for wafer defect classification.

Table 3. Lower and upper bounds of training hyperparameters to be optimized.

Hyperparameter	Lower Limits	Upper Limits
Momentum	0.5	0.9
Initial Learning Rate	0.01	0.1
Maximum Epochs	5	10
L2 Regularization	1×10^{-4}	5×10^{-4}

AOA-Optimized ResNet-18 for Wafer Defect Classification

Input: N, D, UB_j, LB_j

```
01:   Initialize C_Iter = 0;
02:   for i =1 to I do
03:       Randomly generate solution x_i;
04:       C_Iter = C_Iter + 1;
05:   end for
06:   while C_Iter ≤ M_Iter do
07:       Decode the hyperparameters from x_i;
08:       Train the ResNet-18 using transfer learning based on
          hyperparameters decoded from x_i;
09:       Evaluate the accuracy f(x_i) of ResNet-18;
10:       Calculate MOP using Eq. (3);
11:       Calculate MOA using Eq. (1);
12:       for i =1 to I do
13:           for j =1 to D do
14:               if rand > MOA then
15:                   Update x_{i,j}(C_Iter + 1) with Eq. (2);
16:               Else
17:                   Update x_{i,j}(C_Iter + 1) with Eq. (4)
18:               end if
19:           end for
20:           Decode the hyperparameters from x_i;
21            Train the ResNet-18 using transfer learning
              based on hyperparameters decoded from x_i;
22:           Evaluate the accuracy f(x_i) of ResNet-18;
21:           Update x_i, f(x_i), x^{best}, f(x^{best})
22:           C_Iter ← C_Iter + 1;
23:       end for
24:   end while
```

Output: Optimal values of momentum, initial learning rate, maximum epochs and L2 regularization decoded from the best AOA solution represented as x^{best}.

Fig.1 Pseudocode of optimizing the hyperparameters of ResNet-18 using AOA for wafer defect classification.

4. Performance Studies
4.1. *Simulation settings*

The pre-processed WM-811K dataset is randomly partitioned into three subsets for the purposes of training, validation, and testing. Specifically, the dataset allocation

comprises 70% for training, 10% for validation, and 20% for testing. Additionally, all input images are uniformly resized to a dimension of 224×224×3, and the minimum batch size for processing is set as 32.

4.2. *Performance comparisons*

The classification effectiveness of both unoptimized ResNet-18 and AOA-Optimized ResNet-18 in identifying wafer defects is evaluated using several key metrics: recall, accuracy, precision, F1 score, and AUC. Consider *TP*, *TN*, *FP*, and *FN* as the true positive, true negative, false positive, and false negative counts, respectively, generated by the deep learning model during the classification process. Recall, reflecting the model's capability to detect positive samples, is calculated as:

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

Accuracy, indicating the proportion of correctly predicted data, is defined as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

Precision, measuring the model's accuracy in predicting a sample as positive, is expressed as:

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

The F1 score, the harmonic mean of precision and recall, is calculated by:

$$F1\ Score = 2 \times \left(\frac{Precision \times Recall}{Precision + Recall} \right) \quad (8)$$

Lastly, AUC, representing the area under the receiver operating characteristic curve, provides an overall effectiveness measure. The quantitative performance of both the unoptimized and AOA-Optimized ResNet-18 models is detailed in Table 4. Additionally, their qualitative performance is analyzed based on the confusion matrices depicted in Fig. 2.

Table 4. Quantitative performance comparison between optimized and unoptimized ResNet-18

Metrics	Unoptimized ResNet-18	AOA-Optimized ResNet-18
Validation Accuracy	0.8560	**0.8983**
Testing Accuracy	0.8388	**0.8923**
Recall	0.8482	**0.8956**
Precision	0.8751	**0.9059**
F1 Score	0.8523	**0.8972**
AUC	0.9002	**0.9397**

Table 4 illustrates that the AOA-optimized ResNet-18 surpasses the unoptimized model in all assessed metrics. It shows a higher validation accuracy of 0.8983 compared to the unoptimized model's 0.8560, and a testing accuracy of 0.8923 versus 0.8388. Additionally, the recall score of the AOA-optimized model stands at 0.8956, outperforming the unoptimized model's 0.8420. In terms

of precision, the AOA-optimized model achieves a score of 0.9059, exceeding the unoptimized model's 0.8751. The F1 score and AUC for the optimized model are 0.8972 and 0.9397, respectively, both higher than the unoptimized model's scores of 0.8523 and 0.9002. The enhanced validation accuracy of the AOA-optimized ResNet-18 indicates its superior performance in predicting classes of unseen samples during training, while its greater testing accuracy suggests a more accurate prediction of new, unseen data classes. The improved recall and precision scores signify the model's heightened ability to correctly identify more true positive samples and effectively reduce false positives.

(a)

(b)
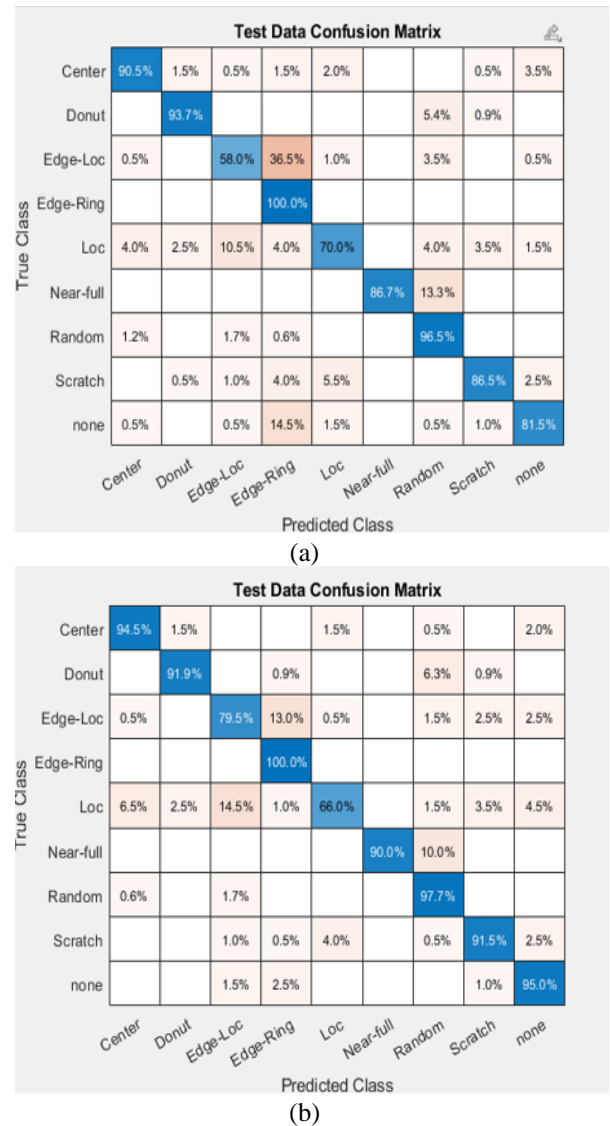
Fig.2 Confusion matrices produced by: (a) Unoptimized ResNet-18 and (b) AOA-Optimized ResNet-18 when classfying wafer defects.

Qualitative results, derived from confusion matrices of both the AOA-optimized and unoptimized ResNet-18 models as depicted in Fig. 2, align with the quantitative findings in Table 4. The AOA-optimized model demonstrates superior accuracy in classifying various

wafer defects, including Center, Edge-Loc, Near-Full, Random, and Scratch, compared to the unoptimized version. Notably, the unoptimized model frequently misclassifies Edge-Loc defects as Edge-Ring, leading to significant errors, and incorrectly labels 14.5% of 'No defect' cases as Edge-Ring defects. Both the quantitative and qualitative analyses suggest that optimizing ResNet-18's hyperparameters through AOA effectively corrects the misclassification issues observed in the unoptimized model, resulting in a marked enhancement in the model's ability to classify wafer defects accurately and efficiently.

5. Conclusion

This study is designed to demonstrate the effectiveness of optimizing hyperparameters in deep learning models for wafer defect classification. In our approach, the Arithmetic Optimization Algorithm (AOA) is employed to fine-tune four key hyperparameters—momentum, initial learning rate, maximum epoch, and L2 regularization rate—of the ResNet-18 model. This model, initially pre-trained on ImageNet, was further refined on a wafer defect dataset via transfer learning. Our extensive simulation studies indicate that the AOA-optimized ResNet-18 model surpasses the unoptimized version in recall, accuracy, precision, F1 score, and AUC, showcasing enhanced capabilities in accurately identifying and classifying wafer defects.

The findings reveal the significance of hyperparameter optimization in developing deep learning models for wafer defect classification. Future research could investigate the applicability of AOA in fine-tuning other deep learning models within this domain. Additionally, exploring AOA for identifying optimal deep learning network architectures offers promising avenues for robust wafer defect classification. Incorporating more features and diverse data sources could further refine the proposed method's efficacy. Finally, subsequent studies should assess the practicality of implementing these optimized deep learning models in industrial settings for real-world wafer defect classification applications.

Acknowledgements

References

1. J. C. Chien, M. T. Wu and J. D. Lee, Inspection and classification of semiconductor wafer surface defects using CNN deep learning networks, Applied Sciences 10(15), 2020, pp. 5340.
2. B. Jdid, W. H. Lim, I. Dayoub, K. Hassan and M. R. B. M. Juhari, Robust automatic modulation recognition through joint contribution of hand-crafted and contextual features, IEEE Access 9, 2021, pp. 104530-104546.
3. M. Alrifaey et al., Hybrid deep learning model for fault detection and classification of grid-connected photovoltaic system, IEEE Access 10, 2022, pp. 13852-13869.
4. T. Berghout, L. H. Mouss, T. Bentrcia and M. Benbouzid, A semi-supervised deep transfer learning approach for rolling-element bearing remaining useful life prediction, IEEE Transactions on Energy Conversion 37(2), 2022, pp. 1200-1210.
5. O. Friha, M. A. Ferrag, M. Benbouzid, T. Berghout, B. Kantarci and K.-K. R. Choo, 2DF-IDS: Decentralized and differentially private federated learning-based intrusion detection system for industrial IoT, Computers & Security 127, 2023, pp. 103097.
6. T. Berghout, M. Benbouzid, Y. Amirat and G. Yao, Lithium-ion battery state of health prediction with a robust collaborative augmented hidden layer feedforward meural network approach, IEEE Transactions on Transportation Electrification 9(3), 2023, pp. 4492-4502.
7. L. S. Chow, G. S. Tang, M. I. Solihin, N. M. Gowdh, N. Ramli and K. Rahmat, Quantitative and qualitative analysis of 18 deep convolutional neural network (CNN) models with transfer learning to diagnose COVID-19 on chest X-ray (CXR) images, SN Computer Science 4(2), 2023, pp. 141.
8. U. Batool, M. I. Shapiai, M. Tahir, Z. H. Ismail, N. J. Zakaria and A. Elfakharany, A systematic review of deep learning for silicon wafer defect recognition, IEEE Access 9, 2021, pp. 116572-116593, 2021.
9. S. Chen, Y. Zhang, M. Yi, Y. Shang and P. Yang, AI classification of wafer map defect patterns by using dual-channel convolutional neural network, Engineering Failure Analysis 130, 2021, pp. 105756.
10. C. Szegedy et al., Going Deeper with Convolutions, 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 2015, pp. 1-9.
11. A. Krizhevsky, I. Sutskever and G. Hinton, ImageNet classification with deep convolutional neural networks, Neural Information Processing Systems 25, 2012.
12. K. Simonyan and A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv 1409.1556, 2014.
13. K. He, X. Zhang, S. Ren and J. Sun, Deep Residual Learning for Image Recognition, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NY, USA, 2016, pp. 770-778.
14. M. F. Ahmad, N. A. M. Isa, W. H. Lim and K. M. Ang, Differential evolution: A recent review based on state-of-the-art works, Alexandria Engineering Journal 61(5), 2022, pp. 3831-3872.
15. M. Shaari et al., Supervised evolutionary programming based technique for multi-DG installation in distribution system, IAES International Journal of Artificial Intelligence 9, 2020, pp. 11.
16. K. M. Ang et al., Modified teaching-learning-based optimization and applications in multi-response machining processes, Computers & Industrial Engineering 174, 2022, pp. 108719.
17. A. Sharma, A. Sharma, V. Jately, M. Averbukh, S. Rajput, and B. Azzopardi, A novel TSA-PSO based hybrid slgorithm for GMPP tracking under partial shading conditions, Energies, 15, 2022, pp. 3164.
18. A. Singh, A. Sharma, S. Rajput, A. K. Mondal, A. Bose and M. Ram, Parameter extraction of solar module using the sooty tern optimization algorithm, Electronics 11, 2022, pp. 564.
19. C. Hassan, V. Durai, S. Sapuan, N. A.A and M. Z. Mohamed Yusoff, Mechanical and crash performance of unidirectional oil palm empty fruit bunch fibre-reinforced polypropylene composite, Bioresources 13, 2018, pp. 8310-8328.

20. E. Natarajan, C. Hassan, C. K. Ang, M. S. Santhosh, S. Ramesh and R. Sasikumar, Modeling of multiwall carbon nanotubes reinforced natrual rubber for soft robotic applications - A comprehessive presentation 46(9), 2021, pp. 3251-3258.

21. A. Machmudah et al. Cyclic path planning of hyper-redundant manipulator using whale optimization algorithm, International Journal of Advanced Computer Science and Applications 12(8), 2021, pp. 677-686.

22. K. Palanikumar, J. Nithyanandam, E. Natarajan, W. H. Lim and S. S. Tiang, Mitigated cutting force and surface roughness in titanium alloy-multiple effective guided chaotic multi objective teaching learning based optimization, Alexandria Engineering Journal 64, 2023, pp. 877-905.

23. K. M. Ang et al., Optimal design of convolutional neural network architectures using teaching and learning-based optimization for image classification, Symmetry 14(11), 2022, pp. 2323.

24. L. Abualigah, A. Diabat, S. Mirjalili, M. Abd Elaziz and A. H. Gandomi, The Arithmetic Optimization Algorithm, Computer Methods in Applied Mechanics and Engineering 376, 2021, pp. 113609.

25. J.-S. R. Jang. MIR-WM811K wafer map, http://mirlab.org/dataSet/public/

Authors Introduction

Mr. Koon Hian Ang

He received the Bachelor of Mechatronics Engineering with Honours in Faculty of Engineering, Technology and Built Environment, UCSI University, Malaysia, in 2023. His research interests are machine learning, deep learning, and optimization algorithm.

Dr. Koon Meng Ang

He received the B.Eng. degree in Mechatronic Engineering with Honours and PhD in Engineering from UCSI University, Malaysia, in 2019 and 2023, respectively. His research interests are swarm intelligence, machine learning and deep learning.

Dr. Chin Hong Wong

He is a Lecturer in Maynooth International Engineering College at Fuzhou University in China. He received his PhD in Electrical and Electronic Engineering from Universiti Sains Malaysia in 2017. His research interests are Energy harvesting and control system.

Dr. Abhishek Sharma

He is a Research Assistant Professor at Graphic Era Deemed to be University in India. He received his PhD from University of Petroluem & Energy Studies in 2022. His research interests are artificial intellgence and power electronics.

Dr. Chun Kit Ang

He is the Dean and Associate Professor in Faculty of Engineering at UCSI University in Malaysia. He received his PhD in Mechanical and Manufacturing Engineering from Universiti Putra Malaysia in 2014. His research interests are artificial intelligence, soft computing, robotics and mechatronics.

Dr. Kim Soon Chong

He is an Assistant Professor in Faculty of Engineering at UCSI University in Malaysia. He received his PhD in Electrical, Eleronics & System Engineering from Universiti Kebangsaan Malaysia in 2022. His research interests are biomedical and healthcare technoologies.

Dr. Sew Sun Tiang

She is an Assistant Professor in Faculty of Engineering at UCSI University in Malaysia. She received her PhD in Electrical and Electronic Engineering from Universiti Sains Malaysia in 2014. Her research interests are optimization and antenna design.

Dr. Wei Hong Lim

He is an Associate Professor in Faculty of Engineering at UCSI University in Malaysia. He received his PhD in Computational Intelligence from Universiti Sains Malaysia in 2014. His research interests are optimization and artificial intelligence.

Tackling Photovoltaic (PV) Estimation Challenges: An Innovative AOA Variant for Improved Accuracy and Robustness

Rayan Mohammed Noor Mohammed Bakhit[1], Abhishek Sharma[2], Tiong Hoo Lim[3], Chin Hong Wong[4,5], Kim Soon Chong[1], Li Pan[1], Sew Sun Tiang[1*], Wei Hong Lim[1*]

[1]*Faculty of Engineering, Technology and Built Environment, UCSI University, Kuala Lumpur 56000, Malaysia*
[2]*Department of Computer Science and Engineering, Graphic Era Deemed to be University, Dehradun 248002, India*
[3]*Faculty of Engineering, Universiti Teknologi Brunei, Bandar Seri Begawan 1410, Brunei Darussalam*
[4]*Maynooth International Engineering College, Maynooth University, Maynooth, Co Kildare, Ireland*
[5]*Maynooth International Engineering College, Fuzhou University, Fujian, 350116, China*
Email: 1001955674@ucsiuniversity.edu.my, abhishek15491@gmail.com, lim.tiong.hoo@utb.edu.bn,
chinhong.wong@mu.ie, 1002060534@ucsiuniversity.edu.my, tiangss@ucsiuniversity.edu.my,
limwh@ucsiuniversity.edu.my

Abstract

Optimizing photovoltaic (PV) cell/module modeling is key to advancing solar power and achieving net zero carbon goals. Challenges in accurate PV parameter estimation arise from environmental variability, aging, and incomplete manufacturer data. Traditional Arithmetic Optimization Algorithm (AOA) often struggles with premature convergence due to imbalanced exploration and exploitation. This paper presents an enhanced AOA variant, incorporating chaotic maps and oppositional-based learning to better balance the optimization process. Our extensive simulations show that this improved AOA variant significantly enhances accuracy and robustness in PV cell/module parameter estimation compared to the conventional method.

Keywords: Photovoltaic module/cell, Parameter estimation, Arithmetic optimization algorithm

1. Introduction

The Intergovernmental Panel on Climate Change (IPCC) has underscored the urgent necessity to address global warming, advocating for significant efforts to attain net zero carbon emissions by 2050. This imperative demands a transition to renewable energy sources, with solar power emerging as a key player due to its sustainability and minimal environmental footprint [1]. Solar energy, captured through photovoltaic (PV) cells and modules, offers versatility, catering to a range of applications from residential heating to extensive solar farms [2]. Yet, the efficacy of solar systems is critically dependent on the performance of PV arrays, which are prone to degradation under harsh outdoor conditions.

Precision in PV system modeling is crucial for enhancing performance. This process encompasses various models, including single-diode, double-diode, and triple-diode, each demanding distinct parameter such as photocurrent and resistances [3]. Often, these parameters are not readily available from manufacturers, necessitating their estimation from experimental data. This task is further complicated by the aging of PV systems and fluctuating environmental factors [3]. The intricate and multimodal characteristics of PV parameter estimation, especially under diverse irradiance and temperature conditions, pose considerable challenges.

Existing photovoltaic (PV) cell/module parameter estimation approaches are broadly categorized into deterministic and metaheuristic methods. Deterministic methods, simpler and less computationally intensive, often fall short in accuracy under varied environmental conditions [3]. In contrast, metaheuristic methods, drawing inspiration from natural phenomena, exhibit superior global search capabilities, ease of implementation, and scalability. Consequently, they have become highly effective for diverse global optimization challenges [4], [5], [6], [7], [8], [9], [10], [11]. These metaheuristic methods also excel in addressing complex, multimodal PV cell/module parameter estimation problems [12], [13], [14], [15], [16]. Despite substantial progress using metaheuristic approaches, challenges persist due to the intricate, nonlinear interplay of PV module parameters and variable operating conditions. Therefore, developing robust metaheuristic methods for more precise PV parameter estimation is vital. Such advancements will enhance PV system optimization, maximizing solar energy utilization and contributing to a sustainable future.

Arithmetic Optimization Algorithm (AOA) [17] is a mathematics-inspired metaheuristic method that employs four basic arithmetic operations (division, multiplication, addition, and subtraction) with varying exploration and exploitation strengths for solving optimization problems.

Rayan Mohammed Noor Mohammed Bakhit, Abhishek Sharma, Tiong Hoo Lim, Chin Hong Wong, Kim Soon Chong, Li Pan, Sew Sun Tiang, Wei Hong Lim

In PV cell/module parameter estimation, the AOA population signifies diverse diode parameter combinations. Since its introduction, AOA and its variants have been effectively applied to various real-world optimization problems [18], [19], [20], [21], [22]. However, AOA, like many metaheuristic methods, relies on a conventional approach for generating its initial population, which often lacks intelligent, systematic initialization. This approach tends to produce initial solutions that are either trapped in local optimum or distant from the global optimum, thus impacting the solution's accuracy and the algorithm's convergence speed [23]. This limitation affects AOA's efficiency in complex, multimodal challenges such as PV cell/module parameter estimation.

In this paper, we introduce an enhanced variant of AOA, termed AOA with Modified Initialization Scheme (AOA-MIS), specifically designed to tackle complex and multimodal PV cell/module parameter estimation problems. The MIS module combines chaotic map strengths and dynamic oppositional learning (DOL) to generate a higher-quality initial population with improved fitness and diversity. Utilizing the non-repetitive and ergodic nature of chaotic maps, the MIS module enhances initial population diversity, thereby increasing the algorithm's robustness for complex issues. Concurrently, the DOL mechanism within MIS module effectively accelerates AOA's convergence by generating opposite solutions for those significantly distant from the global optimum. The efficacy of AOA-MIS, using various diode modeling techniques, is then benchmarked against the original AOA.

2. PV Cell/Module Parameter Estimation Problem

The electrical properties of PV systems can be modeled using different approaches: single diode model (SDM), double diode model (DDM) and triple diode model (TDM) as shown in Fig. 1. Let $j = 1, \ldots, J$ be the diode index in the model, where $J = 1$ is for SDM, $J = 2$ for DDM and $J = 3$ for TDM. For a given output voltage V, the output current I each diode model is determined as:

$$I = I_{ph} - \frac{V + IR_s}{R_{sh}} - \sum_{j=1 \to J} I_{ssdj} \left[e^{\left(\frac{q(V+IR_s)}{n_j kT} \right)} - 1 \right] \quad (1)$$

where I_{ph} and I_{sh} are the photogenerated line current and shunt resistor line current, respectively; I_{ssdj} is the saturation current of the j-th diode, R_s and R_{sh} are the series and shunt resistances, respectively; n_j is the ideality factor of the j-th diode, T is the absolute temperature, k is Boltzmann's constant, and q is the unit charge. Accurate modeling using SDM, DDM, and TDM necessitates the accurate estimation of parameters such as I_{ph}, R_s, R_{sh}, I_{ssdj} and n_j for $j = 1, \ldots, J$.

In PV cell/module parameter estimation using metaheuristic methods, appropriate objective functions are formulated based on the discrepancy between the experimental current and the model's predicted current. Considering X as a candidate solution with undetermined diode parameters, the error function for PV cells is expressed as:
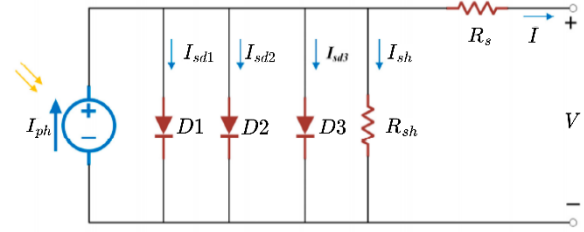
Fig.1 Equivalent circuit of PV using TDM with $J = 3$.

$$\begin{cases} f(X) = I_{ph} - \frac{V + IR_s}{R_{sh}} - \sum_{j=1 \to J} I_{ssdj} \left[e^{\left(\frac{q(V+IR_s)}{n_j kT} \right)} - 1 \right] - I \\ X = [I_{ph}, R_s, R_{sh}, I_{ssdj}, n_j], \quad \text{for } j = 1, \ldots, J \end{cases} \quad (2)$$

Root mean square error (RMSE) is commonly employed as the objective function, defined as:

$$RMSE = \sqrt{\frac{1}{K} \sum_{k=1}^{K} f^2(X)} \quad (3)$$

where k and K represent the indices and total count of measured current data points, respectively. The aim is to find a solution that minimizes the RMSE in PV cell/module parameter estimation.

3. AOA-MIS
3.1. *Proposed MIS module*

The process of generating an initial population in the proposed MIS module is described as follows. Initially, a chaotic population of size N is generated using a modified sine chaotic map, a departure from the traditional initialization schemes. This map, based on deterministic equations, can display stochastic behavior at external levels. The ergodic and non-repetitive nature of chaotic map promotes a more comprehensive search in the solution space, preventing initial solutions from being trapped in local optima and thus enhancing the diversity and robustness of the initial population against premature convergence. Define ϑ_t as the output of a chaotic variable at the t-th iteration, for $t = 1, \ldots, T$, updated as per the modified sine chaotic map with a bifurcation coefficient $\mu = \pi$:

$$\vartheta_{t+1} = \sin(\mu \vartheta_t) \quad (4)$$

The final iteration output ϑ_T initializes the d-th dimension of each n-th chaotic solution, as in Eq. (5), where X_d^L and X_d^U are the lower and upper bounds of the d-th decision variable, respectively. This results in a chaotic population $\mathbf{P^C} = [X_1^C, \ldots, X_n^C, \ldots X_N^C]$.

$$X_{n,d}^C = X_d^L + \vartheta_T (X_d^U - X_d^L) \quad (5)$$

Despite its advantages, chaotic map may still generate initial solutions far from the global optimum, potentially slowing algorithm convergence. To mitigate this, a DOL operator is applied to the chaotic population $\mathbf{P^C}$ to

produce opposite solutions for each *n*-th chaotic solution. This DOL operator broadens the initial population's coverage of the solution space, enhancing the chance of finding fitter solutions. The opposite solution for each dimension, i.e., $X_{n,d}^O$, is calculated using Eq. (6), corresponding to $X_{n,d}^C$, where $r_1, r_2 \in [0,1]$ are randomly generated numbers. This results in an opposition population, $\mathbf{P^O} = [X_1^O, \dots, X_n^O, \dots X_N^O]$.

$$X_{n,d}^O = X_{n,d}^C + r_1 \left[r_2(X_d^U + X_d^L - X_{n,d}^C) - X_{n,d}^C \right] \quad (6)$$

The two populations, $\mathbf{P^C}$ and $\mathbf{P^O}$, are combined into a single set $\mathbf{P^C} \cup \mathbf{P^O}$, resulting in a total population size of $2N$. Each solution within this merged population set is evaluated for fitness using Eq. (3), focusing on RMSE. The solutions are then rearranged based on their fitness values, from the best to the worst. The top N solutions from this sorted set $\mathbf{P^C} \cup \mathbf{P^O}$ are selected as the initial population for the AOA-MIS algorithm, denoted as $\mathbf{P} = [X_1, \dots, X_n, \dots, X_N]$.

3.2. *Iterative Search Mechanisms of AOA-MIS*

After generating the initial population $\mathbf{P}$ with the proposed MIS module, each *n*-th solution of AOA-MIS is iteratively updated using search mechanisms akin to those in the original AOA.

At each iteration, the Math Optimizer Accelerated (MOA) function value is updated for the proposed AOA-MIS to toggle between exploration and exploitation:

$$MOA(C_{Iter}) = Min + C_{Iter} \left(\frac{Max - Min}{M_{Iter}} \right) \quad (7)$$

where C_{Iter} and M_{Iter} represent the current and maximum iteration numbers, respectively; Min and Max are the minimum and maximum values of MOA. Concurrently, the Math Optimizer Probability (MP) function, guiding the search range for each solution and influenced by the critical parameter θ for exploitation efficiency, is updated:

$$MP(C_{Iter}) = 1 - \left(\frac{C_{Iter}}{M_{Iter}} \right)^{1/\theta} \quad (8)$$

Based on MOA's value, a random number $\mathbf{rand1}$ decides the search strategy (exploration or exploitation) at each iteration for updating every *d*-th dimension of the *n*-th solution, $X_{n,d}$. In the exploration phase ($\mathbf{rand1 > MOA}$), Multiplication or Division is randomly chosen:
$X_{n,d}(C_{Iter} + 1)$
$$= \begin{cases} best_d \div (MP + \varepsilon) \times [(X_d^U - X_d^L)\mu + X_d^L], rand2 < 0.5 \\ best_d \times MP \times [(X_d^U - X_d^L)\mu + X_d^L], \quad \text{Otherwise} \end{cases}$$
$$(9)$$

where $rand2$ is a random number between 0 and 1; $best_d$ is the *d*-th dimension of the current best solution; ε is a small positive number preventing division by zero; μ is a control parameter.

In the exploitation phase ($rand1 \leq MOA$), Addition or Subtraction operator updates $X_{n,d}$:

$X_{n,d}(C_{Iter} + 1) =$
$$\begin{cases} best_d - MP \times [(X_d^U - X_d^L)\mu + X_d^L], rand2 < 0.5 \\ best_d + MP \times [(X_d^U - X_d^L)\mu + X_d^L], \quad \text{Otherwise} \end{cases} \quad (10)$$

The proposed AOA-MIS iterates through this search process, following Eqs. (7) to (10), until predefined termination criteria are met, as illustrated in Fig. 2. Upon completion, the optimal diode parameters in the best solution are decoded to address the PV cell/module parameter estimation problems.

AOA-MIS for PV Cell/Module Parameter Estimation

Inputs: $D, N, M_{Iter}, T, Max, Min, \theta$
01:     Initialize $\mathbf{P^C} \leftarrow \emptyset$, $\mathbf{P^O} \leftarrow \emptyset$ and $C_{Iter} \leftarrow 0$;
02:    **for** each *n*-th solution **do**
03:      **for** each *d*-th dimension **do**
04:       Randomly generate $\vartheta_t \in [0,1]$, where $t = 0$;
05:       **while** $t \leq T$ **do**
06:        Update ϑ_t using Eq. (4);
07:        Update $t \leftarrow t + 1$;
08:       **end while**
09:       Calculate $X_{n,d}^C$ with Eq. (5);
10:       Calculate $X_{n,d}^O$ with Eq. (6);
11:      **end for**
12:      Update $\mathbf{P^C} \leftarrow \mathbf{P^C} \cup X_{n,d}^C$ and $\mathbf{P^O} \leftarrow \mathbf{P^O} \cup X_{n,d}^O$;
13:    **end for**
14:    Merge two populations as $\mathbf{P^C} \cup \mathbf{P^O}$;
15:    Fitness evaluation of all solutions stored within the merged population set of $\mathbf{P^C} \cup \mathbf{P^O}$ using Eq. (3);
16:    Sort the solutions within $\mathbf{P^C} \cup \mathbf{P^O}$ from best to worst based on their fitness values;
17:    Select the top N solutions from the sorted $\mathbf{P^C} \cup \mathbf{P^O}$ as the initial population, i.e., $\mathbf{P} = [X_1, \dots, X_n, \dots, X_N]$.
18:    Assign the first solution of $\mathbf{P}$ and its fitness as *best* and $f(best)$, respectively;
19:    **while** $C_{Iter} \leq M_{Iter}$ **do**
20:      Update *MOA* and *MP* with Eqs. (7) and (8);
21:      **for** each *n*-th solution **do**
22:       **if** $rand1 > MOA$ **then** /*Exploration*/
23:        Update $X_{n,d}(C_{Iter} + 1)$ with Eq. (9);
24:       **else** /*Exploitation*/
25:        Update $X_{n,d}(C_{Iter} + 1)$ with Eq. (10);
26:       **end if**
27:       Fitness evaluation of $X_n(C_{Iter} + 1)$ with Eq. (3);
28:       Update the X_n, $f(X_n)$, *best* and $f(best)$, with greedy selection method;
29:      **end for**
30:      $C_{Iter} \leftarrow C_{Iter} + 1$;
31:    **end while**
Output: *best* and the corresponding PV model;

Fig.2 Workflow of proposed AOA-MIS in solving the PV cell/module parameter estimation problems.

4. Results and Discussions
4.1. *Simulation settings*

In this section, the proposed AOA-MIS is applied to solve the PV cell parameter estimation problem involves the test case of R.T.C. France solar cell using SDM, DDM and TDM approaches. The proposed AOA-MIS and original AOA are implemented in MATLAB 2021a on a personal computer consisting of an Intel® Core™

Rayan Mohammed Noor Mohammed Bakhit, Abhishek Sharma, Tiong Hoo Lim, Chin Hong Wong, Kim Soon Chong, Li Pan, Sew Sun Tiang, Wei Hong Lim

i7-HQ CPU, 2.50 GHz, and 16 GB RAM laptop. For all compared techniques, the population size and maximum iteration numbers are set as 30 and 1000, respectively.

4.2. *Performance analysis*

The current and voltage experimental values for the R.T.C. France solar cell were recorded under standard conditions: 1000 W/m² at 33°C. This subsection applies the proposed AOA-MIS to estimate five and seven unknown parameters of the SDM and DDM, respectively, representing the R.T.C. solar cell at these conditions. The AOA-MIS results are compared with those from the original AOA. RMSE serves as the performance index to evaluate AOA-MIS's effectiveness.

Table 1 and Table 2 showcase the optimized parameters derived from AOA-MIS and the original AOA for the R.T.C. France solar cell, using the SDM and DDM, respectively, along with their corresponding RMSE values. Parameters yielding better results, as indicated by lower RMSE values, are highlighted in bold. Table 1 reveals that AOA-MIS's estimation of the five parameters (i.e., I_{ph}, R_s, R_{sh}, I_{ssd} and n) for the SDM leads to higher modeling accuracy for the R.T.C. France solar cell, as evidenced by lower RMSE values, compared to the original AOA. Similarly, Table 2 shows that AOA-MIS's estimation of the seven parameters (i.e., I_{ph}, R_s, R_{sh}, $I_{ssd,1}$, $I_{ssd,2}$, n_1 and n_2) for the DDM yields more competitive RMSE values.

Table 1. Optimized parameters obtained for R.T.C France solar cell using SDM.

Parameters	AOA	AOA-MIS
I_{ph}	0.75483	0.75822
R_s	0.052023	0.039672
R_{sh}	97.805	69.689
I_{ssd}	1.5746×10^{-8}	1.5411×10^{-7}
n	1.2485	1.4316
RSME	5.3572×10^{-3}	$\mathbf{3.2507\times10^{-3}}$

Table 2. Optimized parameters obtained for R.T.C France solar cell using DDM.

Parameters	AOA	AOA-MIS
I_{ph}	0.73552	0.80916
R_s	0.0044392	0
R_{sh}	2.668	1.7305
$I_{ssd,1}$	0	0
$I_{ssd,2}$	0	0
n_1	1.1428	1.0141
n_2	1.9816	1.235
RSME	1.7449×10^{-1}	$\mathbf{1.705\times10^{-1}}$

To further assess the effectiveness of the proposed AOA-MIS in addressing the PV cell/module parameter estimation problems, Table 3 provides a statistical analysis of the RMSE values for the R.T.C. France solar cell using both SDM and DDM models, comparing AOA-MIS and the original AOA. This analysis includes performance metrics such as the minimum (Min), maximum (Max), mean (Mean), and standard deviation (SD) of the RMSE values from multiple simulation runs

for both algorithms. The data in Table 3 reveal that AOA-MIS consistently outperforms the original AOA in terms of Max, Min, and Mean RMSE values. Additionally, AOA-MIS shows superior consistency in achieving lower RMSE values, as evidenced by lower SD values in both SDM and DDM test cases.

Table 3. Statistical results of the RMSE values of R.T.C France solar cell represented by different diode modelling methods

Model	RSME	AOA	AOA-MIS
SDM	Min	5.3572×10^{-3}	$\mathbf{3.2507\times10^{-3}}$
	Mean	3.0009×10^{-2}	$\mathbf{2.4830\times10^{-2}}$
	Max	1.4549×10^{-1}	$\mathbf{6.7476\times10^{-2}}$
	SD	4.2432×10^{-2}	$\mathbf{2.0059\times10^{-2}}$
DDM	Min	1.7449×10^{-1}	$\mathbf{1.7050\times10^{-1}}$
	Mean	1.9299×10^{-1}	$\mathbf{1.7875\times10^{-1}}$
	Max	2.1688×10^{-1}	$\mathbf{1.8938\times10^{-1}}$
	SD	1.2098×10^{-2}	$\mathbf{5.8922\times10^{-3}}$

The superior performance of the proposed AIA-MIS over the original AOA can be attributed to several factors. PV cell/module parameter estimation is a complex, multimodal real-world optimization problem, largely due to the nonlinear relationship between model parameters and varying operating conditions such as temperature and irradiance levels. The complexity of this problem escalates with the number of diodes used in modelling (e.g., five parameters in SDM, seven parameters in DDM and so on), further complicating the optimization task. Table 1, Table 2, and Table 3 suggest that the quality of the initial population is crucial in enabling AOA to accurately model PV cell/module parameters. The original AOA, using a conventional initialization scheme, tends to generate the initial solutions in local or non-optimal regions, leading to premature convergence due to the absence of intelligent methods and knowledge of the surrounding search environment. Moreover, there is a significant risk of initializing solutions far from the global optimum, slowing down the algorithm's convergence speed. In contrast, the MIS module in AOA-MIS generates an initial population of higher quality in terms of fitness and diversity. The chaotic map's non-repetitive and ergodic nature in the MIS module promotes a more comprehensive search of the solution space, minimizing the risk of local optima entrapment and premature convergence. DOL, another key mechanism in MIS, accelerates convergence by enhancing the exploration of search range through generating opposite solutions from those initialized by the chaotic map. This synergistic effect of the chaotic map and DOL in the MIS module bolsters AOA-MIS's robustness in navigating complex, multimodal search spaces, thereby enhancing its accuracy in solving PV cell/module parameter estimation problems.

5. Conclusion

In this paper, we introduce an enhanced version of the Arithmetic Optimization Algorithm, termed AOA-MIS, to address the complex and multimodal challenges of PV

cell/module parameter estimation. The innovation of AOA-MIS lies in integrating chaotic maps and DOL mechanisms into the Modified Initialization Scheme (MIS) module, thereby generating an initial population with improved fitness and diversity. Simulation results reveal that AOA-MIS, with its superior initial population quality, outperforms the original AOA in estimating parameters for the SDM and DDM in representing the R.T.C. France solar cell, tested under standard conditions of 1000 W/m² at 33°C. As one of the future works, AOA-MIS could be further enhanced by incorporating an adaptive search mechanism, potentially increasing its robustness and efficacy in addressing various complexities in PV cell/module parameter estimation.

Acknowledgements

References

1. M. Premkumar and T. R. Sumithira, Design and implementation of new topology for solar PV based transformerless forward microinverter, Journal of Electrical Engineering & Technology 14(1), 2019, pp. 145-155.
2. D. Jung, A. Salmon and P. Gese, Agrivoltaics for farmers with shadow and electricity demand: Results of a pre-feasibility study under net billing in central Chile, AIP Conference Proceedings 2361(1), 2021, pp. 030001.
3. Z. Gu, G. Xiong and X. Fu, Parameter extraction of solar photovoltaic cell and module models with metaheuristic algorithms: A review, Sustainability 15(4), 2023, pp. 3312.
4. E. S. Ghith and F. A. A. Tolba, Real-time implementation of an enhanced proportional-integral-derivative controller based on sparrow search algorithm for micro-robotics system, IAES International Journal of Artificial Intelligence 11(4), 2022, pp. 1395-1404.
5. E. S. Ghith and F. A. A. Tolba, LabVIEW implementation of tuning PID controller using advanced control optimization techniques for micro-robotics system. International Journal of Mechanical Engineering and Robotics Research 11(9), 2022, pp. 653-661.
6. E. S. Ghith and F. A. Tolba, Real-time implementation of an enhanced PID controller based on marine predator algorithm (MPA) for micro-robotics system, 2022 3rd International Conference on Artificial Intelligence, Robotics and Control (AIRC), Cairo, Egypt, 2022.
7. K. M. Ang et al., Modified teaching-learning-based optimization and applications in multi-response machining processes, Computers & Industrial Engineering 174, 2022, pp. 108719.
8. C. Hassan, S. Sapuan, N. Abd. Aziz and M. Z. Mohamed Yusof, Effect of chemical treatment on the tensile properties of single oil palm empty fruit bunch (OPEFB) fibre. Trends in Textile Engineering & Fashion Technology 3(2), 2018, pp. 1-7.
9. E. Natarajan, C. Hassan, C. K. Ang, M. S. Santhosh, S. Ramesh and R. Sasikumar, Modeling of multiwall carbon nanotubes reinforced natural rubber for soft robotic applications - A comprehensive presentation 46(9), 2021, pp. 3251-3258.
10. C. Hassan C. W. Yeo, B. Shari, M. S. Salit and N. Abdul Aziz, Mechanical properties of unidirectional oil palm empty fruit bunch (OPEFB) fiber reinforced epoxy composite. IOP Conference Series: Materials Science and Engineering 206(1), 2017, pp 012045.
11. L. Yao et al., Demand bidding optimization for an aggregator with genetic algorithm. Energies 11(1), 2018, pp. 1-22.
12. M. Premkumar et al., A reliable optimization framework for parameter identification of single-diode solar photovoltaic model using weighted velocity-guided grey wolf optimization algorithm and Lambert-W function, IET Renewable Power Generation 17(11), 2023, pp. 2711-2732.
13. A. Sharma et al., Identification of photovoltaic module parameters by implementing a novel teaching learning based optimization with unique exemplar generation scheme (TLBO-UEGS), Energy Report 10, 2023, pp. 1485-1506.
14. A. Sharma et al., Performance investigation of state-of-the-art metaheuristic techniques for parameter extraction of solar cells/module, Scientific Reports 13, 2023, pp. 11134.
15. A. Sharma et al., An effective method for parameter estimation of solar PV cell using grey-wolf optimization technique, International Journal of Mathematical, Engineering and Management Science 6(3), 2021, pp. 911-931.
16. H. Bakır, Comparative performance analysis of metaheuristic search algorithms in parameter extraction for various solar cell models, Environmental Challenges 11, 2023, pp. 100720.
17. L. Abualigah, A. Diabat, S. Mirjalili, M. Abd Elaziz and A. H. Gandomi, The arithmetic optimization algorithm, Computer Methods in Applied Mechanics and Engineering 376, 2021, pp. 113609.
18. E. S. Ghith and F. A. A. Tolba, Tuning PID controllers based on hybrid arithmetic optimization algorithm and artificial gorilla troop optimization for micro-robotic systems. IEEE Access 11, 2023, pp. 27138-27154.
19. S. Rajput et al., A novel opposition-based arithmetic optimization algorithm for parameter extraction of PEM fuel cell. Electronics 10(2), 2021, pp. 2834.
20. K. M. Ang et al., Training feedforward neural networks using arithmetic optimization algorithm for medical classification, in Advances in Intelligent Manufacturing and Mechatronics. Lecture Notes in Electrical Engineering, eds. M. A. Abdullah et al. Springer, Singapore, vol. 988, 2023.
21. A. A. Mohamed et al., A novel hybrid arithmetic optimization algorithm and salp swarm algorithm for data placement in cloud computing, Soft Computing 27(9), 2023, pp. 5769-5780.
22. X. Shi, X. Yu, and M. Esmaeili-Falak, Improved arithmetic optimization algorithm and its application to carbon fiber reinforced polymer-steel bond strength estimation, Composite Structures 306, 2023, pp. 116599.
23. M. F. Ahamad, N. A. Mat Isa, W. H. Lim and K. M. Ang, Differential evolution with modified initialization scheme using chaotic oppositional based learning strategy, Alexandria Engineering Journal 61(12), 2022, pp. 11835-11858.

Rayan Mohammed Noor Mohammed Bakhit, Abhishek Sharma, Tiong Hoo Lim, Chin Hong Wong, Kim Soon Chong, Li Pan, Sew Sun Tiang, Wei Hong Lim

Authors Introduction

Mr. Rayan Mohammed Noor Mohammed Bakhit

He received the Bachelor of Mechatronics Engineering with Honours in Faculty of Engineering, Technology and Built Environment, UCSI University, Malaysia, in 2023. His research interests are renewable energy, and optimization algorithm.

Dr. Abhishek Sharma

He is a Research Assistant Professor at Graphic Era Deemed to be University in India. He received his PhD from University of Petroluem & Energy Studies in 2022. His research interests are artificial intellgence and power electronics.

Dr. Tiong Hoo Lim

He is a Senior Assistant Professor and Director of Planning Development Office at Universiti Teknologi Brunei in Brunie Darussalam. He received his PhD in Computer Science from University of York, United Kingdom. His research interests are artificial intelligence, advanced technology, competitive and forecasting analysis.

Dr. Chin Hong Wong

He is a Lecturer in Maynooth International Engineering College at Fuzhou University in China. He received his PhD in Electrical and Electronic Engineering from Universiti Sains Malaysia in 2017. His research interests are Energy harvesting and control system.

Dr. Kim Soon Chong

He is an Assistant Professor in Faculty of Engineering at UCSI University in Malaysia. He received his PhD in Electrical, Eleronics & System Engineering from Universiti Kebangsaan Malaysia in 2022. His research interests are biomedical and healthcare technoologies.

Ms. Li Pan

She received her Master of Engineering in Computer Technology degree from Huazhong University of Science and Technology, China in 2008. She is currently a Doctoral research student in UCSI University, Malaysia.

Dr. Sew Sun Tiang

She is an Assistant Professor in Faculty of Engineering at UCSI University in Malaysia. She received her PhD in Electrical and Electronic Engineering from Universiti Sains Malaysia in 2014. Her research interests are optimization and antenna design.

Dr. Wei Hong Lim

He is an Associate Professor in Faculty of Engineering at UCSI University in Malaysia. He received his PhD in Computational Intelligence from Universiti Sains Malaysia in 2014. His research interests are optimization and artificial intelligence.

Deep Learning in Manufacturing: A Focus on Welding Defect Classification with CNNs

Tin Chang Ting[1], Hameedur Rahman[2], Tiong Hoo Lim[3], Chin Hong Wong[4,5], Chun Kit Ang[1], Mohamed Khan Afthab Ahamed Khan[1], Sew Sun Tiang[1*], Wei Hong Lim[1*]

[1]*Faculty of Engineering, Technology and Built Environment, UCSI University, Kuala Lumpur 56000, Malaysia*
[2]*Faculty of Computing and Artificial Intelligence, Air University, Islamabad Capital Territory 44000, Pakistan*
[3]*Faculty of Engineering, Universiti Teknologi Brunei, Bandar Seri Begawan 1410, Brunei Darussalam*
[4]*Maynooth International Engineering College, Maynooth University, Maynooth, Co Kildare, Ireland*
[5]*Maynooth International Engineering College, Fuzhou University, Fujian, 350116, China*
Email: 1002058096@ucsiuniversity.edu.my, rhameedur@gmail.com, lim.tiong.hoo@utb.edu.bn,
chinhong.wong@mu.ie, angck@ucsiuniversity.edu.my, mohamedkhan@ucsiuniversity.edu.my,
tiangss@ucsiuniversity.edu.my, limwh@ucsiuniversity.edu.my

Abstract

Welding is integral to modern manufacturing, yet the complex process often leads to defects, impacting the quality of the final product. Recent advances in deep learning, particularly Convolutional Neural Networks (CNNs), have shown remarkable results in applications like defect recognition. This study evaluated AlexNet, ResNet-18, ResNet-50, ResNet-101, MobileNet-v2, ShuffleNet, and SqueezeNet for their effectiveness in identifying welding defects, using accuracy, precision, sensitivity, specificity, and F-score as metrics. The dataset covered defects like cracks, lack of penetration, porosity, and a no-defect class. Our analysis shows that most of these architectures deliver promising results in accuracy, sensitivity, specificity, precision, and F1-score, highlighting their potential in defect recognition.

Keywords: Convolutional neural network, Classification, Deep learning, Welding defects

1. Introduction

Welding is a crucial industrial process integral to various production sectors, including high-performance industries like aerospace, automotive, marine, and power generation. This complex process involves multiple parameters that directly impact the quality of the weld joint. Factors such as welding procedures, methods, environmental conditions, and the operator's skill level can lead to various defects during the welding of pipelines and pressure vessels [1]. Common defects include blowholes, cracks, incomplete fusion, incomplete penetration, slag inclusions, and undercutting, which significantly compromise the sealing and strength of the products. In addition to welding parameters, unforeseen events in the manufacturing process can also cause weld defects. Therefore, rigorous welding quality inspection and testing are essential during manufacturing, especially for products like pipelines and pressure vessels. This is critical to identify the root causes of welding defects and implement targeted corrective measures to ensure product quality and safety [1].

Considerable research has been devoted to addressing the challenges of weld defect detection and identification. These defects are typically detected using non-destructive testing (NDT) techniques, favored for their non-invasive interaction with specimens. Common NDT methods for weld defect detection fall into various categories, including visual or manual inspection, radiographic testing with ionizing radiation sources (such as gamma rays or X-rays), eddy current testing, ultrasonic testing, and dye penetrant testing [2], [3]. However, each of these techniques has inherent limitations. For example, eddy current testing is only applicable to metallic specimens, while X-ray radiographic testing poses potential health risks due to prolonged exposure to radiation. Additionally, the majority of current weld surface defect recognition relies on manual inspection by NDT experts analyzing these radiographic images. This manual process of interpreting and evaluating images can be complex, subjective, inconsistent, time-consuming, labor-intensive, and prone to errors, particularly when distinguishing between defects with similar features [4]. Consequently, developing an automated inspection solution that offers more efficient and accurate recognition of weld surface defects is essential to overcome the drawbacks of manual inspections.

The adoption of automated machine vision systems, integrating deep learning techniques, offers a promising solution for weld surface image classification challenges. Convolutional Neural Networks (CNNs), renowned for their ability to learn nonlinear relationships between inputs and outputs, mimic the human brain's learning process and have been effectively applied in various real-world scenarios. These applications include signal/image classification [5], [6], [7], cybersecurity [8], [9], medical diagnosis [10], [11], [12], fault detection [13], [14], [15]

Tin Chang Ting, Hameedur Rahman, Tiong Hoo Lim, Chin Hong Wong, Chun Kit Ang, Mohamed Khan Afthab Ahamed Khan, Sew Sun Tiang, Wei Hong Lim

and prediction [16], [17]. Encouraged by deep learning's success, numerous network architectures have been developed, such as AlexNet [18], ResNet-18 [19], ResNet-50 [19], ResNet-101 [19], MobileNet-v2 [20], ShuffleNet [21], and SqueezeNet [22]. Transfer learning, a method for adapting these existing CNN architectures to new tasks, has gained popularity. It utilizes smaller datasets and reduces training time, making it an efficient approach for training CNNs in new application domains.

While previous studies [4], [23], [24], [25], [26], [27] have explored welding defect classification using deep learning, the optimal selection of CNN architectures remains relatively unexplored, leaving many architectures yet to be assessed. Addressing this gap, our paper presents an extensive study evaluating the performance of seven popular CNN architectures: AlexNet, ResNet-18, ResNet-50, ResNet-101, MobileNet-v2, ShuffleNet, and SqueezeNet, in classifying weld defects from digital radiographic images. Our aim is to objectively identify the most effective CNN architecture using various performance metrics. This study could lead to replacing manual inspection methods with a more accurate automated weld defect classification system, potentially reducing production costs and increasing throughput.

2. Methodology
2.1. *Data acquisition and preprocessing*

This study focuses on training and evaluating a deep learning model using the RIAWELC welding defect dataset [28]. Unlike other commonly used datasets such as GDXRay [29] and WDXI [30], RIAWELC offers a larger, open-source dataset, which is beneficial for training CNNs without the risk of overfitting.

The RIAWELC dataset comprises 24,407 radiographic images, each of size 224×224 pixels, categorized into four classes of welding defects: lack of penetration (LP), cracks (CR), porosity (PO), and no defect (ND). A representative image for each defect type is displayed in Fig. 1, and Table 1 details the distribution of images across these classes. The substantial size of RIAWELC dataset supports the development of automated methods for identifying and classifying welding defects, crucial for reliable inspection and quality control.
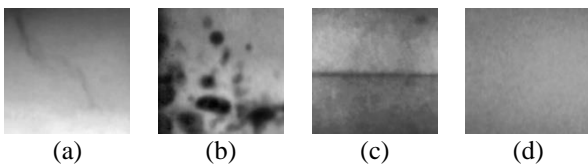
Fig.1 Sample of radiographic images for each welding defect class from RIAWELC dataset: (a) LP, (b) CR, (c) PO and (d) ND.

Table 1. Data distribution of RIAWELC dataset for each welding defect class.

Defect Types	CR	PO	LP	ND
No. of Image	7,635	6,320	4,452	6,000

To facilitate deep learning model training, the RIAWELC dataset is divided into three subsets: 70% for training, 10% for validation, and 20% for testing. Additionally, all radiographic images from the dataset are resized to match the input requirements of the selected pretrained network, as outlined in Table 2, ensuring compatibility with the network architectures.

Table 2. Summary of seven pretrained networks

Pretrained Networks	Depth	Size (MB)	Parameter (Millions)	Input Size
AlexNet	8	227	61.0	224×224
ResNet-18	18	44	11.7	224×224
ResNet-50	50	96	25.6	224×224
ResNet-101	101	167	44.6	224×224
MobileNet-v2	53	13	3.5	224×224
ShuffleNet	50	5.4	1.4	224×224
SqueezeNet	18	5.2	1.24	227×227

2.2. *Transfer learning of pretrained networks*

Training CNNs from scratch for specific tasks presents challenges due to significant resource requirements, such as training time, infrastructure, and input datasets. Transfer learning offers a viable solution by transferring knowledge from one or more source domains to a different target domain. In this study, we utilize transfer learning to extract the learnable parameters from selected pretrained network architectures (namely, AlexNet, ResNet-18, ResNet-50, ResNet-101, MobileNet-v2, ShuffleNet, and SqueezeNet) for the welding defect radiographic images classification tasks.

Specifically, the last three layers of these pretrained networks are replaced with a new fully-connected layer, a SoftMax layer, and a classification output layer. Additionally, the original output layers are substituted with new output layers tailored to four classes: lack of penetration (LP), cracks (CR), porosity (PO), and no defect (ND). These modified networks are then trained using the RIAWELC dataset, as detailed in Table 1.

2.3. *Hyperparameter tuning*

Stochastic Gradient Descent (SGD) is a popular optimizer in deep learning training, valued for its ability to effectively balance accuracy and efficiency. In this study, SGD is employed to train selected pretrained network architectures with the RIAWELC dataset, aiming to minimize the cross-entropy loss function. However, the performance and convergence of SGD are influenced by various training hyperparameters, including momentum, initial learning rate, L2 regularization, maximum epoch, and batch size.

Specifically, the initial learning rate dictates the step sizes in the parameter update process. L2 regularization helps to prevent network overfitting by adding a penalty term to the loss function. The maximum epoch limits the number of iterations to avoid overfitting, and the batch size determines the number of samples used per iteration, impacting the stability, speed, and memory usage of SGD.

To facilitate a fair comparison in performance evaluations, the hyperparameters for all selected pretrained networks are standardized. This includes setting the momentum to 0.9, the initial learning rate to 2×10^{-3}, L2 regularization to 0.5, the maximum epoch to 5, and the batch size to 32.

2.4. *Performance metrics*

The overall performance of all pretrained CNN networks in classifying welding defect radiographic images is evaluated using five key metrics: accuracy, sensitivity, specificity, precision and F1 score. These metrics are derived from the true positive (TP), true negative (TN), false positive (FP), and false negative (FN) outcomes produced during the testing phase of the CNN architectures. The mathematical formulations for each performance metric employed in this study are as follows.

Accuracy, indicating the overall correctness of the CNN architecture's predictions, is calculated as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Sensitivity, measuring the CNN architecture's ability to correctly identify positive results, is computed as:

$$Sensitivity = \frac{TP}{TP + FN} \quad (2)$$

Specificity, gauging the CNN architecture's capacity to accurately identify negative results, is determined as:

$$Specificity = \frac{TN}{TN + FP} \quad (3)$$

Precision, assessing the accuracy of positive predictions among all positive cases identified, is calculated as:

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

F1 Score, the harmonic mean of precision and sensitivity, is defined as:

$$F1\ Score = 2 \times \left(\frac{Sensitivity \times Precision}{Sensitivity + Precision} \right) \quad (5)$$

The abovementioned five performance metrics provide a comprehensive assessment of the CNN architectures' capabilities in distinguishing different classes of welding defects.

3. Results and Discussions

The performance of all selected pretrained networks, including AlexNet, ResNet-18, ResNet-50, ResNet-101, MobileNet-v2, ShuffleNet, and SqueezeNet, in classifying radiographic images of welding defects is comprehensively evaluated. To facilitate a quantitative comparison, each CNN architecture's effectiveness in the welding defect classification task is assessed using the five performance metrics previously mentioned, namely accuracy (Acc.), sensitivity (Sens.), specificity (Spec.), precision (Prec.) and F1 score. The simulation results are detailed in Table 3, where the highest values attained by each CNN architecture for these metrics are highlighted in boldface.

Table 3. Quantitative performance evaluation of all selected CNN architectures for welding defect classification task

CNN architecture	Acc. (%)	Sens. (%)	Spec. (%)	Prec. (%)	F1 (%)
AlexNet	71.43	71.43	90.48	80.21	70.92
ResNet-18	78.57	78.57	92.86	80.56	78.92
ResNet-50	82.14	82.14	94.05	83.93	82.49
ResNet-101	**85.71**	**85.71**	**95.24**	**87.50**	**86.06**
MobileNet-v2	71.43	71.43	90.48	69.00	69.06
ShuffleNet	75.00	75.00	91.67	80.63	75.98
SqueezeNet	71.43	71.43	90.48	75.00	70.92

Table 3 reveals a significant observation across all CNN architectures: the identical values of accuracy and sensitivity when classifying the RIAWELC dataset. This uniformity implies that the models' ability to correctly predict positive instances (sensitivity) significantly impacts their overall accuracy. Among these, ResNet-101 (85.71%) and ResNet-50 (82.14%) stand out with the best and second-best performance in both accuracy and sensitivity, highlighting their superior detection capabilities for welding defects in radiographic images. Conversely, architectures with lower complexity, such as AlexNet, MobileNet-v2, and SqueezeNet, demonstrate less effectiveness, each recording 71.43% in both accuracy and sensitivity.

Moreover, Table 3 indicates that all selected CNN architectures consistently achieve higher specificity and precision compared to their accuracy and sensitivity. This suggests they are more effective in correctly identifying true negatives (specificity) and ensuring accurate positive predictions (precision). Specifically, ResNet-101 exhibits the highest specificity (95.24%), while AlexNet shows the lowest (90.48%), indicating a broad capability across these models to identify non-defective instances accurately. Precision, however, varies more across the models, with ResNet-101 leading at 87.50%, suggesting its higher accuracy in classifying weld defects. Conversely, MobileNet-v2, despite similar accuracy and sensitivity to other models, records the lowest precision (69.00%), implying a higher rate of false positives.

The F1 score, reflecting the harmonic mean of precision and sensitivity, offers a more balanced evaluation of model performance. ResNet-101 and ResNet-50 achieve

Tin Chang Ting, Hameedur Rahman, Tiong Hoo Lim, Chin Hong Wong, Chun Kit Ang, Mohamed Khan Afthab Ahamed Khan, Sew Sun Tiang, Wei Hong Lim

the highest F1 scores (86.06% and 82.49%, respectively), confirming their robust overall performance in classifying the radiographic images of welding defects. In contrast, AlexNet, MobileNet-v2, and SqueezeNet exhibit lower F1 scores (70.92%, 69.06%, and 70.92%, respectively), indicating a compromise between sensitivity and precision in these models.

Analysis of the simulation results in Table 3 reveals that more complex CNN architectures, such as ResNet-101 and ResNet-50, consistently outperform simpler ones like AlexNet and SqueezeNet across all performance metrics. This superior performance of advanced models like ResNet-101 and ResNet-50 highlights the significance of choosing the right model for complex image classification tasks, such as welding defect classification. It suggests that deeper networks are more adept at extracting essential features necessary for accurately identifying welding defects. The enhanced capability of these complex CNN models to discern intricate patterns in data meets the expectations associated with deeper network architectures.

However, it is important to note that while these sophisticated CNN architectures offer improved classification accuracy, they also demand greater computational resources. In contrast, simpler CNN models, despite being less precise, are more computationally efficient and may be preferable in scenarios with limited resources. In practical applications, factors like available computational power, latency considerations, and the criticality of accurately detecting defects should guide the selection of CNN architectures. Balancing these considerations is key to effectively deploying CNNs in real-world tasks.

4. Conclusions

This paper introduces a deep learning-based machine vision inspection algorithm that combines pretrained CNN architectures and transfer learning for classifying four types of welding defects—lack of penetration, cracks, porosity, and no defect—using radiographic images from the RIAWELC dataset. The objective is to thoroughly analyze the efficacy of seven different pretrained CNN architectures: AlexNet, ResNet-18, ResNet-50, ResNet-101, MobileNet-v2, ShuffleNet, and SqueezeNet, in addressing welding defect classification tasks. This analysis employs various performance metrics, including accuracy, sensitivity, specificity, precision, and F1 score. Simulation studies reveal that these pretrained networks exhibit diverse performance levels in defect classification, with ResNet-101 emerging as the most effective, achieving 85.71% accuracy, 85.71% sensitivity, 95.24% specificity, 87.50% precision, and an F1 score of 86.06%. In contrast, MobileNet-v2 shows the least effectiveness, with the lowest scores across all metrics: 71.43% accuracy and sensitivity, 90.48% specificity, 69.00% precision, and an F1 score of 69.06%. These results offer valuable insights into the applicability of pretrained deep learning networks for welding defect classification, providing a significant benefit to

manufacturing companies seeking to enhance their quality control processes. In this study, the hyperparameter settings for all pretrained networks are manually configured. It is anticipated that their classification performance could be further improved by employing advanced metaheuristic search algorithms to optimize these hyperparameter settings, thereby refining the training process of the pretrained networks.

Acknowledgements

References

1. Q. Feng, R. Li, B. Nie, S. Liu, L. Zhao, and H. Zhang, Literature review: Theory and application of in-line inspection technologies for oil and gas pipeline girth weld defection, Sensors 17(1), 2017, pp. 50.
2. L. Yin et al., A novel feature extraction method of eddy current testing for defect detection based on machine learning, NDT & E International 107, 2019, pp. 102108.
3. C. Pei, D. Yi, T. Liu, X. Kou and Z. Chen, Fully noncontact measurement of inner cracks in thick specimen with fiber-phased-array laser ultrasonic technique, NDT & E International 113, 2020, pp. 102273.
4. K. Ding, Z. Niu, J. Hui, X. Zhou, and F. T. S. Chan, A weld surface defect recognition method based on improved MobileNetV2 algorithm, Mathematics 10(19), 2022, pp. 3678.
5. B. Jdid, W. H. Lim, I. Dayoub, K. Hassan and M. R. B. M. Juhari, Robust automatic modulation recognition through joint contribution of hand-crafted and contextual features, IEEE Access 9, 2021, pp. 104530-104546.
6. K. M. Ang et al, Optimal design of convolutional neural network architectures using teaching-learning-based optimization in image classification, Symmetry 14(11), 2022, pp. 2323.
7. K. M. Ang et al., An innovative approach for automated convolutional neural network design for image classification, Mathematics 11(19), 2023, pp. 4115.
8. T. Berghout, M. Benbouzid and Y. Amirat, Towards resilient and secure smart grids against PMU adversarial attacks: A deep learning-based robust data engineering approach, Electronics 12(12), 2023, pp. 2554.
9. T. Berghout and M. Benbouzid, EL-NAHL: Exploring labels autoencoding in augmented hidden layers of feedforward neural networks for cybersecurity in smart grids, Reliability Engineering & System Safety 226, 2022, pp. 108680.
10. L. S. Chow, G. S. Tang, M. I. Solihin, N. M. Gowdh, N. Ramli and K. Rahmat, Quantitative and qualitative analysis of 18 deep convolutional neural network (CNN) models with transfer learning to diagnose COVID-19 on chest X-ray (CXR) images, SN Computer Science 4(2), 2023, pp. 141.
11. A. Qayyum et al., Hybrid 3D-ResNet Deep Learning Model for Automatic Segmentation of Thoracic Organs at Risk in CT Images, 2020 International Conference on

Industrial Engineering, Applications and Manufacturing (ICIEAM), Sochi, Russia, 2020, pp. 1-5.

12. H. Rahman et al. A systematic literature review of 3D deep learning techniques in computed tomography reconstruction. Tomography 9(6), 2023, pp. 2158-2189.

13. M. Alrifaey, W. H. Lim and C. K. Ang, A novel deep learning framework based RNN-SAE for fault detection of electrical gas generator, IEEE Access, 9, 2021, pp. 21433-21442.

14. M. Alrifaey et al., Hybrid deep learning model for fault detection and classification of grid-connected photovoltaic system, IEEE Access 10, 2022, pp. 13852-13869.

15. T. Berghout et al. Federated learning for condition monitoring of industrial processes: A review on fault diagnosis method, challenges and prospect. Electronics 12(1), pp. 158.

16. T. Berghout, M. Benbouzid, Y. Amirat and G. Yao, Lithium-ion battery state of health prediction with a robust collaborative augmented hidden layer feedforward neural network approach. IEEE Transactions on Transportation Electrification 9(3), 2023, pp. 4492-4502.

17. A. A. Abdelhamid et al. Deep learning with dipper throated optimization algorithm for energy consumption forecasting in smart households. Energies 15(23), 2022, pp. 9125.

18. A. Krizhevsky, I. Sutskever and G. E. Hinton, ImageNet classification with deep convolutional neural networks, Communication of the ACM 60(6), 2017, pp. 84-90.

19. K. He, X. Zhang, S. Ren and J. Sun, Deep Residual Learning for Image Recognition, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770-778.

20. M. Sandler et al, MobileNetV2: Inverted Residuals and Linear Bottlenecks. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 2018, pp. 4510-4520.

21. X. Zhang, X. Zhou, M. Lin and J. Sun, ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 2018, pp. 6848-6856.

22. F.Iandola et al. SqueezeNet: AlexNet-level Accuracy with 50x Fewer Parameters and <IMB Model Size, Preprint, submitted November 4, 2016. https://arxiv.org/pdf/1602.07360

23. S. Kumaresan, K. S. Jai Aultrin, S. S. Kumar and M. Dev Anand, Deep learning-based weld defect classification using VGG16 transfer learning adaptive fine tuning. International Journal on Interactive Design and Manufacturing 17, 2023, pp. 2999-3010.

24. W. Hou, Y. Wei, J. Guo, Y. Jin and C. Zhu, Automatic detection of welding defects using deep neural network, Journal of Physics: Conference Series 933, 2017, pp. 012006.

25. W. Du, H. Shen, J. Fu, G. Zhang and Q. He, Approaches for improvement of the X-ray image defect detection of automobile casting aluminium parts based on deep learning, NDT & E International 107, 2019, pp. 102144.

26. S. Kumaresan, K. S. Jai Aultrin, S. S. Kumar and M. Dev Anand, Transfer learning with CNN for classification of weld defect, IEEE Access 9, 2021, pp. 95097-951078.

27. D. Say, S. Zidi, S. M. Qaisar and M. Krichen, Automated categorization of multiclass welding defects using the X-ray image augmentation and convolutional neural network, Sensors 23(14), 2023, pp. 6422.

28. S. Perri, F. Spagnolo, F. Frustaci and P. Corsonello, Welding defects classification through a convolutional neural network, Manufacturing Letters 35(42), 2022, pp. 29-32.

29. D. Mery et al, GDXray: The database of X-ray images for nondestructive testing, Journal of Nondestructive Evaluation 34(42), 2015, pp. 42.

30. W. Guo, H. Qu and L. Liang, WDXI: The Dataset of X-ray Image for Weld Defects, 2018 14th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), Huangshan, China, 2018, pp. 1051-1055.

Authors Introduction

Mr. Tin Chang Ting

He is currently pursuing the Bachelor of Mechatronics Engineering with Honours in Faculty of Engineering, Technology and Built Environment, UCSI University, Malaysia. His research interests are machine learning, deep learning, and optimization algorithm.

Dr. Hameedur Rahman

He is an Associate Professor in Faculty of Computing and AI at AIR University in Pakistan. He received his PhD in Computer Science from Universiti Kebangsaan Malaysia in 2018. His research interests are Virual/Augmented Rality, Image Processing, Data Mining, Artificial Intelligence, Natural Language Processing and CyberSecurity.

Dr. Tiong Hoo Lim

He is a Senior Assistant Professor and Director of Planning Development Office at Universiti Teknologi Brunei in Brunie Darussalam. He received his PhD in Computer Science from University of York, United Kingdom. His research interests are artificial intelligence, advanced technology, competitive and forecasting analysis.

Dr. Chin Hong Wong

He is a Lecturer in Maynooth International Engineering College at Fuzhou University in China. He received his PhD in Electrical and Electronic Engineering from Universiti Sains Malaysia in 2017. His research interests are Energy harvesting and control system.

Tin Chang Ting, Hameedur Rahman, Tiong Hoo Lim, Chin Hong Wong, Chun Kit Ang, Mohamed Khan Afthab Ahamed Khan, Sew Sun Tiang, Wei Hong Lim

Dr. Chun Kit Ang	He is the Dean and Associate Professor in Faculty of Engineering at UCSI University in Malaysia. He received his PhD in Mechanical and Manufacturing Engineering from Universiti Putra Malaysia in 2014. His research interests are artificial intelligence, soft computing, robotics and mechatronics.

Dr. Mohamed Khan Afthab Ahmed Khan	He is an Assistant Professor in Faculty of Engineering at UCSI University in Malaysia. His research interests are artificial intelligence, robotics, control, medical rehabilitation and power electroncs.

Dr. Sew Sun Tiang	She is an Assistant Professor in Faculty of Engineering at UCSI University in Malaysia. She received her PhD in Electrical and Electronic Engineering from Universiti Sains Malaysia in 2014. Her research interests are optimization and antenna design.

Dr. Wei Hong Lim	He is an Associate Professor in Faculty of Engineering at UCSI University in Malaysia. He received his PhD in Computational Intelligence from Universiti Sains Malaysia in 2014. His research interests are optimization and artificial intelligence.

Enhancing Global Optimization Performance of Arithmetic Optimization Algorithm with a Modified Population Initialization Scheme

Tin Chang Ting[1], Hameedur Rahman[2], Meng Choung Chiong[1], Mohamed Khan Afthab Ahamed Khan[1], Cik Suhana Hassan[1], Farah Adilah Jamaludin[1], Sew Sun Tiang[1*], Wei Hong Lim[1*]

[1]*Faculty of Engineering, Technology and Built Environment, UCSI University, Kuala Lumpur 56000, Malaysia*
[2]*Faculty of Computing and Artificial Intelligence, Air University, Islamabad Capital Territory 44000, Pakistan*
Email: 1002058096@ucsiuniversity.edu.my, rhameedur@gmail.com, ChiongMC@ucsiuniversity.edu.my, mohamedkhan@ucsiuniversity.edu.my, suhana@ucsiuniversity.edu.my, FarahAdilah@ucsiuniversity.edu.my, tiangss@ucsiuniversity.edu.my, limwh@ucsiuniversity.edu.my

Abstract

Arithmetic Optimization Algorithm (AOA) is widely used to solve global optimization problems. However, it often faces premature convergence challenges in complex optimization scenarios. A key factor affecting AOA's performance is the solution quality of the initial population. The conventional initialization scheme, despite its prevalence, lacks reliability in ensuring high-quality solutions due to inherent stochastic processes. To address this issue, we propose a modified initialization scheme that improves initial population quality by integrating chaotic maps and oppositional-based learning. Through extensive simulation studies, we demonstrate that the enhanced AOA, equipped with this new initialization scheme, exhibits superior performance in solving a range of benchmark functions with improved accuracy.

Keywords: Arithmetic optimization algorithm, Population initialization, Chaotic map, Oppositional-based learning

1. Introduction

Optimization is pivotal in real-life engineering design, where it seeks the most effective solutions while accommodating various stakeholders' criteria. Accurate problem definition and modeling are essential for resolving these design challenges. This process includes establishing clear objectives, identifying both technical and non-technical constraints, and optimizing decision variables. Nonetheless, real-world engineering design problems frequently present complexity and challenges due to factors like high dimensionality, numerous constraints, conflicting objectives, and data uncertainty.

In the era of Industrial Revolution 4.0 (IR 4.0), engineering systems have become more intricate, often involving non-differentiable, nonlinear, multimodal, and non-continuous functions. While traditional optimization methods such as Newton's method are prevalent, they face limitations in addressing these complex, real-world engineering design challenges. These methods often depend heavily on initial solutions and are typically suited to specific problem types, limiting their scalability for diverse, complicated optimization tasks [1]. Most traditional methods also rely on gradient information to identify optimal solutions, which is not always feasible in real-world scenarios that frequently present as black box functions [2]. Furthermore, the limited global search capability of traditional optimization methods increases the risk of premature convergence.

Recognizing the limitations of traditional optimization methods, there is an urgent need in the Industrial Revolution 4.0 era to develop more intelligent and robust optimization algorithms. These algorithms must be capable of providing efficient solutions to a broad spectrum of increasingly complex optimization problems. Metaheuristic Search Algorithms (MSAs) have emerged as effective solutions, drawing on search mechanisms inspired by natural phenomena. Existing MSAs fall into four categories based on their natural inspirations [3]: evolutionary algorithms, swarm intelligence, human-based algorithms, and physics-based algorithms. Compared to traditional methods, MSAs offer several advantages, including potent global search capabilities, straightforward implementation, and enhanced scalability. They exploit the unique strengths of their respective inspirations, enabling them to effectively address a variety of complex optimization challenges as outlined in [4], [5], [6], [7], [8], [9], [10], [11], [12], [13].

The Arithmetic Optimization Algorithm (AOA) [14], introduced in 2021, is a physics-based algorithm with search mechanisms inspired by the four basic arithmetic operations: division, multiplication, addition, and subtraction. These operations vary in their exploration and exploitation strengths, essential for solving optimization problems. Like other MSAs, achieving a proper balance between exploration and exploitation is crucial for enhancing AOA's performance. Consequently, several enhancement schemes, including hybridization

Tin Chang Ting, Hameedur Rahman, Meng Choung Chiong, Mohamed Khan Afthab Ahamed Khan, Cik Suhana Hassan, Farah Adilah Jamaludin, Sew Sun Tiang, Wei Hong Lim

[15], [16], [17], the introduction of new learning mechanisms [18], [19], [20] and others, have been developed to improve the balance in AOA variants. Since its introduction, both the original AOA and its enhanced variants have been successfully applied to a wide array of engineering optimization problems [15], [16], [19], [21], [22], [23], [24], [25].

The quality of the initial population is a key factor in determining the performance of MSAs. Despite numerous variants of the AOA being introduced in recent years, many still rely on conventional methods to randomly generate their initial populations. This conventional approach, while simple to implement, does not intelligently leverage environmental information around the solution regions during initialization [3]. As a result, it often mistakenly places some solutions in local or suboptimal regions, leading to premature convergence. Additionally, if initial solutions are far from the global optimum, the algorithm's convergence speed may be compromised. These drawbacks can significantly impact the robustness and effectiveness of AOA in addressing complex real-world optimization challenges.

This paper presents the Multi-Chaotic Dynamic Oppositional Learning (MCDOL) module as an enhancement to the AOA's population initialization scheme, culminating in a new variant named MCDOL-AOA. The MDCOL module is designed to produce an initial population with enhanced solution quality, both in terms of accuracy and convergence rate, by harnessing the strengths of multiple chaotic maps and the dynamic oppositional-based learning (DOL) mechanism. Notably, the diverse ergodic properties of various chaotic maps are employed for a comprehensive solution space search, enhancing population diversity and mitigating the risk of initializing solutions in local optima. Additionally, the inclusion of DOL in MCDOL module promotes more rapid global optimum identification by expanding the exploration of solution regions. The effectiveness of MCDOL-AOA is evaluated through benchmarking against the original AOA across a range of benchmark functions with varying characteristics.

2. Methodology
2.1. *Proposed MCDOL module*

In this subsection, we detail how the proposed MCDOL module generates an initial population of superior quality. The MCDOL module leverages multiple chaotic maps to create a diverse set of initial solutions, forming a chaotic population. This approach contrasts with conventional initialization schemes, which often result in poorly distributed initial solutions. The deterministic nature of chaotic maps, coupled with their ability to exhibit stochastic behavior through ergodic properties, allows for a more comprehensive exploration of the solution space. Unlike some prior studies [26], [27], [28] that rely on a single chaotic map, this paper explores the benefits of using multiple chaotic maps for population initialization. We hypothesize that each map's unique ergodic characteristics can enhance the algorithm's performance

for specific optimization problems. Therefore, the synergistic integration of multiple chaotic maps promises to significantly improve the algorithm's robustness in addressing various complex problems.

The MCDOL module incorporates five distinct chaotic maps — Circle, Logistic, Piecewise, Sine, and Tent — for generating a chaotic population. Each chaotic map is randomly chosen to generate different dimensions of the initial solutions within this population. To facilitate this, define ϑ_t as the output of a chaotic variable at the t-th iteration, where $t = 1, \ldots, T$. The population initialization in the MCDOL module, using these multiple chaotic maps, follows five criteria: (a) the Circle map is used if $0 \leq \vartheta_0 < 0.2$, (b) the Logistic map for $0.2 \leq \vartheta_0 < 0.4$, (c) the Piecewise map for $0.4 \leq \vartheta_0 < 0.6$, (d) the Sine map for $0.6 \leq \vartheta_0 < 0.8$, and (e) the Tent map for $0.8 \leq \vartheta_0 \leq 1.0$. The specific equations for these five chaotic maps, Circle, Logistic, Piecewise, Sine, and Tent, are detailed in Eqs. (1) to (5).

$$\vartheta_{t+1} = \mathrm{mod}\left(\vartheta_t + 0.2 - \left(\frac{0.5}{2\pi}\right)\sin(2\pi\vartheta_t), 1\right) \quad (1)$$

$$\vartheta_{t+1} = 4\,\vartheta_t(1 - \vartheta_t) \quad (2)$$

$$\vartheta_{t+1} = \begin{cases} \dfrac{\vartheta_t}{P}, & 0 \leq \vartheta_t < P \\ \dfrac{\vartheta_t - P}{0.5 - P}, & P \leq \vartheta_t < 0.5 \\ \dfrac{1 - P - \vartheta_t}{0.5 - P}, & 0.5 \leq \vartheta_t < 1 - P \\ \dfrac{1 - \vartheta_t}{P}, & 1 - P \leq \vartheta_t < 1 \end{cases} \quad (3)$$

$$\vartheta_{t+1} = \sin(\pi\vartheta_t) \quad (4)$$

$$\vartheta_{t+1} = \begin{cases} \dfrac{\vartheta_t}{0.7} & \vartheta_t < 0.7 \\ \dfrac{10}{3}(1 - \vartheta_t) & \vartheta_t \geq 0.7 \end{cases} \quad (5)$$

Let X_d^L and X_d^U denote the lower and upper bounds of the d-th dimensional decision variable, respectively, for $d = 1, \ldots, D$. At the final iteration $t = T$, a chaotic value ϑ_T is generated from a randomly selected chaotic map. This value initializes the d-th dimension of each n-th chaotic solution, as outlined in Eq. (6). The resulting chaotic population, denoted as $\mathbf{P^C} = \left[X_1^C, \ldots, X_n^C, \ldots X_N^C\right]$, encompasses all solution members formed using multiple chaotic maps.

$$X_{n,d}^C = X_d^L + \vartheta_T (X_d^U - X_d^L) \quad (6)$$

While various chaotic maps exhibit differing levels of robustness to local optima, they may still generate initial solutions distant from the global optimum, potentially slowing the algorithm's convergence. To address this issue, a DOL operator is applied to the chaotic population $\mathbf{P^C}$, generating opposite solutions for each n-th chaotic solution. This DOL operator is expected to broaden the initial population's coverage of the solution space,

increasing the likelihood of identifying fitter solutions. The opposite solution for each dimension, denoted as $X_{n,d}^O$, is determined using Eq. (7), corresponding to $X_{n,d}^C$, with $r_1, r_2 \in [0, 1]$. Consequently, this generates an opposition population, $\mathbf{P^O} = [X_1^O, \dots, X_n^O, \dots X_N^O]$.

$$X_{n,d}^O = X_{n,d}^C + r_1 \left[r_2(X_d^U + X_d^L - X_{n,d}^C) - X_{n,d}^C \right] \quad (7)$$

The two populations, one generated using multiple chaotic maps ($\mathbf{P^C}$) and the other using the DOL operator ($\mathbf{P^O}$), are merged to form a combined population set $\mathbf{P^C} \cup \mathbf{P^O}$ with a total size of $2N$. The fitness of each solution member in this merged population set is evaluated according to the predefined objective function. These solution members are then ordered based on their fitness values, from the best to worst performing ones. The top N solution members from this ordered population set $\mathbf{P^C} \cup \mathbf{P^O}$ are chosen as the initial population for the proposed MCDOL-AOA algorithm, represented as $\mathbf{P} = [X_1, \dots, X_n, \dots, X_N]$.

2.2. *Iterative Search Processes of MCDOL-AOA*

After generating an initial population $\mathbf{P}$ of superior quality using the MCDOL module, each n-th solution in MCDOL-AOA is iteratively updated through search mechanisms similar to those in the original AOA.

During each iteration, the Math Optimizer Accelerated (MOA) function value is adjusted in MCDOL-AOA to alternate between exploration and exploitation phases:

$$MOA(C_{Iter}) = Min + C_{Iter} \left(\frac{Max - Min}{M_{Iter}} \right) \quad (8)$$

where C_{Iter} and M_{Iter} represent the current and maximum iteration counts; Min and Max are the minimum and maximum values of MOA. Simultaneously, the Math Optimizer Probability (MP) function, which dictates the search range for each solution and is influenced by the critical parameter θ for exploitation efficiency, is updated as:

$$MP(C_{Iter}) = 1 - \left(\frac{C_{Iter}}{M_{Iter}} \right)^{1/\theta} \quad (9)$$

A random number $\mathbf{rand1}$ determines the search strategy (exploration or exploitation) at each iteration for updating the d-th dimension of the n-th solution, $X_{n,d}$. During the exploration phase ($\mathbf{rand1} > \mathbf{MOA}$), either the Multiplication or Division operator is selected:

$$
X_{n,d}(C_{Iter} + 1)
$$
$$
= \begin{cases} best_d \div (MP + \varepsilon) \times [(X_d^U - X_d^L)\mu + X_d^L], rand2 < 0.5 \\ best_d \times MP \times [(X_d^U - X_d^L)\mu + X_d^L], \quad \text{Otherwise} \end{cases}
$$
$$(10)$$

where $rand2$ is a random number between 0 and 1; $best_d$ denotes the d-th dimension of the current best solution; ε is a small positive number to prevent division by zero, and μ is a control parameter.

In the exploitation phase ($rand1 \le MOA$), either the Addition or Subtraction operator is used to update $X_{n,d}$:

$$
X_{n,d}(C_{Iter} + 1) =
$$
$$
\begin{cases} best_d - MP \times [(X_d^U - X_d^L)\mu + X_d^L], rand2 < 0.5 \\ best_d + MP \times [(X_d^U - X_d^L)\mu + X_d^L], \quad \text{Otherwise} \end{cases} \quad (11)
$$

As illustrated in Fig. 1, MCDOL-AOA continues this iterative search, following Eqs. (8) to (11), until pre-set termination criteria are fulfilled. Upon completion, optimal decision variables in the best solution are decoded to solve the specific optimization problems.

MCDOL-AOA for Global Optimization
Inputs: $D, N, M_{Iter}, T, Max, Min, \theta$
01: Initialize $\mathbf{P^C} \leftarrow \emptyset, \mathbf{P^O} \leftarrow \emptyset$ and $C_{Iter} \leftarrow 0$;
02: **for** each n-th solution **do**
03: **for** each d-th dimension **do**
04: Randomly initialize $\vartheta_t \in [0,1]$, where $t = 0$;
05: **if** $0 \le \vartheta_0 < 0.2$ **then**
06: Select Circle map in Eq. (1);
07: **elseif** $0.2 \le \vartheta_0 < 0.4$ **then**
08: Select Logistic map in Eq. (2);
09: **elseif** $0.4 \le \vartheta_0 < 0.6$ **then**
10: Select Piecewise map in Eq. (3);
11: **elseif** $0.6 \le \vartheta_0 < 0.8$ **then**
12: Select Sine map in Eq. (4);
13: **elseif** $0.8 \le \vartheta_0 \le 1.0$ **then**
14: Select Tent map in Eq. (5);
15: **end if**
16: **while** $t \le T$ **do**
17: Update ϑ_t with the selected chaotic map;
18: $t \leftarrow t + 1$;
19: **end while**
20: Generate $X_{n,d}^C$ using Eq. (6);
21: Generate $X_{n,d}^O$ using Eq. (7);
22: **end for**
23: Update $\mathbf{P^C} \leftarrow \mathbf{P^C} \cup X_{n,d}^C$ and $\mathbf{P^O} \leftarrow \mathbf{P^O} \cup X_{n,d}^O$;
24: **end for**
25: Merge the two population sets as $\mathbf{P^C} \cup \mathbf{P^O}$;
26: Evaluate the fitness values of all solutions stored within $\mathbf{P^C} \cup \mathbf{P^O}$;
27: Rearrange the solutions stored within $\mathbf{P^C} \cup \mathbf{P^O}$ from best to worst based on their fitness values;
28: Select the top N solutions from the sorted $\mathbf{P^C} \cup \mathbf{P^O}$ as the initial population, i.e., $\mathbf{P} = [X_1, \dots, X_n, \dots, X_N]$.
29: Assign the first solution of $\mathbf{P}$ and its fitness as $best$ and $f(best)$, respectively;
30: **while** $C_{Iter} \le M_{Iter}$ **do**
31: Update MOA and MP with Eqs. (8) and (9);
32: **for** each n-th solution **do**
33: **if** $rand1 > MOA$ **then** /*Exploration*/
34: Update $X_{n,d}(C_{Iter} + 1)$ with Eq. (10);
35: **else** /*Exploitation*/
36: Update $X_{n,d}(C_{Iter} + 1)$ with Eq. (11);
37: **end if**
38: Fitness evaluation of $X_n(C_{Iter} + 1)$;
39: Update the $X_n, f(X_n)$, $best$ and $f(best)$,
40: with greedy selection method;
41: **end for**
42: $C_{Iter} \leftarrow C_{Iter} + 1$;
43: **end while**
Output: $best$ and $f(best)$

Fig.1 Workflow of proposed MCDOL-AOA in solving the global optimization problems.

Tin Chang Ting, Hameedur Rahman, Meng Choung Chiong, Mohamed Khan Afthab Ahamed Khan, Cik Suhana Hassan, Farah Adilah Jamaludin, Sew Sun Tiang, Wei Hong Lim

3. Results and Discussions

3.1. *Simulations settings*

In this section, we compare the performance of MCDOL-AOA with the original AOA using 23 benchmark functions, each with distinct characteristics as outlined in [17]. Functions F1 to F7 are scalable unimodal functions, while F8 to F13 are scalable multimodal functions, all set at a dimension size of $D = 100$. Functions F14 to F23, on the other hand, are fixed-dimension multimodal functions with D ranging from 2 to 6. Both MCDOL-AOA and the original AOA were implemented on MATLAB 2021a, running on a personal computer equipped with an Intel® Core™ i7-HQ CPU at 2.50 GHz and 16 GB RAM. For both algorithms, the population size and maximum number of iterations are set at 30 and 1000, respectively.

3.2. *Performance analysis*

The evaluation of MCDOL-AOA's performance, and its comparison with the original AOA across all 23 benchmark functions, is detailed in Table 1. We employ two performance metrics: mean error (E_{mean}) and standard deviation (SD), to assess the algorithms' accuracy and consistency in solving these functions. Lower values of E_{mean} and SD are preferable, indicating the algorithm's consistent and accurate resolution of the given benchmark functions.

Table 1. Performance comparison of MCDOL-AOA and original AOA using 23 benchmark functions with different characteristics.

Fun	Original AOA		MCDOL-AOA	
	E_{mean}	SD	E_{mean}	SD
F1	2.048e-04	5.662e-05	**1.938e-04**	5.371e-05
F2	0.0112	0.0115	**0.0107**	0.0011
F3	**0.0915**	0.0229	0.0934	0.0202
F4	0.0523	0.0076	**0.0516**	0.0089
F5	98.090	0.0899	**98.060**	0.1158
F6	15.131	0.7747	**14.930**	0.8160
F7	**2.609e-05**	3.060e-05	3.574e-05	3.667e-05
F8	**-1.53e+04**	646.631	**-1.53e+04**	631.509
F9	6.244e-05	1.478e-05	**6.195e-05**	1.080e-05
F10	**0.0016**	1.452e-04	**0.0016**	1.767e-04
F11	**0.0085**	0.0249	0.0176	0.0519
F12	0.891	0.0655	**0.0888**	0.0740
F13	9.906	0.0030	**9.904**	0.0030
F14	**9.864**	4.317	11.192	**2.979**
F15	0.0055	0.0125	**0.0028**	0.0075
F16	**-1.032**	**5.375e-12**	**-1.032**	8.016e-12
F17	**0.398**	5.620e-07	**0.398**	**3.516e-07**
F18	18.3	28.080	**16.5**	25.318
F19	**-3.863**	5.441e-05	**-3.863**	**4.105e-06**
F20	**-3.294**	**0.0511**	-3.274	0.0592
F21	**-8.554**	**2.522**	-8.141	2.979
F22	-8.077	3.155	**-8.350**	3.267
F23	-8.590	3.094	**-8.751**	3.091

The results in Table 1 showcase MCDOL-AOA's impressive accuracy, outperforming or matching the original AOA in 16 of the 23 benchmark functions, as

evidenced by lower E_{mean} values. In the scalable unimodal function's category (F1 to F7), MCDOL-AOA surpasses the original AOA in 5 of the 7 functions based on E_{mean}. For the 6 scalable multimodal functions (F8 to F13), MCDOL-AOA demonstrates superior E_{mean} values in 4 functions (F8, F9, F12, and F13) and equals the original AOA in one function (F10). When dealing with the fixed-dimension multimodal functions (F14 to F23), MCDOL-AOA outperforms the original AOA in 4 functions (F15, F18, F22, and F23) and equals its performance in 3 other functions (F16, F17, and F19).

Analysis of the simulation results in Table 1 indicates that MCDOL-AOA significantly outperforms the original AOA in solving larger-scale benchmark functions (F1 to F13) with $D = 100$, evidenced by lower E_{mean} values in 9 out of 13 functions. However, the performance of MCDOL-AOA is relatively on par with the original AOA in fixed-dimension problems (F14 to F23) with smaller dimensions ($D = 2, 3, 4$, and 6). These findings suggest that the MCDOL module in MCDOL-AOA effectively produces initial populations with superior fitness and diversity compared to the original AOA, yielding enhanced optimization results. The benefits of the MCDOL module are particularly notable in larger dimensional sizes (e.g., $D = 100$), as utilized in this study. The multiple chaotic maps' non-repetitive and ergodic properties within the MCDOL module facilitate a more exhaustive exploration of the solution space across various problem types, thereby minimizing the risk of local optima entrapment and premature convergence. Additionally, the DOL mechanism within the MCDOL module contributes to faster algorithmic convergence by broadening solution space exploration through the generation of opposite solutions from those initiated by the chaotic maps.

4. Conclusion

In this paper, we introduce an enhanced version of the AOA, termed MCDOL-AOA, designed to solve complex global optimization problems with improved accuracy. The novelty of MCDOL-AOA resides in its integration of multiple chaotic maps and DOL mechanisms into a modified initialization scheme, the MCDOL module. This module aims to generate an initial population of superior quality, focusing on fitness and diversity. Simulation results demonstrate that MCDOL-AOA, benefiting from the enhanced initial population quality provided by the MCDOL module, surpasses the original AOA in solving 23 distinct benchmark functions. Notably, the performance gains of MCDOL-AOA are more pronounced in solving functions with larger dimensions, specifically $D = 100$. This underscores the efficacy of the multiple chaotic maps and DOL mechanisms within the MCDOL module in boosting the algorithm's robustness against premature convergence and enhancing its convergence speed. Future work will explore the application of MCDOL-AOA to real-world engineering optimization problems, including machine

learning model training and scheduling optimization, to assess its practicality and effectiveness.

Acknowledgements

References

1. J. C. Nash, On best practice optimization methods in R, Journal of Statistical Software 60(2), 2014, pp. 1-14.
2. B. Petrongolo and A. P. Chirstopher, Looking into the black box: A survey of the matching function, Journal of Economic Literature 39(2), 2001, pp. 390-431.
3. M. F. Ahamad, N. A. Mat Isa, W. H. Lim and K. M. Ang, Differential evolution with modified initialization scheme using chaotic oppositional based learning strategy, Alexandria Engineering Journal 61(12), 2022, pp. 11835-11858.
4. E. S. Ghith and F. A. A. Tolba, Real-time implementation of an enhanced proportional-integral-derivative controller based on sparrow search algorithm for micro-robotics system, IAES International Journal of Artificial Intelligence 11(4), 2022, pp. 1395-1404.
5. E. S. Ghith and F. A. A. Tolba, LabVIEW implementation of tuning PID controller using advanced control optimization techniques for micro-robotics system. International Journal of Mechanical Engineering and Robotics Research 11(9), 2022, pp. 653-661.
6. E. S. Ghith and F. A. Tolba, Real-time implementation of an enhanced PID controller based on marine predator algorithm (MPA) for micro-robotics system, 2022 3rd International Conference on Artificial Intelligence, Robotics and Control (AIRC), Cairo, Egypt, 2022.
7. C. Hassan, S. Sapuan, N. Abd. Aziz and M. Z. Mohamed Yusof, Effect of chemical treatment on the tensile properties of single oil palm empty fruit bunch (OPEFB) fibre. Trends in Textile Engineering & Fashion Technology 3(2), 2018, pp. 1-7.
8. E. Natarajan, C. Hassan, C. K. Ang, M. S. Santhosh, S. Ramesh and R. Sasikumar, Modeling of multiwall carbon nanotubes reinforced natural rubber for soft robotic applications - A comprehensive presentation 46(9), 2021, pp. 3251-3258.
9. C. Hassan C. W. Yeo, B. Shari, M. S. Salit and N. Abdul Aziz, Mechanical properties of unidirectional oil palm empty fruit bunch (OPEFB) fiber reinforced epoxy composite. IOP Conference Series: Materials Science and Engineering 206(1), 2017, pp 012045.
10. L. Yao et al., Demand bidding optimization for an aggregator with genetic algorithm. Energies 11(1), 2018, pp. 1-22.
11. K. M. Ang et al., Modified teaching-learning-based optimization and applications in multi-response machining processes, Computers & Industrial Engineering 174, 2022, pp. 108719.
12. A. Sharma et al., Identification of photovoltaic module parameters by implementing a novel teaching learning based optimization with unique exemplar generation scheme (TLBO-UEGS), Energy Report 10, 2023, pp. 1485-1506.
13. A. Sharma et al., An effective method for parameter estimation of solar PV cell using grey-wolf optimization technique, International Journal of Mathematical, Engineering and Management Science 6(3), 2021, pp. 911-931.
14. L. Abualigah, A. Diabat, S. Mirjalili, M. Abd Elaziz and A. H. Gandomi, The arithmetic optimization algorithm, Computer Methods in Applied Mechanics and Engineering 376, 2021, pp. 113609.
15. A. A. Mohamed et al., A novel hybrid arithmetic optimization algorithm and salp swarm algorithm for data placement in cloud computing, Soft Computing, 27(9), 2023, pp. 5769-5780.
16. M. Issa, Enhanced arithmetic optimization algorithm for parameter estimation of PID controller, Arabian Journal for Science and Engineering, 48(2), 2023, pp. 2191-2205.
17. L. Abualigah and A. Diabat, Improved multi-core arithmetic optimization algorithm-based ensemble mutation for multidisciplinary applications, Journal of Intelligent Manufacturing 34(4), 2023, pp. 1833-1874.
18. S. B. Aydemir, A novel arithmetic optimization algorithm based on chaotic maps for global optimization, Evolutionary Intelligence 16(3), 2023, pp. 981-996.
19. W.-K. Hao, J.-S. Wang, X.-D. Li, H.-M. Song and Y.-Y. Bao, Probability distribution arithmetic optimization algorithm based on variable order penalty functions to solve combined economic emission dispatch problem, Applied Energy, 316, 2022, pp. 119061.
20. Y.-J. Zhang, Y.-F. Wang, Y.-X. Yan, J. Zhao and Z.-M. Gao, LMRAOA: An improved arithmetic optimization algorithm with multi-leader and high-speed jumping based on opposition-based learning solving engineering and numerical problems, Alexandria Engineering Journal 61(12), 2022, pp. 12367-12403.
21. E. S. Ghith and F. A. A. Tolba, Tuning PID controllers based on hybrid arithmetic optimization algorithm and artificial gorilla troop optimization for micro-robotic systems. IEEE Access 11, 2023, pp. 27138-27154.
22. S. Rajput et al., A novel opposition-based arithmetic optimization algorithm for parameter extraction of PEM fuel cell. Electronics 10(2), 2021, pp. 2834.
23. K. M. Ang et al., Training feedforward neural networks using arithmetic optimization algorithm for medical classification, in Advances in Intelligent Manufacturing and Mechatronics. Lecture Notes in Electrical Engineering, eds. M. A. Abdullah et al. Springer, Singapore, vol. 988, 2023.
24. X. Shi, X. Yu, and M. Esmaeili-Falak, Improved arithmetic optimization algorithm and its application to carbon fiber reinforced polymer-steel bond strength estimation, Composite Structures 306, 2023, pp. 116599.
25. A. Sharma et al., Performance investigation of state-of-the-art metaheuristic techniques for parameter extraction of solar cells/module, Scientific Reports 13, 2023, pp. 11134.
26. Y. Wang, S. Gao, Y. Yu, Z. Wang, J. Cheng and T. Yuki, A gravitational search algorithm with chaotic neural oscillators, IEEE Access 8, 2020, pp. 25938-25948.
27. Z. Xu, S. Gao, H. Yang and Z. Lei, SCJADE: Yet another state-of-the-art differential evolution algorithm, IEEJ Transactions on Electrical and Electronic Engineering 16(4), 2021, pp. 644-646.
28. A.B. Ozer, CIDE: chaotically initialized differential evolution, Expert Systems with Applications 37(6), 2010, pp. 4632-4641.

Tin Chang Ting, Hameedur Rahman, Meng Choung Chiong, Mohamed Khan Afthab Ahamed Khan, Cik Suhana Hassan, Farah Adilah Jamaludin, Sew Sun Tiang, Wei Hong Lim

Authors Introduction

Mr. Tin Chang Ting

He is currently pursuing the Bachelor of Mechatronics Engineering with Honours in Faculty of Engineering, Technology and Built Environment, UCSI University, Malaysia. His research interests are machine learning, deep learning, and optimization algorithm.

Dr. Hameedur Rahman

He is an Associate Professor in Faculty of Computing and AI at AIR University in Pakistan. He received his PhD in Computer Science from Universiti Kebangsaan Malaysia in 2018. His research interests are Virual/Augmented Rality, Image Processing, Data Mining, Artificial Intelligence, Natural Language Processing and CyberSecurity.

Dr. Meng Choung Chiong

He is an Assistant Professor in Faculty of Engineering at UCSI University in Malaysia. He received his PhD in Chemical Enigneering from Universiti Teknologi Malaysia. His research interests are artificial intelligence, mechatronics and alternative fuels combustion.

Dr. Mohamed Khan Afthab Ahmed Khan

He is an Assistant Professor in Faculty of Engineering at UCSI University in Malaysia. His research interests are artificial intelligence, robotics, control, medical rehabilitation and power electroncs.

Dr. Cik Suhana Hassan

She is an Assistant Professor in Faculty of Engineering at UCSI University in Malaysia. She obtained her PhD in PhD in Mechanical Engineering from Universiti Putra Malaysia. Her research interests are artifical intelligence, bio composite and finite element analysis.

Dr. Farah Adilah Binti Jamaluddin

She is an Assistant Professor in Faculty of Engineering at UCSI University in Malaysia. She obtained her PhD in PhD in Eletrical Power Engineering from Universiti Putra Malaysia. Her research interests are artifical intelligence, electrical insulation and high voltage engineering.

Dr. Sew Sun Tiang

She is an Assistant Professor in Faculty of Engineering at UCSI University in Malaysia. She received her PhD in Electrical and Electronic Engineering from Universiti Sains Malaysia in 2014. Her research interests are optimization and antenna design.

Dr. Wei Hong Lim

He is an Associate Professor in Faculty of Engineering at UCSI University in Malaysia. He received his PhD in Computational Intelligence from Universiti Sains Malaysia in 2014. His research interests are optimization and artificial intelligence.

Enhancing Precision Object Detection and Identification for Autonomous Vehicles through YOLOv5 Refinement with YOLO-ALPHA

Guandong Li[1], Yanzhe Xie[1], Yuhao Lu[1], Zongyan Wen[1], Jingzhen Fan[1], Yuankui Huang[1], Qinghong Ma[1], Wei Hong Lim[2], Chin Hong Wong[1*]

[1]*Maynooth International Engineering College, Fuzhou University, Fujian, China*
[2]*Faculty of Engineering, Technology and Built Environment, UCSI University, 1, Jalan Puncak Menara Gading, UCSI Heights, 56000 Cheras, Kuala Lumpur, Malaysia.*
E-mail: guandong.li.2022@mumail.ie, yanzhe.xie.2022@mumail.ie, yuhao.lu.2022@mumail.ie, zongyan.wen.2022@mumail.ie, jingzhen.fan.2022@mumail.ie, yuankui.huang.2022@mumail.ie, qinghong.2021@mumail.ie,, limwh@ucsiuniversity.edu.my, chinhong.wong@mu.ie

Abstract

Advancing swiftly in contemporary society, the rapid growth of autonomous driving technology suggests its potential adoption across continents. The realization of fully autonomous driving relies on proficiently detecting, classifying, and tracking road objects such as pedestrians and vehicles. This research employs the YOLOv5 neural network, enhancing it with YOLO-ALPHA. Modifications, encompassing freeze and attention mechanisms, serve to refine accuracy and expedite training. Furthermore, adjustments to the activation function aim to stabilize precision and recall. The integration of an FCN based on semantic segmentation theory contributes to improved accuracy in detecting road conditions during autonomous driving. Consequently, this enables the successful and highly accurate functionality of automatic identification.

Keywords: Object detection, Autonomous vehicle, YOLOv5, FCN, Attention mechanism, Freeze mechanism, Activation function

1. Introduction

In recent years, the rapid advancement of technology has led to a surge in the popularity of autonomous driving systems. Real-time image processing is a pivotal technology in autonomous driving. Despite reaching an advanced stage, autonomous driving technology faces persistent challenges in low recognition accuracy and sluggish real-time image processing, giving rise to safety concerns. These issues give rise to safety concerns. Although automated driving technology has reached a certain level of maturity, it encounters limitations in adverse weather conditions, complex environments, and nighttime operations. automatic driving sensor recognition accuracy tends to decline, compromising autonomous vehicle safety. Consequently, there is a pressing need to enhance automatic driving technology to bolster autonomous vehicle reliability and safety.

This paper introduces an attention mechanism into the object detection algorithm, fine-tunes activation functions, adds a freezing mechanism and conducts a comparative analysis with the FCN (Fully Convolutional Networks) algorithm. This algorithm offers an enhanced detection speed and accuracy. The implications of this research are highly relevant to the enhancement of autonomous driving safety.

Utilizing the PASCAL VOC (The PASCAL Visual Object Classes) dataset, this study systematically alters various modules and introduces distinct mechanisms, this paper changes a different module and adds a disparate mechanism each time. Then, by evaluating precision, recall, mAP, and mAP0.5 indices of the adapted models, it identifies the optimal placement for incorporating an attention mechanism, freezing layers to prolong training duration, and selecting the activation function that demonstrated superior performance on the PASCAL VOC dataset.

This paper presents a novel detection technique named YOLO-ALPHA. The approach involves incorporating an attention mechanism at various layers within the YOLO network. Specifically, inserting the Squeeze-and-Excitation (SE) attention mechanism (Self-Attention Mechanism) after the Concat module on the 16th layer of the network could enhance both the precision and recall of the model. Additionally, adding a freeze mechanism in the YOLO backbone could accelerate curve fitting. Moreover, replacing the combined structure of SiLU (Sigmoid-weighted Linear Units) and LeakyReLU with the SiLU activation function contributes to a smoother precision and recall curve, thereby, improving the stability of the model by 0.169%.

2. Literature Review

Guandong Li, Yanzhe Xie, Yuhao Lu, Zongyan Wen, Jingzhen Fan, Yuankui Huang, Qinghong Ma, Wei Hong Lim, Chin Hong Wong

Convolutional neural networks (CNNs) often suffer from information loss, resulting in gradient disappearance and detection inaccuracies during image processing. Researchers continuously strive to mitigate semantic information loss and enhance overall network performance. Lan et al. [1] conducted research to improve the YOLO network, focusing on addressing losses in pedestrian information during image processing. They introduced additional passthrough layers, including a Route layer and Reorg layer, or reconstructed the primitive YOLOv2 network. The resulting YOLO-R network specializes in pedestrian detection, effectively improving accuracy and reducing false detection rates. Wu et al. [2] presented a vehicle detection system in CARLA, introducing YOLOv5-Ghost with adjusted layer structures for reduced computational complexity. The study achieved improved detection accuracy and speed. Lu et al. [3] enhanced a YOLOv5-based model for crack and vehicle detection datasets, addressing sample imbalance and small object presence. Dodia and Kumar [4] compared modern object detectors with YOLO for vehicle detection, suggesting enhancements to improve vehicle prediction capacity. Kaloev et al. [5] investigated activation functions in deep learning, emphasizing their role in introducing nonlinearity. Xiao et al. [6] proposed intelligent layer freezing during training to accelerate training on various networks. Zhao et al. [7] improved an infrared detection model based on YOLOv5s, adding an SE-Net module. Kaymak et al. [8] experimented with FCN architectures in the context of autonomous driving. The study concludes that applying YOLO to vehicle detection faces challenges, leading researchers to improve the algorithm through various modifications.

3. Design Methodology

Fig. 1 shows the flow chart to tackle imminent challenges related to object identification, classification, and tracking in the domain of autonomous driving. Given the effectiveness of YOLOv5 in object identification tasks, as evidenced by Xiao [9] and Kaymak [8], and recognizing the utility of FCN in autonomous driving scenarios, this research chooses to employ both methods for real-time image processing in autonomous driving contexts. To ensure a fair comparison of results across different networks using the control variate method, a consistent dataset, PASCAL VOC 2012, was used to train and validate both YOLO and FCN networks. PASCAL VOC 2012 encompasses 20 classes commonly encountered in autonomous driving scenarios, such as bicycle, bus, car, motorbike, and person.

Following multiple rounds of training and prediction in YOLO, this work collects and scrutinizes detection and classification outcomes using TensorBoard, relying on index curves such as precision, recall, mAP, and mAP0.5 curvature. Higher precision values denote the network's proficiency in accurately classifying items, while elevated recall signifies the network's efficacy in rectifying previous classification errors. The mAP and mAP0.5 metrics evaluate the detection accuracy across

all classes present in the dataset. A lower curvature in the curve indicates enhanced stability and a more consistently changing trend, signifying ideal model training stability and convergence.
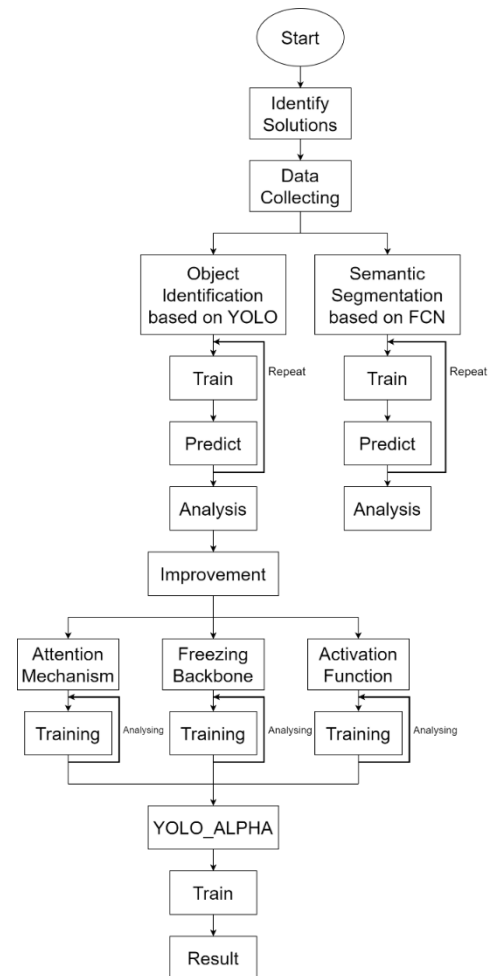
Fig. 1 Flow chart

Then, the project aims to enhance the YOLO algorithm by adding an attention mechanism, freezing backbone, and activation function.

Incorporating an attention mechanism enables the network to focus on crucial input features while disregarding inconsequential ones. Through the integration of SE attention mechanisms in different layers of the YOLO structure and subsequent training, the research evaluates the performance enhancement trend by scrutinizing index curves and determining the optimal location for implementing the SE attention mechanism.

Freezing the backbone is a training strategy aimed at expediting training speed by halting parameters in the backbone upgrade. By comparing the result indexes of modules employing freezing backbone, freezing all layers, and not freezing layers separately, the project seeks to identify the optimal freezing method.

Activation functions amplify the nonlinearity of convolutional neural networks. By replacing the original activation function with SiLU, Mish, LeakyReLU, and ReLU in the modified YOLOv5s, the project utilizes the

average curvature of the precision curve to compare and assess the smoothness and stability of the network.

Ultimately, the study amalgamates the proven optimal strategies from the preceding analysis to formulate YOLO-ALPHA, an enhanced version of YOLOv5. The performance of YOLO-ALPHA is then juxtaposed with that of YOLOv5 to affirm its improvement.

4. Results and Discussions

The object detection output images from YOLOv5s depict vehicles on the road using bounding boxes. However, these bounding boxes lack information regarding the distance between vehicles, a crucial consideration in congested road scenarios, particularly during peak commuting hours, a vital aspect of autonomous driving. To precisely discern the outlines of nearby vehicles and pedestrians, pixel-wise object detection and classification with FCN are employed. Fig. 2 illustrates the contrast in detection results between YOLOv5s and FCN. The depiction from YOLOv5s shows vehicles with pink bounding boxes, outlining the edges of cars as straight lines that do not accurately match the actual boundaries. On the other hand, FCN, with pixel-wise detection, unveils the precise boundaries of cars, enabling autonomous vehicles to effectively avoid collisions.
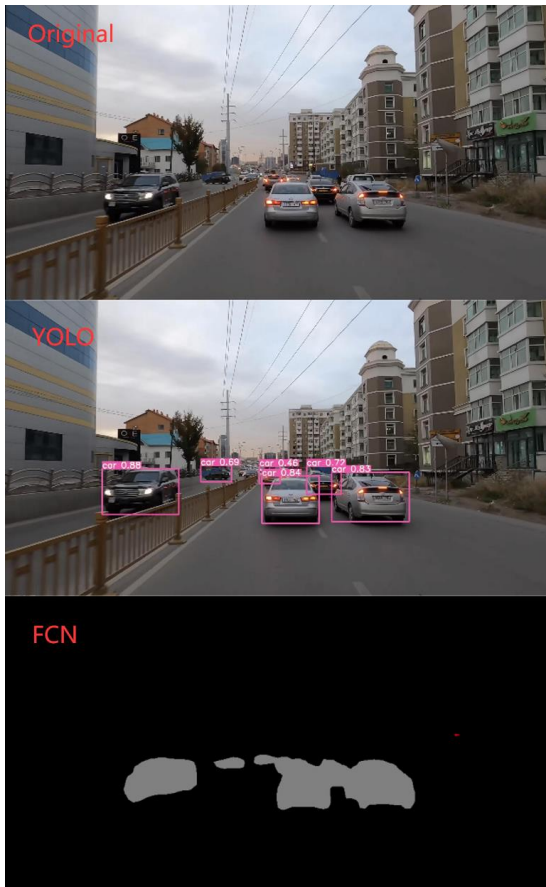
Fig. 2 Detection result between YOLOv5s and FCN

Fig. 3 shows minimal differences observed between the frozen backbone curve and the original model curve in terms of mAP, precision, and recall at the highest training rounds. Despite a reduction in training time by 2 hours, image accuracy slightly decreases from 0.7166 to 0.6716. The key conclusion drawn is that freezing trained parameters enhances training speed without compromising accuracy. Freezing all layers completely leads to virtually no accuracy, emphasizing the importance of trained parameters. Freezing the backbone during training enhances training efficiency, as the unchanging feature extraction network consumes less memory, leading to faster training.
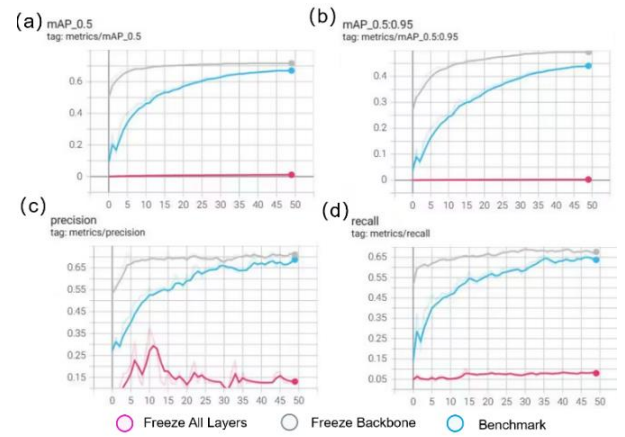
Fig. 2 Freezing Mechanism graphs for (a) mAP_0.5, (b) mAP_0.5:0.95, (c) Precision and (d) Recall

Experimental results suggested that adding the SE attention mechanism to the bottom or middle layers of the network yields optimal outcomes. Fig. 4 illustrates that adding the attention mechanism between layers 16 and 17 achieves higher accuracy than the original version, validating the hypothesis. Fig. 5 further confirms that placing the SE attention mechanism between the middle and bottom layers produces the best results.
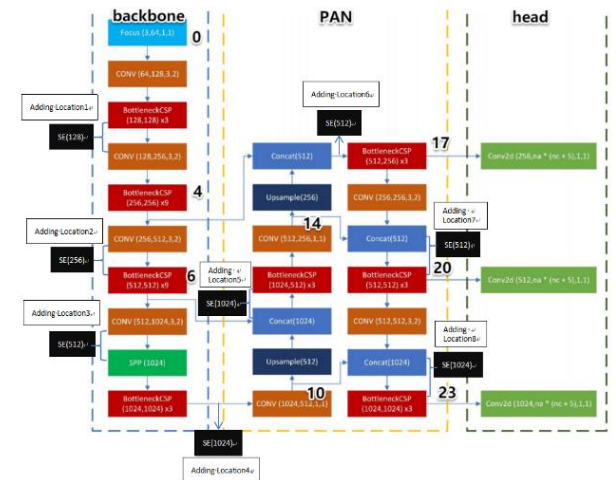
Fig. 3 SE adding locations in the original YOLOv5 network

Guandong Li, Yanzhe Xie, Yuhao Lu, Zongyan Wen, Jingzhen Fan, Yuankui Huang, Qinghong Ma, Wei Hong Lim, Chin Hong Wong
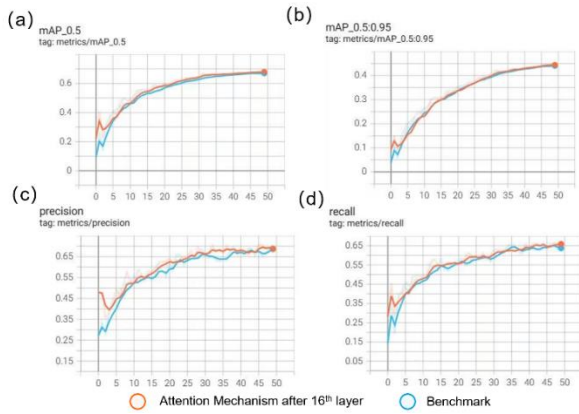
Fig. 4 Attention Mechanism graphs for (a) mAP_0.5, (b) mAP_0.5:0.95, (c) Precision and (d) Recall

Five activation functions (ReLU, Leaky ReLU, Mish, SiLU, and the original activation function) undergo comparative analysis. Fig. 6 displays the precision, recall, mAP, and mAP0.5 indexes at epoch 50 for networks employing various activation functions. The SiLU activation function is selected for the PASCAL VOC dataset due to its smoother curve and increased stability, as tabulated in Table 1.
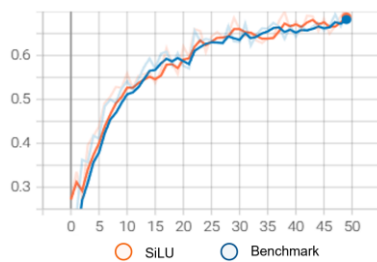
Fig. 5 Activation Function

Table 1 Activation Function Data

Epoch = 50	mAP	mAP0.5	Precision	Recall
benchmark	0.6716	0.4426	0.7064	0.6286
LeakyReLU	0.6437	0.4114	0.6531	0.6238
Mish	0.6649	0.4397	0.6783	0.6363
ReLU	0.6353	0.4058	0.7007	0.5739
SiLU	0.6694	0.4426	0.6961	0.6270

Fig. 7 depicts a less satisfactory outcome when combining the optimal freezing mechanism, attention mechanism, and activation function. The precision initially starts high but continuously decreases, indicating performance degradation. Removing the attention mechanism results in Fig. 8 that surpass those of the original YOLOv5 network, affirming the performance enhancement of the upgraded YOLO-ALPHA network.

5. Conclusion

In this investigation, YOLOv5s is employed to carry out object detection, classification, and tracking, with the FCN (Semantic Segmentation) technique serving as a supplementary method to precisely delineate object

boundaries. Enhancements, including the incorporation of attention mechanisms, freezing the network's backbone, and modifying the activation function, are evaluated based on precision, recall, mAP, and mAP0.5 curves. These identified mechanisms and modules, known to augment the performance of YOLOv5s, are amalgamated into an upgraded version referred to as YOLO-ALPHA. However, the collective optimal modules identified in previous research resulted in a degradation of performance. Following meticulous analysis, the freezing backbone mechanism is pinpointed as the cause. Subsequent refinement by removing the freezing backbone mechanism reveals that YOLO-ALPHA showcases performance improvements in comparison to the original YOLOv5s.
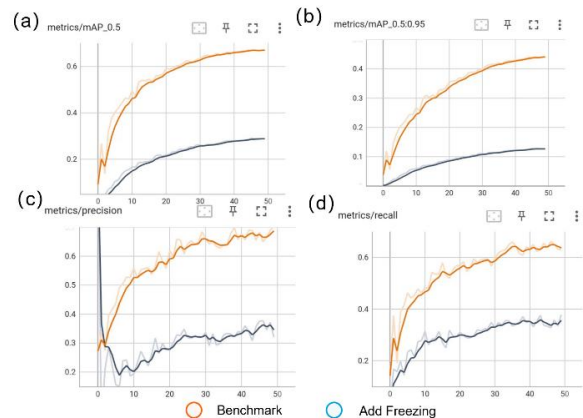
Fig. 6 Unsuccessful Result for (a) mAP_0.5, (b) mAP_0.5:0.95, (c) Precision and (d) Recall
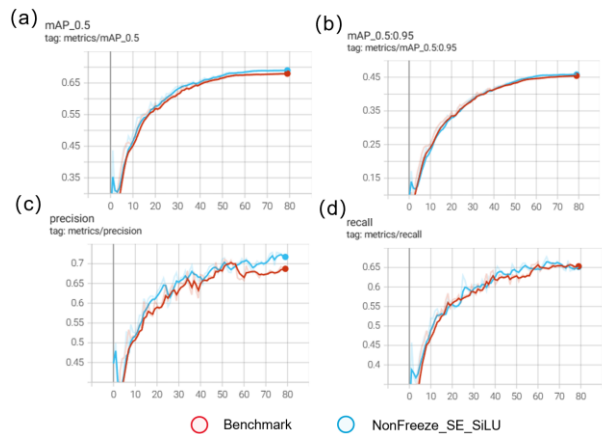
Fig. 7 Result of Applying SE and SiLU excluding Freezing Mechanism graph for (a) mAP_0.5, (b) mAP_0.5:0.95, (c) Precision and (d) Recall

References

1. W. Lan, J. Dang, Y. Wang, and S. Wang, "Pedestrian detection based on yolo network model," in 2018 IEEE International Conference on Mechatronics and Automation (ICMA), 2018, pp. 1547–1551.
2. T.-H. Wu, T.-W. Wang, and Y.-Q. Liu, "Real-time vehicle and distance detection based on improved

yolo v5 network," in 2021 3rd World Symposium on Artificial Intelligence (WSAI), 2021, pp. 24–28.

3. Z. Lu, L. Ding, Z. Wang, L. Dong, and Z. Guo, "Road condition detection based on deep learning yolov5 network," in 2023 IEEE 3rd International Conference on Electronic Technology, Communication and Information (ICETCI), 2023, pp. 497–501.

4. A. Dodia and S. Kumar, "A comparison of yolo based vehicle detection algorithms," in 2023 International Conference on Artificial Intelligence and Applications (ICAIA) Alliance Technology Conference (ATCON-1), 2023, pp. 1–6.

5. M. Kaloev and G. Krastev, "Comparative analysis of activation functions used in the hidden layers of deep neural networks," in 2021 3rd Inter- national Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), 2021, pp. 1–5.

6. X. Xiao, T. Bamunu Mudiyanselage, C. Ji, J. Hu, and Y. Pan, "Fast deep learning training through intelligently freezing layers," in 2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), 2019, pp. 1225–1232.

7. H. Zhao, Z. Liang, D. Cai, and Y. Wang, "An improved method for infrared vehicle and pedestrian detection based on yolov5s," in 2022 International Conference on Machine Learning, Cloud Computing and Intelligent Mining (MLCCIM), 2022, pp. 377–383.

8. Kaymak and A. Uçar, "Semantic image segmentation for autonomous driving using fully convolutional networks," in 2019 International Artificial Intelligence and Data Processing Symposium (IDAP), 2019, pp. 1–8.

9. B. Xiao, J. Guo, and Z. He, "Real-time object detection algorithm of autonomous vehicles based on improved yolov5s," in 2021 5th CAA International Conference on Vehicular Control and Intelligence (CVCI), 2021, pp. 1–6.

Authors Introduction

Mr. Guandong Li

He is currently pursuing Bachelor of Robotics and Intelligent Devices as 3rd year student in the Department of Maynooth International Engineering College, Fuzhou University, China. His research interests are embedded systems and machine learning.

Mr. Yanzhe Xie

He is currently pursuing Bachelor of Robotics and Intelligent Devices as 3rd year student in the Department of Maynooth International Engineering College, Fuzhou University, China. His research interest is robotics and machine learning

Mr. Yuhao Lu

He is currently pursuing Bachelor of Robotics and Intelligent Devices as 3rd year student in the Department of Maynooth International Engineering College, Fuzhou University, China. His research interests are data science and embedded system.

Mr. Zongyan Wen

He is currently pursuing Bachelor of Robotics and Intelligent Devices as 3rd year student in the Department of Maynooth International Engineering College, Fuzhou University, China. His research interests are embedded systems and machine learning.

Mr. Jingzhen Fan

He is currently pursuing Bachelor of Robotics and Intelligent Devices as 3rd year student in the Department of Maynooth International Engineering College, Fuzhou University China. His research interests are software development and system learning.

Mr. Yuankui Huang

He is currently pursuing Bachelor of Robotics and Intelligent Devices as 3rd year student in the Department of Maynooth International Engineering College, Fuzhou University, China. His research interests are electronic information and machine learning.

Mr. Qinghong Ma

He is currently pursuing Bachelor of Robotics and Intelligent Devices as 3rd year student in the Department of Maynooth International Engineering College, Fuzhou University, China. His research interests are electronic information and machine learning.

Dr. Wei Hong Lim

He is an Associate Professor in Faculty of Engineering at UCSI University in Malaysia. He received his PhD in Computational Intelligence from Universiti Sains Malaysia in 2014. His research interests are optimization and artificial intelligence.

Guandong Li, Yanzhe Xie, Yuhao Lu, Zongyan Wen, Jingzhen Fan, Yuankui Huang, Qinghong Ma, Wei Hong Lim, Chin Hong Wong

Dr. Chin Hong Wong

He is a Lecturer at Maynooth International Engineering College at Fuzhou University, China. He received his PhD in Electrical and Electronic Engineering from Universiti Sains Malaysia in 2017. His research interests are Energy harvesting, signals and systems, and machine learning.

A Game Framework Based on the Disinformation Warfare in Russo-Ukrainian War

Jumpei Ono
Aomori University, Kobata, Aomori, Aomori, 030-0943, Japan

Takashi Ogata
Iwate Prefectural University, Sugo, Takizawa, Iwate, 020-0611, Japan
Email: j.ono@aomori.ac.jp, t-ogata@iwate-pu.ac.jp

Abstract

From 2022, when Russia began its invasion of Ukraine, to the present, the number of speeches and information directly/indirectly related to Russo-Ukrainian war has continued to increase in mass communication venues such as SNS and TV. This study proposes the idea of a game system that simulates the spread of disinformation by focusing on the information published by Ukraine as a list of disinformation speakers, especially on the topics they talk about. The game system is intended to provide players with psychological immunity against disinformation by simulating the spread of disinformation.

Keywords: Disinformation, Russo-Ukrainian war, Game System

1. Introduction

Russo-Ukrainian war has two parts. On February, 2014, Russia annexed Crimea. Also, in February 2022, Russia has invaded the Donbass region of Ukraine. In this war, the discourse and information directly and indirectly related to Russo-Ukrainian war continues to grow in mass communication spaces such as social networking sites and television. In this context, Russia sees the information space as another battlefield, with NATO [1] and Ukraine (https://cpd.gov.ua/) reporting on the situation. Ogata [2], [3], [4], [5] has collected discourses on Russo-Ukrainian war and summarized them as "Russo-Ukrainian War as Narrative Warfare"[6], while Ono and Ogata have published articles on this war, proposing the concept of a narrative-generating game based on Russo-Ukrainian war[7], [8].

This study proposes the idea of a game system (we call the game system as simply "game" in the following description) that simulates the spread of disinformation by focusing on the discourses compiled by Ogata [2], [3] [4] [5] and published by Ukraine as a disinformation speaker list, especially the topics that the speakers talk about. The game is designed to help players develop psychological immunity to disinformation by simulating the spread of disinformation.

2. Background

Disinformation is information that influences a person's perception and distorts their decision making.

For example, fake news, which spreads sensational or conspiratorial content, including falsehoods, is a typical example of disinformation. Today, battles that make full use of disinformation, in addition to the conventional domain of warfare, are positioned as a new domain of warfare and are called cognitive warfare [9]. On the other hand, Ogata, who believes that falsehood is not the only means of attacking opponents in battles that affect the perception of the target, has proposed the term "narrative warfare" as a domain in which attacks are made regardless of whether the content is true or false [6]. For example, one method might be to present a distorted conclusion as the conclusion of a logic that only mentions the truth about individual elements. Maan have focused on narrative warfare, and they have been exposed to it from the time of the emergence of the Islamic State to the present day [10]. Narrative warfare in Russo-Ukrainian war has also received attention [11].

Attempts to position players on the disseminators of disinformation have been experimented with by Roozenbeek and Linden [12] with *Bad News* [13]. *Bad News* is a web game that provides a virtual experience of a fake news speaker. In this game, each player follows the game's instructions and gradually increase their influence on public opinion by selecting one of fake new candidates. The player can efficiently increase their influence by selecting more appropriate fake news, and their influence decreases if they select an inappropriate fake for the occurred event. Roozenbeek and Linden found that this game allows each player to "cognitive resistance against fake news strategies." There are similar games on the market, such as *Plague Inc.* [14]. *Plague Inc.* is a simulation game. Each player in the game becomes an

epidemic and aims to wipe out humans. The player can select the game mode of fake news.

3. Proposed Game System

The structure of the game is presented after an overview of the proposed game. At last, an example of the game play is shown.

3.1. Overview of the proposed game

The game proposed in this paper is a game in which the player is a Russian spreading disinformation about the Ukrainian model. The goal of each player is to convert more Ukrainian faction to the Russian faction ("faction" in this paper does not mean a particular political party, people who believe in the state.). The game presents the player with events that occurred during Russo-Ukrainian War, and the player chooses the disinformation to spread accordingly. The Ukrainian faction is a set of non-player characters (NPCs) that are not controlled by the player, and each NPC has a trust parameter. This trust represents the level of trust in Ukraine, and the disinformation chosen by the player will fluctuate this value. The Ukrainian faction will switch to the Russian faction if the trust parameter falls below a certain value.

This work does not include factors such as the spread of misinformation ("misinformation" is false information that is spread without malicious intent), the phenomenon of disinformation having a greater impact due to biased information on social networking sites, or the range of influence of disinformation or the number of people affected. These elements are considered important for simulating disinformation and will be discussed as future topics.

3.2. The flow in a game play

A player becomes the sender of disinformation and changes the parameters of non-player characters (NPCs) made up of several people, to change the NPC's faction.

The game repeats the turn as many times as the number of events that make up the scenario, with (i) the generation of events, (ii) the selection of disinformation, and (iii) the calculation of effects as a single turn. At the end of the turn, the amount of trust in the Ukrainian faction of NPCs is scored.

When choosing disinformation, the player chooses one of the disinformation candidates. The mechanism changes the parameters of all NPCs based on the selected disinformation.

3.3. Components of the proposed game

The proposed game has non-player characters (NPCs), events, and disinformation. These attributes are used to calculate the score that evaluates the player, or are values related to the game play.

Non-Player Character (NPC)

- Trust: if this value is 0, the NPC has completely lost trust in the faction. This value also indicates resistance to false information.
- Characteristics: What personality traits the NPC has. If there are no characteristics worth mentioning, it is set to null. Currently, it indicates if the NPC has conspiracy theorist traits.

Event

- Characteristics: Classification of the event. Disinformation is more influential if it matches the "Characteristics" in the disinformation.

Disinformation

- Strength: The degree to which the disinformation affects the NPC's trust.
- Reality: The degree to which the disinformation is factual, indicating the likelihood that the NPC will accept the disinformation; if this value is low, the likelihood of acceptance is low.
- Characteristics: Classification of the disinformation. If the disinformation is appropriate to the nature of an event, it will have a greater influence on the NPC's trust.

Disinformation in the proposed game based on *СПІКЕРИ, ЯКІ ПРОСУВАЮТЬ СПІВЗВУЧНІ РОСІЙСЬКІЙ ПРОПАГАНДІ НАРАТИВИ* (*Speakers who promote narratives that are in tune with Russian propaganda*, https://cpd.gov.ua/reports/spikery-yaki-prosuvayut-spivzvuchni-rosijskij-propagandi-naratyvy-2/). This list was revised once, and before and after the revision, excluding duplicates, the list contained 89 speakers. In this paper, if the same speaker sent the same disinformation before and after the revision, it was classified as a single source, and if different speakers sent the same disinformation, they were classified as different sources. The classification was then done in a bottom-up fashion and organized as disinformation used in the game. Each classification and the example of its breakdown are provided as the Appendix.

The trust value of the NPC is reduced based on the following Eq. (1) and. Eq. (2) First determine whether the NPC is affected by the disinformation (Eq. (1)), and then determine the degree of influence (Eq. (2)). If the NPC is committed to a conspiracy theory, the trust value used in the Eq. (1) and Eq. (2) Regarded as half the value.

$$P = 100 + (R - T) \qquad (1)$$
$$A = S - T/4 \qquad (2)$$

P: Probability of decreasing trust, A: Amount of trust decreases, R: Reality in selected disinformation, T: Trust that the NPC has, S: Strength in selected disinformation

3.4. A game session example and a short discussion

The following shows the example of a game session. There are five NPCs. One NPC in the five NPCs believes in the conspiracy theory. The session has six events. "Disinformation" in the following description means the fake information selected by a player. "NPCs" in the description includes the changed "Trust" in the respective NPC ("Number").

Event 1: "Russia have launched an invasion."
Disinformation: "The U. S. is provoking/inspiring Russia."
NPCs: Number 0: Trust 82 (conspiracy thinking) / Number 1: Trust 95 (nil) / Number 2: Trust 95 (nil) / Number 3: Trust 95 (nil) / Number 4: Trust 95 (nil)
Average: 92.4, Total trust: 462/500

Event 2: "Russia occupied a city."
Disinformation: "Russia has not attacked civilians."
NPCs: Number 0: Trust 62 (conspiracy thinking) / Number 1: Trust 88 (nil) / Number 2: Trust 88 (nil) / Number 3: Trust 88 (nil) / Number 4: Trust 88 (nil)
Average: 82.8, Total trust: 414/500

Event 3: "A massacre occurred."
Disinformation: "The massacre in Bucha is a fake."
NPCs: Number 0: trust 9 (conspiracy thinking) / Number 1: trust 50 (nil) / Number 2: trust 88 (nil) / Number 3: trust 50 (nil) / Number 4: trust 50 (nil)
Average: 49.4, Total trust: 247/500

Event 4: "The war continues."
Disinformation: "Ukraine/Zelensky wants war."
NPCs: Number 0: trust 0 (conspiracy thinking) / Number 1: trust 32 (nil) / Number 2: trust 80 (nil) / Number 3: trust 32 (nil) / Number 4: trust 32 (nil)
Average: 35.2, Total trust: 176/500

Event 5: "The war continues."
Disinformation: "Russia will win the war."
NPCs: Number 0: trust 0 (conspiracy thinking) / Number 1: trust 10 (nil) / Number 2: trust 70 (nil) / Number 3: trust 10 (nil) / Number 4: trust 10 (nil)
Average: 20.0, Total trust: 100/500

Event 6: "Russia has retreated."
Disinformation: "They should not hunt down Russia."
NPCs: Number 0: trust 0 (conspiracy thinking) / Number 1: trust 0 (nil) / Number 2: trust 57 (nil) / Number 3: trust 0 (nil) / Number 4: trust 0 (nil)
Average: 11.4, Total trust: 57/500

In the above session, the value of trust to Ukraine of the conspiracy theorist (Number 0) is rapidly reduced. In the other NPCs, the NPC of Number 2 continues to have the value of trust to Ukraine. However, the values of trust of the other NPCs, Number 1, 3, and 4, finely become 0. When we see the entire process, we know the accelerated growth of suspicion to Ukraine.

In the game, the value of trust is also the power to resist the influence of disinformation. NPCs are repeatedly exposed to some type of disinformation and the influence of disinformation for all NPCs are gradually increased.

4. Conclusion

In this paper, we proposed the idea of a game system related to disinformation in the context of Russo-Ukrainian war and developed a simple prototype dealing with disinformation. We used one pattern of disinformation used actually in Russo-Ukrainian war. This game framework, including several NPCs and a human player, simulates the spreading process of disinformation. The goal of a player in one session is to change NPCs from Ukrainian faction to the Russian faction. The player may acquire the knowledge to prevent influences of disinformation through the game process. In the future, we would like to extend the scale of the game.

Acknowledgements

References

1. E. Lange-Ionatamišvili, Analysis of Russia's information campaign against Ukraine. Executive Summary. Executive Summary, Riga: NATO Strategic Communications Centre of Excellence, 2014.
2. T. Ogata, The Narrative warfare of Russo-Ukrainian war in Japan: Introduction, The Japanese Society for Artificial Intelligence SIG-LSE0C31, 70, 9-82, 2022.
3. T. Ogata, Survey, analysis, and critique of the social discourse in contemporary Japan based on an example: Post-narratological study using the phenomenon of "Hashimoto Toru", Proc. of the 39th Annual Meeting of Japanese Cognitive Science Society, 632-640, 2022.
4. T. Ogata, War and narrative generation amateur talk: Beyond the landscape, in sad Japan where there is also a World's Commander Hashimoto-kun, The Japanese Society for Artificial Intelligence SIG-LSE0C31, 17-142, 2022.
5. R. Abe, J. Ono and T. Ogata, Concept acquisition and sentence generation of Russian-Ukrainian war based on linguistic data from TV news and commentary programs, The Japanese Society for Artificial Intelligence SIG-LSE0C31, 71, 161-175, 2023.
6. T. Ogata, Monogatarisen Toshite no Russo Ukraine Senso, Shin'yosya, 2023.
7. J. Ono and T. Ogata, Narrative generation game based on the Ukraine War, The Japanese Society for Artificial Intelligence SIG-LSE0C31, 69, 143-148, 2022.
8. J. Ono and T. Ogata, Designing a narrative generation game based on the Russian Invasion of Ukraine, Proceedings of the 2023 International Conference on Artificial Life and Robotics (ICAROB2023), 732-736, 2023.
9. Y. Koizumi, K. Kuwabara, and K. Komiyama, Nise Joho Senso: Anata no Atama no Naka de Okoru Tatakai, 2023.

10. Mann, A., Narrative Warfare. Narrative Strategies Ink., 2018.

11. N. Aleksejeva (Reserch coodinator) and A. Carvin (Ed.), Narrative Warfare: How the Kremlin and Russian News Outlets Justified a War of Aggression against Ukraine, 2023.

12. J. Roozenbeek and S. V. D. Linden, Fake news game confers psychological resistance against online misinformation, Palgrave Communications, 5 (65), 2019. DOI: https://doi.org/10.1057/s41599-019-0279-9

13. Cambridge Social Decision-Making Lab., Bad News Game. https://www.getbadnews.com/en/play, 2018.

14. Ndemic Creations, Plague Inc., Ndemic Creations, 2021.

Appendix: The types of disinformation

The following categories are based on the *СПІКЕРИ, ЯКІ ПРОСУВАЮТЬ СПІВЗВУЧНІ РОСІЙСЬКІЙ ПРОПАГАНДІ НАРАТИВИ* (*Speakers who promote narratives that are in tune with Russian propaganda*, https://cpd.gov.ua/reports/spikery-yaki-prosuvayut-spivzvuchni-rosijskij-propagandi- naratyvy-2/). "~~~" is a space to insert a specific word, one of the U.S., NATO, Ukraine, or Zelensky.

1. ~~~ is provoking/inspiring Russia
2. It is a proxy war between ~~~ and Russia
3. The West is running a propaganda campaign
4. Russia-hatred (Russophobia)
5. NATO expansion is bad
6. Ukraine is suppressing Russian speakers in the country
7. Ukraine/Zelensky wants war
8. Ukraine is committing acts against humanity
9. Bucha massacre is a fake
10. Ukraine has a biological weapons laboratory
11. We should not give arms to Ukraine
12. Ukraine started war before February 24/Maidan Revolution
13. Ukraine is a Nazi
14. Ukraine should make concessions
15. Ukraine did not comply with the Minsk agreements
16. Russia will win the war
17. Sanctions are not getting through to Russia
18. Russia has not attacked civilians
19. ~~~ should negotiate with Russia/Putin
20. ~~~ should not blame Russia/Putin
21. ~~~ should not hunt down Russia
22. Crimea is Russian territory
23. Russia has freedom of speech

For example, 1. "~~~ is provoking/inspiring Russia" includes the following contents. Other types also include many contents. Small amount of contents are classified as "Other" and was not discussed in this paper.

The U.S. and Ukraine have actively provoked Russia into aggression./ The U.S. and Ukraine have aggressively provoked Russia into aggression. / Putin has been humiliated and Russia persecuted for decades. / The West provoked Russia into war with Ukraine. / Zelensky provoked Russia into launching special operations against it. / The U.S. has provoked Putin for years. / The events in Ukraine are a provocation by the U.S. and NATO. / The U.S., through NATO, provoked Russia to war in Ukraine. NATO and the U.S. provoked Putin. NATO provoked Putin. NATO provoked Russia The West provoked Putin. / The West provoked Russia. / The U.S. provoked Russia. /The West provoked Putin. / Arming Ukraine provokes Russia. / Zelensky is provoking Putin. / NATO has crossed the "red line." / The U.S. and UK have crossed Putin's "red line." / For eight years, Ukrainians have been provoking Russia. / Putin's decision to invade Ukraine was NATO's fault. / NATO intentionally deployed on the Russian border. / The U.S. and Europe provoked the Russian president. / The U.S. is responsible for Russia's attack on Ukraine. / The U.S. needs to stop the madness in Ukraine provoked by the U.S.

Authors Introduction

Dr. Jumpei Ono

He received his bachelor's degree from the Faculty School of Software and Information Science, Iwate Prefectural University in 2010. He received his MS and PhD from the Graduate School of Software and Information Science, Iwate Prefectural University in 2014 and 2018, respectively. He worked as an information and communication technology instructor at the Vocational School of Digital Arts Sendai from 2018 to 2020. He is currently an assistant professor at the Faculty of Software and Information Technology, Aomori University.

Dr. Takashi Ogata

He received his BSS from Waseda University in 1983, his MS from Tsukuba University in 1992, and his PhD from the University of Tokyo in 1995. He has garnered industrial experience from working at software development companies since 1983. He has been an associate professor in the Faculty of Engineering at Yamanashi University since 1997 and is currently a professor in the Faculty of Software and Information Science at Iwate Prefectural University, where he has worked since 2005.

A Comparative Analysis of Eye Tracking between Veteran and Novice during Radiological Interpretation

Yuka Naito
Graduate School of Economics, Chuo University, Higashinakano Hachioji-shi, 192-0393, Japan

Jun Nakamura
Faculty of Global Management, Chuo University, Higashinakano Hachioji-shi, 192-0393, Japan

Yoshinobu Ishiwata
Graduate School of Medicine, Yokohama City University, 3-9, Fukura, Yokohama Shi Kanazawa Ku, 236-0004, Japan
Email: a19.3gem@g.chuo-u.ac.jp, jyulis.77f@g.chuo-u.ac.jp, ishi_y@yokohama-cu.ac.jp

Abstract

With the development of AI, AI may substitute for physicians in radiological interpretations in the future. The purpose of this paper is to identify factors that could form the foundation for algorithms enabling AI to explore pathological findings. This study focuses on the "angles of gaze trajectory" and conducts a comparative analysis between veteran and novice. The results indicate commonalities and variances in the angles of gaze trajectory between veteran and novice, particularly 10, 350, and 180 degrees. Both Veteran and Novice have a relatively higher frequency of gaze movement around 10 and 350 degrees compared to other angles. Veteran has a relatively higher frequency of gaze movement around 180 degrees compared to other angles, while novice has a relatively lower frequency around 180 degrees.

Keywords: Eye-tracking, Angles of eye trajectory, Artificial Intelligence (AI)

1. Introduction

Machine learning and deep learning advancements, coupled with high-performance GPUs, have led to a significant breakthrough in the field of artificial intelligence [1]. Particularly in image recognition, AI is considered to surpass human capabilities [2]. In the realm of healthcare, a critical application of AI's image recognition capabilities is evident, with one prominent example being the interpretation of CT images. Given the current shortage of radiologists, AI is expected to achieve a level of expertise in image interpretation that can serve as a viable alternative to human physicians. To attain this capability, various factors contributing to the training of AI in image interpretation must be identified. One such factor is the ocular movement of physicians during image interpretation. This paper aims to discuss the significance of physicians' gaze patterns as one of these factors, emphasizing the need to uncover and understand these aspects to enhance the learning process of AI in image interpretation.

2. Previous Research

Numerous studies have explored gaze measurements of physicians during the interpretation of radiographic images, a certain cue is presented on the screen, and an analysis is conducted on how participants can detect this cue in terms of gaze trajectory and movement velocity [3]. Studies comparing the gaze patterns of veteran and novices are also prevalent. For instance, an analysis of gaze fixation points during chest X-ray image interpretation shows that experts promptly identify lesions and repeatedly confirm their presence while also scanning other areas, whereas novices tend to fixate on structurally complex regions [4]. In research on overlooking and missing findings during medical image interpretation, it has been demonstrated that experts thoroughly observe photographs, strongly focusing on the accurate location of anomalies, while novices exhibit more inconsistent observation patterns, with their gaze often diverted to complex positions where organs overlap [5]. This line of research on tacit knowledge and expertise in skill-demanding scenarios has unveiled that veterans and novices differ in terms of "where to look," influencing their subsequent performance [6]. As a preliminary step, it is necessary to extract and quantitatively analyze the gaze patterns of experts before linguistically formulating such characteristics. Therefore, the purpose of this paper is to conduct a comparative analysis between novices and veterans regarding the angles of gaze trajectories, which could serve as the foundation for algorithms exploring findings in artificial intelligence (AI). Furthermore, the objective is to propose recommendations for AI to learn

from the disparities in gaze trajectories between novices and veterans, ultimately enabling more accurate and effective image diagnostics in the future.

3. Experiment

The gaze measurement experiment was carried out as follows.
● Participants: One veteran radiologist (with over 10 years of experience) and one novice radiologist (with 1~3 years of experience).
● Data:15 cases with abnormal findings.
● Experimental method: Tobii Pro Lab eye-tracking software was employed. The camera was fixed. Participants were presented with CT images on a computer screen and performed interpretation of CT images by scrolling through the images using a mouse, following their usual reading methods.

4. Analysis Methods

Duchowski provides guidelines regarding data collection and interpretation in the context of eye tracking. The book introduces applications and instances of success in fields such as medicine, psychology, and human factors engineering, underscoring the effectiveness of employing eye tracking in the medical domain [7]. Furthermore, Holmqvist et al. offer an in-depth exposition of data analysis methods in eye tracking research. Of particular note, fixation points serve as a frequently employed metric [8]. However, the analyzed metric in this paper is the angle of the gaze movement as it sequentially moves from one fixation point to the next in succession. (Henceforth, this shall be referred to the angle of eye movement.) The reason for this is to confirm whether the direction of eye movement, namely the trajectory of the gaze, when radiologists interpret CT images is consistent with the characteristics of everyday human gaze as elucidated in previous research [9]. The characteristics of human eye movement elucidated in this previous research are detailed in the subsequent discussion section. The employed eye-tracking software in this study automatically aggregates parameters such as fixation points, fixation counts, fixation duration, and coordinates of fixation points, subsequently generating this information as output data. The trajectories connecting successive fixation points (hereinafter referred to as gaze trajectories) are visually indicated by red lines, enabling direct observation.

The angles of gaze movement were defined based on the eye-tracking data obtained using Tobii Pro Lab and analyzed using a custom-developed program. In this context, angles of gaze movement refer to the direction changes of gaze trajectories as it progresses across different fixation points. The concept is illustrated in the Fig. 1, where black circles represent fixation points (pixel), and connecting lines depict the gaze trajectory between these points. When the coordinate shifts from

(971, 810) to (988, 867), and subsequently to (974, 816), the angle of gaze movement is denoted by Angle 1 in Fig. 1. The directional movement is measured clockwise. Similarly, when moving from (988, 867) to (974, 816) and further to (1013, 837), the angle of gaze movement is represented as Angle 2 in Fig. 1.

Fig.1 Angles of gaze movement

5. Results

The results of the measured angles of gaze movement, as captured by the software, are presented in Fig. 2 and Fig. 3.

Fig.2 Angles of gaze movement in veteran (with abnormal findings)

Fig.3 Angles of gaze movement in novice (with abnormal findings)

A comparison between veteran and novice revealed notable differences in angles of gaze movement. In the graphs, the numbers displayed on the outer circumference represent the angles, while the numbers within the circumference, namely the radii, indicate the

frequency of the corresponding angles. Skilled practitioners exhibited a higher frequency of gaze movement angles around 180 degrees within the 360-degree spectrum. This is demonstrated in Fig. 2. In Fig. 2, angles such as 10 degrees, 350 degrees, 190 degrees, and 180 degrees were observed with higher frequencies. When examining the results for novice, as shown in Fig. 3, angles such as 350 degrees and 10 degrees were more pronounced in their gaze movement. Comparing the results presented in Fig. 2 and Fig. 3, it is evident that veteran frequently exhibit gaze movement angles near 180 degrees, a pattern not as pronounced in novice practitioners. This suggests that veteran tend to pause and carefully observe between fixation points when shifting their gaze, while novice appear to transition their gaze more promptly from one point to another.

6. Discussion

The paper highlighted differences in gaze movement between veteran and novice during interpretation. The analysis of angles of gaze movement, a hypothesis was formulated. The hypothesis suggests that skilled practitioners pause for careful observation between fixation points when shifting their gaze, whereas novices shift their gaze more promptly from one point to another, lacking this intermediate pause.

Considering that novices possess less experience than skilled practitioners, they might find it challenging to predict the location of anomalies. This could potentially explain why novices tend to survey the entire CT images thoroughly and quickly. In contrast, veteran can infer the approximate location of anomalies, allowing them to concentrate on specific areas. Interviews with experienced practitioners aligned with this explanation. Inexperienced individuals may have attempted to observe a broad area as much as possible. As a result, the gaze of novice may have exhibited movements along the Z-axis (the head-foot direction in CT images), while simultaneously surveying the XY plane comprehensively. Conversely, veteran tend to direct their gaze along the Z-axis (head-foot direction). Therefore, in a series of similar CT images, should an anomaly be detected, it could be recognized as a flicker. According to the previous research [9], the characteristics of human gaze encompass distinct roles for central vision and peripheral vision, where grasping stimuli occurs through peripheral vision and subsequently confirming details via central vision constitutes a fundamental process. Veterans shift their gaze along the Z-axis (head-foot direction) with peripheral vision, and subsequently enhance the resolution with their central vision when anomalies are detected. This phenomenon aligns with the description provided by adept physicians as "recognizing anomalies as flickers."

Additionally, the gaze patterns of veteran revealed repeated back-and-forth movements along lines near 180 and 360 degrees. Similar repetitive movements along horizontal and vertical lines have been observed

in previous studies [9]. The research has been elucidated that repetitive movements along horizontal lines are more frequent than those along vertical lines. In other previous research [10], it has revealed that the capability to accurately perceive information is higher in the horizontal direction of eye movements compared to the vertical direction. This paper also found that novice exhibited more frequent back-and-forth movements along horizontal lines than vertical ones. This phenomenon can be attributed to reasons such as the relatively easier execution of horizontal movements compared to vertical ones, as well as the cautious tendency of novices due to the difficulty in anticipating the location of abnormal findings. Consequently, novices may opt to observe a broader scope of the CT image.

7. Conclusions

The purpose of this paper is to conduct a comparative analysis between novices and veterans regarding the angles of gaze trajectories, which could serve as the foundation for algorithms exploring findings in artificial intelligence (AI). Furthermore, the objective is to propose recommendations for AI to learn from the disparities in gaze trajectories between novices and veterans, ultimately enabling more accurate and effective image diagnostics in the future. This paper specifically analyzed the angles of gaze movement and speed of veteran and novice radiologists. However, it is important to note that this research had limitations in terms of a limited number of participants and sample size. Future studies should involve a larger sample size to further investigate the elements of algorithms during radiological interpretation.

8. Acknowledgements

References

1. T. Hirasawa, Y. Ikenoyama, M. Ishioka, K. Namikawa, Y. Horiuchi, H. Nakashima, T. Tada, and J Fujisaki, Endoscopic AI Diagnosis of Gastric Cance, The journal of the Japan Laser Medical Society, 42(4), 2021-2022, pp.250-260
2. T. Hirasawa, T. Tada, and J. Fujisaki, The Future of Gastric Cancer Diagnosis with the Utilization of Artificial Intelligence (AI), The journal of the Japan Endoscope for Fire Extinguishers Society, 62(12), 2020, pp.3031-3140
3. N. Tsutsumi, M. Abe, T. Someya, and T. Matsumoto, Fundamental Consideration for Eye Movement Analysis during X-ray Image Interpretation. Journal of the Japan Society of Medical Image and Information Science, 17(3), 2000, pp.152-160

4. T. Matsumoto, T. Iinuma, Y. Tateno, K. Fukuhisa, K. Ohata, M. Kaneko and S. Ikeda, Gaze Point Analysis of Physicians during Interpretation of Chest X-ray Images, Medical Imaging Technology, 35(4), 1990, pp.212-216

5. T. Terashita, T. Sato, T. Ogura, and K. Doi, Development and Evaluation of an Eye-tracking-based Application for Measuring Medical Image Interpretation Abilities. Transactions of the Japanese Society for Artificial Intelligence, 23(1), 2021, pp.43-46

6. R. Fukuda, K. Yoshida, Y. Ono, J. Matsubara, and M. Kudo, Feature Extraction of Skilled Caregivers using Eye Tracking, 2011

7. Andrew T. Duchowski, Eye Tracking Methodology, Theory and Practice, Springe, 2017

8. K. Holmqvist, M Nystrom, R. Andersson, R. Dewhurst, H/ Jarodzka, and J. Weijer, Eye Tracking: A Comprehensive Guide to Methods and Measures, Oxford University Press, 2011

9. K. Hashimoto, K. Ushiki, M. Nakamura, G. Watanabe, and N. Ogawara, Consideration on Stimulus Attractiveness and Arrangement in Video Playback. Journal of the Information Processing Society of Japan: Human and Computer Interaction, 2006(3), 2006, pp.75-81

10. K. Murata, T. Miyake, and M. Moriwaka, Effectiveness of Gaze Input Systems - Identification of Conditions to Ensure Point Accuracy and Speed and the Influence of Movement Direction. Ergonomics, 45(4), 2009, pp.226-235

Authors Introduction

Ms. Yuka Naito

She is a 1st year graduate student at the faculty of Economics, Chuo University, Japan. The areas of his research interest are Experimental Economics, Behavioral Economics, Cognitive Psychology, and Artificial Intelligence.

Dr. Jun Nakamura

He is a professor at the Faculty of Global Management, Chuo University, Japan. Previously he was working as a professor at the Graduate School of Engineering Management, Shibaura Institute of Technology. His main research areas are Cognitive Science and Humanities.

Development of Notification System to Prevent Working Productivity from Declining Caused by Increased Carbon Dioxide Concentration

Yohei Kamoda
Chuo University, Faculty of Global Management, 742-1 Higashinakano, Hachioji-City
Tokyo, 192-0393, Japan

Jun Nakamura
Chuo University, Faculty of Global Management, 742-1 Higashinakano, Hachioji-City
Tokyo, 192-0393, Japan
Email: kamoda.yohei@gmail.com, jyulis.77f@g.chuo-u.ac.jp

Abstract

In association with the change of work style to remote, the home environment should be considered in terms of carbon dioxide concentration (CDC) which has negative effects on the human body, such as less cognitive abilities. To challenge this problem, the authors developed an alert system that supports remote workers to be notified of an increase in CDC, by combining Raspberry Pi, CDC sensor, and Slack. As a result, the developed alert system was able to support the user in keeping the CDC in the room below the set threshold. In addition, the system shows that a significant increase in CDC can be observed in the room when it comes to insufficient ventilation.

Keywords: Carbon Dioxide Concentration (CDC), Raspberry Pi, Working Productivity

1. Introduction

In recent years, the spread of the Internet and the coronavirus have prompted many companies to adopt remote work. According to the Pew Research Center (2022), remote work rates have not declined much even after the peak of the pandemic [1]. In addition, many workers have a favorable attitude toward remote work after the pandemic (Da Silva et al., 2023) [2]. Therefore, remote work will likely continue to be promoted after the end of pandemic.

In consideration of the situation mentioned above, working productivity is noteworthy. Generally, the promotion of remote work in the wake of the coronavirus is interpreted as an increase in productivity of work. In fact, there is an argument that the shift to remote work leads to increased productivity, by thinking from the perspective of the amount of work, work-life balance, and job satisfaction (Kurdy, 2023) [3]. However, most of the papers referring to productivity in remote work simply compare productivity before the introduction of remote work with after doing that. In other words, there is little discussion from the perspective of changes in productivity caused by the home environment.

When pursuing productivity in remote work, a home environment that has become the working environment should be focused. Krivonossova (2022) points out that distractions in remote work are caused by the home environment, such as TVs,

books, and beds [4]. However, different from the perspective above, this paper focuses on the composition of indoor air. According to Azuma (2018), increased indoor CDC due to human breathing and other factors can cause negative impacts on the human body, such as dizziness, headaches, and reduced cognitive and problem-solving abilities [5]. Therefore, remote workers need to maintain indoor air quality by ventilating and other ways. However, the increase in CDC cannot be perceived easily, and it is difficult to know the appropriate timing to ventilate. Therefore, there is a possibility that the increase in CDC is causing a decrease in working productivity even if remote workers do not perceive it.

The purpose of this paper is to prevent an increase in CDC and a decrease in working productivity by using the system developed to notify the appropriate timing of ventilation.

2. Previous Research to Set Threshold

The unit used to express CDC is the "ppm (parts per million)". To determine the CDC threshold for the notification, it should be referred to the indoor CDC standards published by various public agencies.

The Wisconsin Department of Health Services (2023) states that 400 to 1,000 ppm is a reasonable standard for CDC indoors. Moreover, it points out that from 1,000 to 2,000 ppm will cause complaints

Yohei Kamoda, Jun Nakamura

of drowsiness and poor air quality, and from 2,000 ppm and above will cause health problems such as headaches, drowsiness, decrease in concentration to work, and mild nausea [6]. In addition, Japan's Ministry of Health, Labour and Welfare (2016) has set the environmental health management standard for CDC at 1,000 ppm in its Building Sanitation Law because CDC levels above 1,000 ppm cause fatigue, headaches, tinnitus, and breathlessness in the human body [7].

From the above, it is reasonable to set the threshold used in this paper at 1,000 ppm. Thus, the standard for notification used in the developed system in this paper can be set at 1,000 ppm.

3. Specifications of Developed System

3.1. System Overview

Fig. 1 shows a photograph of the overall CDC measurement and notification system created in this paper, and Fig. 2 shows a wiring diagram of the system. The system consists of a small computer called "Raspberry Pi model 4B" and two modules called "MH-Z19C" and "LCD 1602". The computer and display are connected by an i2c communication interface, "PCF8574".

Fig. 1. Overall Picture of CDC Measurement and Notification System

Fig. 2. Wiring Diagram of the System

The functions implemented in this system can be divided into measurement and notification. These two functions are intended to solve the problem stated in Chapter 1.

3.2. Measurement Function

The CO2 sensor module, MH-Z19C, was used to measure CDC. In terms of the program code to measure CDC, the function from the library

available on GitHub created by Ueda (2018) is utilized to simplify source code [8]. There are two measurement functions using this. The first is to record measurement results as csv file for analysis. The second is to immediately check the results at any given time through a small display.

3.3. Notification Function

In the system, Slack, the communication tool in business, was used for the CDC notifications. The purpose of this is to send a notification to the user to urge them to ventilate when the CDC value exceeds the threshold level of 1,000 ppm.

In addition, three conditions were set for notification. First, the Raspberry Pi must be connected to the Internet. Second, the notification is sent when the measured CDC value exceeds 1,000 ppm. Third, the judgment on whether to send a notification is made every five minutes.

4. Methods of Evaluation and Data Measurement

4.1. Environment during Measurement of Carbon Dioxide Concentration Data

The data measurements in this paper were conducted by the authors in their room. A simplified figure of the room where the data measurements were taken is shown in Fig. 3. To measure CDC as close as possible to the air that a person breathes, the sensor was placed close to the person. At this time, the position of the sensor was adjusted to avoid direct contact with the exhaled breath to obtain accurate data. Also, the "natural ventilation system" in this paper refers to the 24-hour ventilation system in the bathroom outside the "door", and the room is ventilated by taking in fresh air through the gaps around the "door" in the experimental environment.

When measuring data, test measurements were conducted under the same environment for at least 60 minutes prior to the start of the measurement to confirm that the CDC remained stable at less than 1,000 ppm. In addition, the working human was not away from the seat during the measurement and was assumed to be engaged in constant breathing activity.

Fig. 3. Environment of Room during Measurement

4.2. System Evaluation

In the evaluation of the system, the effect of the notification system on the increase in CDC is evaluated. Assuming a situation of inadequate ventilation, the "natural ventilation system" is stopped. Then, the "Window" in Fig. 3 is opened at the time of the notification to urge the user to ventilate the room. As a result, the system will be evaluated whether it has been able to contribute to reducing the CDC in the room below 1,000 ppm.

5. Results

Under the environment and conditions described in Chapter 4, an experiment was conducted by opening the "Window" and ventilating the room when the measured CDC value in the room exceeded 1,000 ppm, and a Slack notification was received. Table 1 shows the date and time of the measurement and the total measurement time, and Table 2 shows the elapsed time until when the notification was sent and the conducted ventilation time. Also, the CDC transition obtained from this measurement is shown in Fig. 4. In Fig. 4, the period during which the ventilation was conducted is colored.

Table 1. Date and Total Time of Measurement for Fig. 1

Start date and time	End date and time	Total Time
2023/9/27 19:00	2023/9/28 0:00	300 minutes

Table 2. Elapsed Time from Start of Measurement to Notification and Ventilation Time

Elapsed Time (minutes)	Ventilation Time (minutes)
65	15
95	15
120	30
190	35

Fig. 4. Transition of Carbon Dioxide Concentration when Ventilation is Conducted at the Time of Notification

As shown in Table 2 and Fig. 4, it can be said that ventilation at the time of the notification contributes to keeping the CDC in the room below 1,000 ppm. However, when the "Window" was closed, the CDC rapidly rose and exceeded 1,000 ppm. In the end, 135 minutes of ventilation time was required out of 300 minutes. However, it would be fair to say that positive results were obtained for the question, "Can the notification system help to keep the CDC in the room below 1,000 ppm?"

6. Discussion

It is fair to say that the system developed in this paper can help to prevent the increase of CDC in a room, as shown in the results of Chapter 5. In addition, it also showed that a room with inadequate ventilation has a high possibility of increasing CDC causing not only reduced productivity but also health problems. In other words, the system developed in this paper can prevent the decline in working productivity by knowing the adequate timing to ventilate. Also, the social value of this system in realizing the goal of improving productivity in remote operations can be well recognized.

However, this paper does not actually compare working productivity. Therefore, it is necessary to consider the possibility that Slack notifications may interfere with concentration on work. From this viewpoint, this system will not necessarily lead to increased working productivity based on the only verification in this paper.

In addition, the system developed in this paper only notifies the user of appropriate ventilation timing. Thus, the control of CDC in the room depends on the opening and closing of windows by the user. This is sufficient for remote work but needs further expansion to be utilized in other locations. As Fig. 4 shows, in rooms with inadequate ventilation, ventilation by opening and closing windows only becomes a temporary measure. Considering that the CDC rose

immediately after the window was closed, a ventilation system that automatically adjusts the CDC independent of user behavior would be needed to maintain the appropriate working environment.

Also, according to OPEN ACCESS GOVERNMENT (2021), since the COVID-19 pandemic, workers' awareness of flexible work styles and better working environments has increased [9]. Therefore, the system developed in this paper can be used to improve the working environment at a more advanced stage than the promotion of remote work. The importance of this system will be recognized, and its demand will rise in the current situation that COVID-19 has promoted not only remote work but also awareness of the work environment.

7. Conclusion

The purpose of this paper is to prevent an increase in CDC and a decrease in working productivity by using the system developed to notify the appropriate timing of ventilation.

The CDC notification system developed in this paper could be utilized to improve the working environment for remote workers. As shown in Fig. 4, the system was able to help keep the CDC in the room below 1,000 ppm by encouraging ventilation at the appropriate time. However, it does not indicate whether it was possible to prevent working productivity decline. Therefore, further verification of working productivity will be necessary in the future.

The data measurements in this paper were conducted in the limited environment of the authors' room, and the number of experimental subjects is insufficient. Therefore, it will be necessary in the future to acquire a large amount of data and to consider detailed environmental conditions such as the location of ventilation fans, actual flow of air, the number of people, room size, and temperature. This would provide detailed findings and suggestions for ventilation.

Acknowledgement

References

1. Pew Research Center. (2022, February 19). *COVID-19 Pandemic Continues To Reshape Work in America.*
2. Da Silva, A. D., Georgarakos, D., & Weißler, M. (2023). How people want to work – preferences for remote work after the pandemic. *ECB Economic Bulletin,* (1).
3. Kurdy, D. M., Al-Malkawi, H.-A. N., & Rizwan, S. (2023). The impact of remote working on employee productivity during COVID-19 in the UAE: the moderating role of job level. *Journal of Business and Socio-economic Development. 3*(4), 339-352.
4. Krivonossova, M. (2022, September 9). *6 Pros and Cons of Remote Work.* University Wire.
5. Azuma, K. (2018). Effects of inhalation exposure to carbon dioxide on human health in indoor environment. *Indoor Environment, 21*(2), 113-120.
6. Wisconsin Department of Health Services. (2023, March 29). *Carbon Dioxide*. Retrieved December 3, 2023, from https://www.dhs.wisconsin.gov/chemical/carbondioxide.htm
7. Ministry of Health, Labour and Welfare. (2016, October 31). [ventilation requirements for indoor air quality] shitsunai kuuki sitsu no tame no hitsuyou kankiryou (in Japanese).
8. Ueda, T. (2018, November 29). *Mhz-19*. GitHub. Retrieved December 3, 2023, from https://github.com/UedaTakeyuki/mh-z19/wiki/mh_z19
9. OPEN ACCESS GOVERNMENT. (2021, March 2). *Why is it important to have a great working environment?*

Authors Introduction

Yohei Kamoda

He is a student of Faculty of Global Management at Chuo University. He is the president of the "V.Code" programming club at the university. He also holds various certifications, including the Fundamental Information Technology Engineer Examination.

Dr. Jun Nakamura

He is a professor of Faculty of Global Management at Chuo University in Japan. He graduated from the Department of Technology Management for Innovation, School of Engineering, University of Tokyo, with highest honors. His area of research is cognitive sciences.

Visualization of the Skilled Physician's Gaze characteristic during Diagnosis

Taiki Sugimoto
Graduate School of Economics, Chuo University, Higashinakano Hachioji-shi, 192-0393, Japan

Jun Nakamura
Faculty of Global Management, Chuo University, Higashinakano Hachioji-shi, 192-0393, Japan

Yoshinobu Ishiwata
Graduate School of Medicine, Yokohama City University, 3-9, Fukura, Yokohama Shi Kanazawa Ku, 236-0004, Japan
Email: a19.3gem@g.chuo-u.ac.jp, jyulis.77f@g.chuo-u.ac.jp, ishi_y@yokohama-cu.ac.jp

Abstract

This study examines gaze movement differences in diagnosis between skilled and beginner physicians, aiming to identify factors influencing diagnostic speed and developing human resource and enhance artificial intelligence's diagnostic capabilities. Results reveal that experienced physicians, on average, spent 61% less diagnostic time than beginners, covering 49% of the gaze distance on the X and Y-axes. (comparison of the cumulative distance of gaze on the screen). Despite increased movement on the Z-axis (comparison of the scrolling speed of the CT scan results), skilled physicians moved 2-3 times faster, effectively narrowing attention and identifying specific areas.

Keywords: Experimental Economics, Eye-tracking, Cognitive Science, Artificial Intelligence (AI)

1. Introduction

Examination work in the medical industry is generally considered time-consuming and highly labor-intensive to train one CT service technician to acquire the knowledge necessary to prepare interpretation reports [1]. In addition, a nationwide survey of the actual status of the reading system for chest interpretation of radiograms for medical examinations, based on a questionnaire targeting 239 radiologists in 2018 [2], also indicated that it is difficult to train radiologists in a short period of time and requires a certain level of clinical and reading experience. In addition, the results of the survey indicated that the burden on radiologists is increasing due to the aging of the workforce and that the spread of CT scans will further reduce the manpower needed for the interpretation of radiograms, which may lead to the breakdown of the current radiogram system.

This has led to increased awareness of human resource development and medical AI development. Medical AI is being developed using deep learning. However, because deep learning has a black box problem in which the decision-making process cannot be identified by humans, the development of medical AI that can explain the process is being explored in the medical field, where safety is a high priority [3].

Considering the above issues, this study does not aim to contribute to the development of medical AI based on deep learning. This study aims to understand humans through the diagnosis of skilled physicians, and based on the understanding, aims to contribute to the realization of diagnosis like that of skilled physicians by beginner physicians and AI.

Diagnosis using imaging test results requires time and effort to master, resulting in a difference in the speed at which skilled physicians and beginner physicians can make a diagnosis. This is thought to be due to the difference in knowledge and experience between skilled and beginner physicians based on previous research [2]. This paper visualizes and analyzes the eye movement of both physicians resulting from the difference in knowledge and experience.

Humans are thought to possess the property of being able to make automatic inferences about stimuli with which they have frequent contact through practice, and this process of developing automaticity in judgment through practice is called procedural processing [4]. Procedural processing has been found to enable quicker decisions, improve reaction speed, and emphasize rational and practiced judgment [5]. It is thought that such procedural judgment is also made by experienced physicians when they use imaging test results to make diagnoses.

It is difficult to verbalize and provide effective guidance for diagnoses that are made unconsciously using such procedural judgments. Therefore, in this paper, by using a method [6] to visualize the eye movements of experienced physicians and beginner physicians during diagnosis by measuring eye gaze and to analyze the differences from a cognitive perspective, this paper aims to clarify the factors that lead to the speed with which

Taiki Sugimoto, Jun Nakamura, Yoshinobu Ishiwata

experienced physicians make diagnoses.

2. Purpose

The purpose of this paper is to contribute to human resource development and Medical AI development by clarifying factors that lead to the speed of making a diagnosis through a cognitive approach, such as measuring the eye movement of experienced physicians and beginner physicians during the interpretation of radiograms.

This paper is a study of novelty and academic significance. Conventional economics has focused on the results of decision-making and attempted to understand human behavior through mathematical decision-making models. However, this paper attempts to understand human behavior through a cognitive approach such as eye measurement by visualizing and analyzing the diagnostic process during the interpretation of a radiogram by using imaging tests. In addition, unlike previous studies that analyzed decision-making by comparison, this paper is unique in that it analyzes decision-making in the process of developing automaticity [4] as mentioned above.

This paper has social significance in that it aims to contribute to the development of beginner physician and development of medical AI by making the gaze at the time of diagnosis visible and clarifying factors that lead to the speed of giving a diagnosis.

3. Hypothesis

The hypothesis was developed as differences in the way experienced physicians and beginner physicians move their gaze during the interpretation of radiograms.

Experienced physicians have a narrower range of eye movement when searching for the presence or absence of symptoms/points of symptoms compared to beginner physicians. Since both interpretations of radiograms by experienced physicians and batting by experienced hitters have in common that they are performed using procedural judgment [4], a process in which the aforementioned automaticity of judgment has developed, previous studies [7] in which the gaze of experienced hitters while batting was studied were consulted. The results showed that experienced batters fix their gaze on the entire pitcher in their peripheral vision, fixing their gaze by on the point of attention prediction, rather than on a wide area as beginner batters gaze at a wide range. Based on this finding, it was hypothesized that the gaze movements of experienced physicians would be narrower than those of beginner physicians.

4. Methods

In this paper, experiments were conducted under the following conditions.

Yokohama City University Hospitals
- Experienced (more than 10 years career) 1 person

- Beginner (1~3 years of career) 1 person

With the cooperation of 2 physicians
- Abnormal findings 15 people (cases)
- No findings 15 people (cases)

A total of 30 people (cases) were tested for interpretation of head CT scan results, and eye gaze was measured using Tobii Pro Lab Screen Edition.

When interpreting the 30 cases, both physicians were not informed of the findings, the examination images were random, and both physicians diagnosed the cases in the same order, beginning and ending the diagnosis of each case at the physician's own discretion.

In comparing the experienced physician and beginner physician in this paper, the following two types of analyses were conducted.
- Comparison of the cumulative distance of gaze on the screen (X-axis and Y-axis movement): Abnormal findings - 15 cases.
- Comparison of the scrolling speed (Z-axis movement) of CT scan results: Abnormal findings - 3 cases.

The cumulative distance of the gaze (movement of the X-axis and Y-axis) was performed with reference to the velocity-detection method [6], which measures the distance of gaze movement on the screen. In addition, in this paper only analyzed abnormal findings (15 cases) from 30 cases.

The scrolling speed of the imaging test results (movement of the Z-axis) is an analysis unique to this paper; the movement of the Z-axis is an expression also used in the interpretation of radiogram and corresponds to the scrolling speed (diagnostic time and movement) between images of CT test results that capture cross sections. In the analysis, the diagnostic time for each image was recorded at the time the examination images were switched. The recording was done manually, and the error in the timing of when the physician switched images was kept within a range of 0 microseconds to 4 microseconds or less. If the physician did not change the scroll direction within 5 images, the time was recorded at intervals of 5 images, and if the physician did change the scroll direction within 5 images, the time was recorded each time an image was switched.

Both analyses were conducted with the purpose of focusing on gaze movements in the X-axis, Y-axis, and Z-axis to investigate whether the range of gaze movements of the hypothesized experienced physician's eye would be narrower than the range of gaze movements of the beginner physician's eye.

5. Results

Comparison of cumulative distance of gaze (X-axis and Y-axis movements)

Fig.1, Fig.2 below shows a comparison of the cumulative distance of gaze when the same case was diagnosed by an experienced physician and a beginner physician. Due to space limitations, the figures for each case are not shown here, but rather the total cumulative distance and the total normalized diagnosis time for all

cases, table summarizing their sum, mean, and variance, and figure showing the mean cumulative distance for the cases with findings (15 cases). The horizontal axis represents the time of diagnosis.

Table 1. X-axis and Y-axis movement in Abnormal Finding cases.

	Experienced_normalized time	Experienced_cumulative distance of gaze	Beginner_normalized time	Beginner_cumulative distance of gaze
Case1	66.99	31042.43	100	100626.18
Case4	84.16	99326.15	100	87496.00
Case5	57.23	115339.97	100	128587.59
Case6	43.91	66014.34	100	180581.15
Case10	70.13	63618.04	100	112969.58
Case13	64.24	42570.61	100	73038.85
Case14	55.58	43520.29	100	90969.10
Case15	58.08	44683.93	100	96657.33
Case17	52.90	40321.40	100	84090.70
Case18	39.92	27541.30	100	173406.17
Case19	77.37	52913.18	100	85824.34
Case20	61.02	34938.71	100	134411.07
Case21	46.54	34720.16	100	168388.14
Case24	83.93	66331.85	100	154028.48
Case27	86.07	76429.26	100	166290.33
SUM	948.07	839311.63	1500	1837364.99
AVERAGE	63.20	55954.11	100	122491.00
VARIANCE	205.25	608630602.91	0	1325636396.42

Fig.1. Average of X-axis and Y-axis movement

Fig.2. Case 4 X-axis and Y-axis movement

The horizontal axis represents diagnostic time. Among the diagnoses made by experienced physician and beginner physician diagnosticians, the diagnosis completion time of the one with the longest diagnosis time was set to 100 and normalized. The vertical axis represents the cumulative distance of eye movements until the end of the diagnosis, calculated by dividing all eye movements in the screen where the examination image is displayed in Tobii Pro Lab by the coordinates of the X-axis and Y-axis.

A comparison of the results for 15 cases revealed the following three points.

- Experienced physician require an average of 61% of the diagnostic time of beginner physician for all 15 cases. (Table 1, Fig.1).
- Experienced physician will require an average of 49% of the cumulative gaze distance of the beginner physician for all 15 cases except Case 4 (Fig.2).

- Experienced physician never exceeded the cumulative distance of the beginner physician's gaze for more than half of the diagnostic time in all 15 cases except Case 4 (Fig.2).

Comparison of scrolling speed (Z-axis movement) of imaging test results

Fig.3 below is representative Figure that shows a comparison of the representative Z-axis gaze shift of the gaze when the same case was diagnosed by an experienced physician and a beginner physician.

Fig. 3. Z-axis movement

The following two points were clarified.

- Experienced Physician move back and forth 1~1.5 times between imaging test results until beginner physician confirm all imaging test results (Fig.3).
- Compared to the Beginner Physician, who tended to move back and forth along the Z-axis at many points and was not able to narrow down the range of attention, the Experienced Physician tended to identify the points of attention and gradually narrow down the range of attention along the Z-axis (Fig.3).

6. Discussion

The purpose of this paper is to contribute to human resource development and medical AI development by clarifying factors that lead to the speed of making a diagnosis through a cognitive approach, such as measuring the eye movement of experienced physicians and beginner physicians during the interpretation of radiograms.

In analyzing the gaze, this paper hypothesized that Experienced physicians would have a narrower range of eye movement when searching for the presence or absence of symptoms and the location of symptoms than beginner physicians.

As a result of the analysis, it can be said that the hypothesis is valid according to the following two analysis results for the movement of the X-axis and Y-axis.

- Experienced physician will require an average of 49% of the cumulative gaze distance of the beginner physician for all 15 cases except Case 4 (Fig.2).

Taiki Sugimoto, Jun Nakamura, Yoshinobu Ishiwata

- Experienced physician never exceeded the cumulative distance of the beginner physician's gaze for more than half of the diagnostic time in all 15 cases except Case 4 (Fig.2).

In addition, about Z-axis,

- Compared to the Beginner Physician, who tended to move back and forth along the Z-axis at many points and was not able to narrow down the range of attention, the Experienced Physician tended to identify the points of attention and gradually narrow down the range of attention along the Z-axis (Fig.3).

This is related to the hypothesis that the range of eye movement of Experienced physicians is narrower than the range of eye movement of beginner physicians.

However, in all the cases investigated in this paper, the number of eye movements back and forth on the Z-axis for Experienced physicians was greater than that for beginner physicians Thus, in addition to substantiating the hypothesis, the experienced physicians made fewer eye movements between the X-axis and Y-axis than the beginner physicians, but made more eye movements in the Z-axis, suggesting that the experienced physician may perceive and diagnose imaging test results in a three-dimensional rather than a planar manner.

These findings and discussions in this paper will lead to human understanding through the skilled physician's diagnosis and contribute to the future development of human resource training and medical AI.

7. Conclusions

The purpose of this paper was to contribute to human resource development and medical AI development by focusing on the differences in eye movement between experienced physicians and beginner physicians during diagnosis and to identify factors that lead to the speed at which they make a diagnosis.

As a result of analyzing eye movement during diagnosis in two ways: cumulative distance of eye movement (movement on the X-axis and Y-axis) and scrolling speed of imaging test results (movement on the Z-axis), it was found that experienced physician was faster than beginner physician in making a diagnosis, moving less distance on the X and Y axes, and usually moving less than the distance of beginner physician. On the Z-axis, however, they tended to identify the area of focus and gradually narrow the range of attention but traveled more distance and in a shorter time than the beginner physician.

Next, the limitations of this paper are discussed. Due to technical problems, this paper has not been able to analyze the Z-axis shift of all cases. Besides improving the reliability of the analysis results, it would be desirable to continue analyzing the Z-axis shift, which may lead to new discoveries. In addition, this paper only compares one experienced physician and one beginner physician each. In the future, it is desired to increase the number of subjects and to include intermediate physicians who are between the levels of the two physicians in this paper to increase the reliability of the results.

Finally, in terms of future prospects, this paper will share the results of this analysis with physicians, and by deepening the discussion, explore effective human resource development methods and way of developing medical AI.

8. Acknowledgements

References

1. H. Ichikawa, H. Yamaguchi, T. Noda, and K. Tanaka, The effect of CT interpretation assistance by radiological technologists in off hours emergency medical care. Journal of Japanese Society for Emergency Medicine 17(4), 2014, pp. 535-542
2. H. Mitomo, T. Nakayama, K. Ashizawa, C. Endo, T. Kobayashi, M. Sato, K. Shibuya, T. Sobue, D. Takenaka, K. Nishii, M. Harada, S. Maeda, Y. Maruyama, H. Miura, K. Murata, and M. Sagawa, Nationwide Questionnaire on the Actual Status of an Interpretation System for Lung Cancer Screening by Chest Radiography. The Japan Lung Cancer Society 58 (4), 2018, pp. 243-251
3. M. Terasaki, Y. Terasaki, and A. Shimizu, Artificial Intelligence (AI) and Pathological Diagnosis. Journal of Nippon Medical School 19 (2), 2023, pp. 72-77
4. E. R. Smith, Model of social inference processes. Psychological Review 91(3), 1984, pp. 392–413
5. M. A. Zarate and E.R. Smith, Person Categorization and Stereotyping. Social Cognition 8(2), 1990, pp. 161–185
6. R. A. Monty, and J. W. Senders, Eye Movements: On-line measurement, analysis, and control. Hillsdale: Psychology Library Editions: Perception, 1976
7. T. Katou, and T. Fukuda, Visual Search Strategies of baseball batters during the preparatory phase of batting. The Japanese journal of ergonomics 38(6), 2002, pp. 333-340

Authors Introduction

Mr. Taiki Sugimoto

He is a 1st year graduate student at the faculty of Economics, Chuo University, Japan. The areas of his research interest are Experimental Economics, Behavioral Economics, Cognitive Science, and Artificial Intelligence.

Dr. Jun Nakamura

He is a professor at the Faculty of Global Management, Chuo University, Japan. Previously he was working as a professor at the Graduate School of Engineering Management, Shibaura Institute of Technology. His main research areas are Cognitive Science and Humanities.

Toward Temporal Taste Sensing: Contrasting Human Experience with Machine Perception in Food Appreciation

Hiroki Fxyma

Kobe University, 1-2 Tsurukabuto, Nada-ku, Kobe, Japan

Abstract

This paper explores the temporal aspects of food and beverage appreciation, particularly focusing on the interplay between taste sensors and human sensory experiences. It begins by discussing the fundamental principle of taste sensors, which analyze chemical substances but often fail to capture the complexities of human taste perception. To address this limitation, the concept of "artificial tongues" is introduced, aiming to replicate human taste perception more accurately. The paper then delves into the temporalities inherent in our eating experiences, categorized into three dimensions: the time of the object, the time of embodied cognition, and the time of the phenomenon. Four temporal aspects of the phenomenon are proposed: non-temporal, linear temporality, circular temporality, and point temporality, which are shaped by the eater's attention. The discussion further examines the motivations behind the temporality of food, highlighting cognitive and physical constraints such as memory and oral capacity. Additionally, the paper investigates temporalities in dessert appreciation, using the example of a strawberry mille-feuille to illustrate how the arrangement of ingredients and the sequence of consumption create unique temporal experiences. Finally, it concludes by emphasizing the need for philosophical and theoretical advancements to bridge the gap between human sensory experiences and machine sensing, thereby enhancing our understanding of taste perception in both realms.

Keywords: Temporality, Artificial tongues, Taste sensors, Aesthetics

1. Introduction

1.1. *Taste sensor and the temporality*

The basic principle of taste sensors is to analyze chemical substances that are taste components. Sugar meters and salt meters are the most basic of such sensors. However, a precise analysis of the chemical composition of food does not represent the taste that humans perceive. The characteristics of human taste perception include, for example, the fact that humans perceive different chemical substances (sucrose and aspartame) as the same taste stimulus ("sweetness") as well as the inhibitory/synergistic effect of one taste on the perceived intensity of another taste. To reflect such cognitive characteristics of human perception and output "the taste perceived by humans" instead of simply detecting taste substances, sensors are referred to as artificial tongues. As Toko [1] reviews, there are two types of commercialized electronic tongues in the world. One is the taste sensing systems SA402B and TS-5000Z, commonly known as the taste sensor, and the other is the Astree II e-tongue.

These electronic tongues are intelligent sensors that reflect the human taste perception system in their component analysis as information processing, but their detection of temporality is limited to a small portion. The SA402B and TS-5000Z developed by Toko et al. can analyze taste as it is perceived on the tongue (initial taste) and the taste that remains after swallowing (aftertaste). These are computed based on the ease with which taste components detach from the lipid membrane, which is a taste sensor, thereby detecting, for example, a rich aftertaste. However, none of the current taste sensors measure changes in perception over time, such as fatigue with the same taste. Also, effects such as contrast between ingredients (e.g., strawberry after cream) must be measured separately and evaluated by humans.

1.2. *Temporalities in our eating experience*

Taste and smell are gaining status in aesthetics, but theories concerning their basic appreciation are insufficient. Fxyma [2] proposed the basic theory for temporality of food and beverage appreciation. According to the study, the temporality underlying food appreciation is classified into three categories: the time of the object, the time of the embodied cognition, and the time of the phenomenon. Next, Fxyma proposed four aspects of the time of the phenomenon: non-temporal, linear temporality, circular temporality, and point temporality. These four temporalities are not determined a priori but generated by the eater's attention.

2. Motivations of the Temporality of food

Temporality as a phenomenon in the human eating experience arises from cognitive and physical constraints. In this section, I outline the primary constraints: memory

(cognitive constraints) and oral capacity (physical constraints).

The first constraint is memory. Memory is generally divided into sensory memory, short-term memory, and long-term memory. The retention period of sensory memory varies by sensory modality, but is short, ranging from less than one second to several seconds. In contrast, its retention capacity is larger than that of short-term memory. Sensory information is then stored as short-term memory, medium-term memory, and then long-term memory. Human senses cannot reproduce stimuli. We can recall pain and taste, but this is not a reproduction of the pain or taste "itself. If it were possible to reproduce the pain itself, we would experience intense pain again every time we explain the situation of a serious injury.

On the other hand, a machine's sensors can reproduce the stimuli as numerical values; the values measured 10 years ago can once again be input into the sensors; the wine opened 10 years ago and the wine opened now can be tasted "at the same time". The human being seems to be comparing the two, but one is the sensor data of the actual wine in front of us, while the other is the edited and partly erased/enhanced long-term memory of the wine of 10 years ago.

This means that the time possessed by the machine is basically point temporality in Fxyma's temporality classification. The present and the past appear simultaneously in the here and now. It is not Linear temporality because there is no passage of time or order. It is also different from Circular temporality because it recreates the same stimulus in the past. Circular temporality is about finding repetition in different experiences. For example, spring is repetitive, but the content of the spring experience is not identical every year. The reproduction of exactly the same stimuli is different from the concept of Circular temporality. Thus, the simultaneous appearance of the here and now as a condensed time of the present and the past is Point Temporality.

This constraining characteristic of human memory to be edited, forgotten, and reinforced over time gives rise to Linear Temporality or Circular temporality. When eating a parfait, the first strawberry, the next taste of cream, and the subsequent strawberries are not experienced separately, but have a linear continuity or a circular temporality structure as repetitions of strawberries. When a person eats a parfait, temporal memory is passed on from one bite to the next; if the same taste continues in the mouth, the person becomes bored, and if the sweet flavor of the cream is followed by the sour jam, there is a contrast. In the repetition of strawberries and cream, the strawberry is presented with a different role each time it appears. Here, more than the summation of the elements, an aesthetic experience emerges.

The second constraint is the physical structure of the nasal and oral cavities. Generally, the volume of the oral cavity is smaller than that of the stomach, and humans divide food during meals, sending the entire amount of

one meal to the stomach in several to several tens of times. In the case of eating the entire amount of food at once, there is no temporality. The division of food itself becomes the basic unit of temporality, such as order and repetition. For example, beef steak is divided into several pieces with a knife, finely crushed with several times the number of chewing, and sent to the stomach.

Human meals are embedded in multiple cyclical temporalities. The time span ranges from seconds, such as repeated chewing and swallowing, to levels such as three meals a day (rise and fall of blood sugar in some hours), and complex cyclical cycles such as food aversion/preference learning (in some hours to days), or seasonal foods (in a year). When analyzing food with taste sensors, it is liquefied with a mixer and the composition of flavor components is analyzed. Here, temporalities such as division with a knife or repetitive chewing are disregarded, and analysis is conducted as timeless, ignoring point time or time passage.

The analysis times increase significantly when the number of steps in the implicit integral is expanded. In the implicit method, the solutions of simultaneous equations are obtained by inserted Newton-Raphson method, and numerical substitutions into symbolic matrix within this process require the computation processing time. The symbolic matrix is enlarged by increasing the number of implicit stages, which takes more computation time. Therefore, it is important to implement an efficient substitution process into a symbolic matrix for the implementation of higher order implicit methods in MBD analysis.

3. Temporalities in Dessert Appreciation

In this section, I will examine the process by which humans experience food, using one dish of dessert as an example. This dessert shown in Fig. 1 is provided at the pastry shop *Assiette Dessert Maruyama*, located in Nishinomiya City, Japan.

Fig.1 The Strawberry Millefeuille of *Assiette Dessert Maruyama.*

The strawberries scattered on the plate play different roles in the beginning and the end of the dish. Of course, they also play a role in giving an impact at the moment

they are presented. However, this function is a visual one. The strawberries eaten at the beginning of the meal function to tell the taste of the strawberries or to introduce the flavor of the strawberries used as an ingredient. In order to understand the function of the strawberries in the latter stages, it is necessary to interpret the flavor of the millefeuille in the main body of the dish. The millefeuille is characterized by the horizontal structure of the cut layers, rather than the more common style in which layers of pie and cream are piled up vertically.

In a typical layered millefeuille, one must cut the pastry with a fork to achieve the intended size, which is not only cumbersome but also risks disrupting the balance between the pie pastry and the strawberries by causing cream to spill out. However, in *Maruyama*'s millefeuille, a bite-sized portion is pre-cut, allowing one to enjoy the pie pastry, cream, and strawberries atop without compromising their balance.

In *Maruyama*'s mille-feuille, the meticulous calculation behind this bite-sized composition becomes fully evident as one progresses through consuming it. As the layers of cream and pastry are meticulously arranged side by side, defying the conventional definition of a mille-feuille, scooping up a set with a fork automatically results in an accompanying strawberry, completing a mouthful. The pastry is not of the rigid variety but rather consists of delicate, thin layers stacked upon one another. When pressed against the custard cream with the tongue and teeth, the pastry dissolves effortlessly without requiring a bite, instantly enveloping the tongue in the smoothness of the cream. There is no sensation of strawberry juice separating from the custard cream and flowing out. While the strawberries provide a fibrous texture and a refreshing acidity to the tip of the tongue, the carefully controlled juice harmoniously melds with the pastry and cream, creating an impression of "strawberry custard cream" rather than separating out.

The perfect harmony of pastry, cream, and strawberries creates an impression of "strawberry custard cream" where everything melts together in the mouth, to the extent that one might almost forget the awareness of "eating strawberries" as they progress through the middle and later stages. To prevent this forgetting, the role of the strawberries scattered on the plate from the middle to the end comes into play. The strawberries scattered on the plate ensure a return to the main theme that might otherwise be lost due to the remarkable harmony.

This mille-feuille is likely intended to be consumed from left to right. Evidence of this can be found in the absence of cream at the right end, indicating a progression. In the overall linear progression of time, each bite-sized millefeuille forms a condensed, Point Temporality through the perfect harmony of its ingredients. While the repetition of millefeuille layers from left to right introduces a recursive Cyclical Temporality, there comes a moment when continuity threatens to turn into monotony and a sense of timeless monotony (i.e., Atemporality). At this juncture, the strawberries scattered on the plate serve as "privileged moments [3], bringing the consciousness of the eater back to the cyclical time.

4. Conclusion

This paper examines the differences between human sensory experiences and machine sensing through the description of the experience of eating a dessert on a plate. I argue that the human experience of eating is temporal in nature, primarily motivated by two constraints: the limitations of memory and oral capacity. In contrast, machine taste sensing, without these constraints, is an experience characterized by "point time" or "timelessness," where past, present, and future information are analyzed homogeneously.

The current limitations of sensing in reflecting the temporal aspects of the human eating experience, especially linear or cyclical time, do not stem from inadequacies in sensor capabilities or the scientific knowledge supporting them.

Rather, the deficiency lies in the philosophical and theoretical understanding of what should be sensed. We lack theories to explain the experience and phenomena of consuming a dessert or savoring a glass of wine, which limits taste sensing to mere component analysis. Taste sensors themselves are capable of detecting taste components more "accurately" than humans, "objectively" across various conditions, and with an unparalleled sensitivity to subtle taste components. If taste sensors are to transcend mere component analysis and aim to reproduce "eating" as an intellectual behavior of machines, what is required is philosophical and theoretical leadership in science.

References

[1] K. Toko, Y. Tahara, M. Habara, Y. Kobayashi, and H. Ikezaki, "Taste Sensor: Electronic Tongue with Global Selectivity," in *Essentials of Machine Olfaction and Taste*, 1st ed., T. Nakamoto, Ed., Wiley, 2016, pp. 87–174. doi: 10.1002/9781118768495.ch4.

[2] H. Fxyma, "The Temporality of the Aesthetic Appreciation of Food and Beverages," *Contemp. Aesthet.*, 2024.

[3] E. Souriau, "Time in the plastic arts," *J. Aesthet. Art Crit.*, vol. 7, no. 4, pp. 294–307, 1949.

Authors Introduction

Dr. Hiroki Fxyma

He is currently an Assistant Professor at Kobe University. He completed his doctoral and master's degrees at Keio University. His current research topic is the aesthetic appreciation of taste and its cognitive process.

Deep Learning Based Prediction of Heat Transfer Coefficient Using Spectrogram Images from Boiling Sound

Fuga Mitsuyama

Department of Mechanical and Control Engineering, Kyushu Institute of Technology,
1-1 Sensui-cho, Tobata-ku, Kitakyushu, Fukuoka, 804-0015, Japan

Ren Umeno

Department of Mechanical and Control Engineering, Kyushu Institute of Technology,
1-1 Sensui-cho, Tobata-ku, Kitakyushu, Fukuoka, 804-0015, Japan

Tomohide Yabuki

Department of Mechanical and Control Engineering, Kyushu Institute of Technology,
1-1 Sensui-cho, Tobata-ku, Kitakyushu, Fukuoka, 804-0015, Japan

Tohru Kamiya

Department of Mechanical and Control Engineering, Kyushu Institute of Technology,
1-1 Sensui-cho, Tobata-ku, Kitakyushu, Fukuoka, 804-0015, Japan
E-mail: mitsuyama.fuga621@mail.kyutech.jp, umeno.ren696@mail.kyutech.jp,
yabuki.tomohide556@mail.kyutech.jp, kamiya@cntl.kyutech.ac.jp

Abstract

Cooling methods based on boiling have attracted attention as a thermal solution for electronic equipment. In this situation, it is necessary to measure the heat transfer coefficient (HTC) to design more efficient cooling systems. In this paper, we propose a method to predict of the HTC from boiling sound data using deep learning techniques. The accuracy improved by 1.12% compared to the conventional method through the development of Convolutional Neural Network (CNN) incorporate Convolutional Block Attention Module (CBAM).

Keywords: Boiling Sound, Heat Transfer Coefficient (HTC), Convolutional Neural Network (CNN), CBAM.

1. Introduction

The heat generation density of CPUs and GPUs is increasing year by year as they become more highly integrated [1]. Under these circumstances, the cooling method using water boiling is currently attracting attention. Boiling cooling has a high heat transfer coefficient (HTC) compared to gas forced convection cooling using heat sinks and fans [2], which is still in use today. Furthermore, since water serves as the boiling medium, it is environmentally friendly and economical. Since the boiling absorbs heat through the evaporation of the boiling liquid and the convection induced by the bubble motion [3], the HTC significantly depends on the number and size of the bubbles. In fact, the measured HTC shows a significant change when the number and shape of the bubbles are varied, as shown in Fig. 1. Therefore, it is important to accurately measure the HTC to develop more efficient and reliable cooling systems.

This paper proposes a method for predicting the HTC from spectrogram images of the boiling sound using Convolutional Neural Network (CNN). This method is based on the idea that there is a relationship between boiling sound and HTC. We conduct a classification task by categorizing the HTCs obtained through experiments into five classes. Then, we apply Convolutional Block

Fuga Mitsuyama, Ren Umeno, Tomohide Yabuki, Tohru Kamiya

Attention Module (CBAM) [4] to a CNN model to improve the classification accuracy.

2. Method

2.1. *Creation of dataset*

We experimented with sound data acquisition to create a dataset of spectrograms. Then, this acquired sound data was transformed into a spectrogram using the Fast

Fig. 1. Boiling state transitions and associated changes in the quantity and shape of bubbles

Fourier Transform (FFT) [5] to generate a 200 x 200 [pixels] image. Fig. 2 gave an overview of the

Fig. 2. Overview of experimental setup for sound data acquisition

experimental setup. A 1-mm-thick, 10-mm-square copper plate soldered to the top of a copper heater block into which cartridge heaters are inserted was used as a heat-transfer wall. Thermocouples inserted in the copper block were used to the surface measure temperature and heat flux. Before the boiling experiment, the pure water used as the boiling medium was degassed for more than 1 hour. Boiling sound was measured using Miniature Hydrophone Type 8103® [6] at a sampling frequency of 50 [kHz]. Table 1 shows the number of data for each the HTC class.

2.2. *Classification by CNN*

The model is based on the CNN model developed by Tabata et al [7] and Sinha, K.N.R et al [8]. A basic block of CNN in Fig. 3 consists of a convolution layer (with a kernel size of 3 x 3 and 96 channels), batch normalization, ReLU functions (as activation functions) and average pooling (with a kernel size of 2 x 2). Then, four stacks of these blocks downsample the input images and construct

a robust model against noise. We specify the labels for each class ranging from 10–20, 20–30, 30–40, and 40–50 [kW/m^2/K]. These label numbers represent the HTC.

2.3. *Convolutional Block Attention Module (CBAM)*

In order to improve the accuracy of the conventional CNN model, we incorporate CBAM and compare its accuracy with that of the conventional model. The structure of the CNN model with CBAM is shown in Fig. 3. The CBAM architecture is based on the paper by S. Woo et al [4], where it is found that higher accuracy results can be obtained by connecting the Channel Attention Module and the Spatial Attention Module in series. The Channel Attention Module is a mechanism that learns and weights the importance of information in each channel based on the correlation of features between channels. The Spatial Attention Module is a mechanism that learns and weights the importance of information in the spatial axis, allowing the extraction of important information in the spatial direction in the feature map.

3. Experimental Result and Discussion

3.1. *Evaluation*

The evaluation method is the Five-Fold Cross Validation and average accuracy, which is the average of the accuracy of each class in the five-class classification and the average of the accuracy for all data in each dataset.

Table 1 The number of data in each class

Class [kW/m^2/K]	Number
10–20	880
20–30	440
30–40	440
40–50	900
50–60	420
Total	3080

Fig. 3. Configuration of the proposed model with CBAM

(Accuracy is defined as the proportion of correct and predicted labels that match).

3.2. *Result*

The results of the proposed method using the datasets obtained from the experiments in Section 2.1 are shown in Table 2 and compared with the those of the conventional method. The accuracy in each class decreased by 1.09% in the 10–20 [kW/m^2/K] class and by 0.95% in the 30–40 [kW/m^2/K] class but increased significantly in the other classes. In particular, the accuracy in the 40–50 [kW/m^2/K] class improved by 3.38%. The average accuracy of the proposed method is 91.61% compared to 90.49% for the conventional method. That's an improvement of 1.12%.

3.3. *Discussion*

The main reason why the average accuracy was improved except for the 10–20 [kW/m^2/K] and 30–40 [kW/m^2/K] classes is that CBAM was able to extract the response of the constant-frequency component of the horizontal line, which is unique to each HTC. That allowed CBAM to maintain high accuracy even for noisy data sets. On the other hand, the inaccuracy could be attributed to the presence of various noises added during the creation of dataset, which aimed to construct a model that can be robust to noise. We have concluded that the spectrograms obtained from the experiments in Section 2.1 may not reflect the unique frequency characteristics found at each HTC. The spectrograms shown in Fig. 4

provide typical data for these classes, and we can recognize horizontal lines of constant frequency near the center of the images. On the other hand, Fig. 5 is an example of images that could not be classified correctly, showing that the features appeared in Fig. 4 are not clearly represented or are faded due to noises in Fig. 5. The CBAM mechanism may have made it more difficult to classify images like those in Fig. 5 by over-extracting features from data with typical features like those in Fig. 4.

4. Conclusion

In this paper, the incorporation of CBAM into the conventional CNN model decreased accuracy by 1.09% in the 10–20 [kW/m^2/K] class and by 0.95% in the 30–40 [kW/m^2/K] class, but significantly increased accuracy in the other classes, with the average accuracy improving by 1.12%. Future work will include increasing the number of experiments to build a CNN model that is more robust to noise, as one reason for the decrease in accuracy is the limited number of datasets. In addition, we will identify features common to each class and build a model that takes these features into account, with the aim of predicting the HTC in more detail and further improving accuracy.

Fuga Mitsuyama, Ren Umeno, Tomohide Yabuki, Tohru Kamiya

Table 2 Results of five classes classification from spectrograms using two CNN model

	10 ~ 20	20 ~ 30	30 ~ 40	40 ~ 50	50 ~ 60	Average Accuracy [%]
CNN	95.01	84.90	88.45	90.08	89.80	90.49
CNN + CBAM	93.92	86.90	87.50	93.46	91.15	91.61

Fig. 4. Typical spectrogram of 10–20 and 30–40 [kW/m^2/K] class

Fig. 5. Spectrogram of failed classification in the 10–20 and 30–40 [kW/m^2/K] class

Reference

1. N. Unno, "[Towards t he Realization of Boiling Cooling Devices Using Bubble Refinement Boiling] *Kihoubisaikahuttou wo tukatta huttoureikyakudebaisu jitugen ni mukete*", JSME TED Newsletter, No.89, pp. 2-4, 2019. (In Japanese)
2. M. Suzuki, K. Kawaguchi, T. Terao, E. Tanaka, H. Tanaka, "Boiling Refrigerant Type Compact Cooling Unit for Computer Chip (Comparison with Other Type Cooling Unit for Computer Chip)", The Japan Society of Mechanical Engineers, pp. 97-98, 2000.
3. T. Tanaka, K. Miyazaki, T. Yabuki, "Observation of heat transfer mechanisms in saturated pool boiling of water by high-speed infrared thermometry", International Journal of Heat and Mass Transfer, 170, 121006, 2021.
4. S. Woo, J. Park, JY. Lee, IS. Kweon, "Cbam: Convolutional block attention module", Proceedings of the European conference on computer vision (ECCV), pp. 3-8, 2018.
5. H.J. Nussbaumer, "Fast Fourier Transform and Convolution Algorithms", Springer-Verlag, pp. 80-111, 1982.
6. MrHBK, "MINIATURE HYDROPHONE TYPE 8103", https://www.bksv.com/-/media/literature/Product-Data/bp0317.ashx, Nov. 22, 2023.
7. M. Tabata, H. Lu, T. Kamiya, S. Mabu, S. Kido, "Automatic Classification of Respiratory Sound Considering Hierarchical Structure", 2022 The 22nd International Conference on Control, Automation and Systems (ICCAS), pp. 2-3, 2022.
8. Sinha, K.N.R et al., "Deep Learning the Sound of Boiling for Advance Prediction of Boiling Crisis", Cell Reports Physical Science, pp. 3-12, 2021.

Authors Introduction

Mr. Fuga Mitsuyama

He received his Bachelor's degree in Engineering in 2023 from the Faculty of Engineering, Kyushu Institute of technology in Japan. He is currently a master student in Kyushu Institute of Technology, Japan.

Mr. Ren Umeno

He received his Master's degree in Engineering in 2023 from the Faculty of Engineering, Kyushu Institute of technology in Japan. He is currently a doctoral student in Kyushu Institute of Technology, Japan.

Dr. Tomohide Yabuki

He is an Associate Professor in the Department of Mechanical and Control Engineering at Kyushu Institute of Technology and is a PRESTO researcher of Japan Science and Technology Agency (JST). His research interests are in the areas of boiling heat transfer, MEMS thermal measurement techniques, and single cell thermal analysis.

Dr. Tohru Kamiya

He received his B.A. degree in Electrical Engineering from Kyushu Institute of Technology in 1994, the Masters and Ph.D. degree from Kyushu Institute of Technology in 1996 and 2001, respectively. He is a professor in the Department of Mechanical and Control Engineering at Kyushu Institute of Technology. His research interests are focused on image processing, medical application of image analysis, etc. .

A Study on Classification of Faulty Motor Sound Using Convolutional Neural Networks

Md Shafayet Jamil
Graduate School of Engineering, University of Miyazaki, 1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan

Praveen Nuwantha Gunaratne
Interdisciplinary Graduate School of Agriculture and Engineering, University of Miyazaki, 1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan

Hiroki Tamura
Faculty of Engineering, University of Miyazaki, 1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan
Email: z322t01@student.miyazaki-u.ac.jp, ti20060@student.miyazaki-u.ac.jp, htamura@cc.miyazaki-u.ac.jp

Abstract

Classification of sound has its usage in various fields in today's world. In this paper we will go through the sound classification techniques for the detection of faulty machines with the help of the sound data produced by the machine. The focus is towards determining the pertinency of audio classification methods to detect faulty motors by their sounds; both in noisy and noise-free scenarios; so that the requirement of human inspection can be reduced in factories and industries. Noise reduction plays such an important role in improving accuracy of detection some researchers simulated data by adding noise for benchmarking their models. Hence noise reduction is widely used in audio classification tasks. Among various available methods, we have implemented an autoencoder for noise reduction. We have conducted the classification tasks on both noisy and denoised data using Convolutional Neural Network. Accuracy of classification on the data denoised using autoencoder is compared with the noisy ones. For classification, we used spectrogram, Mel-frequency cepstral co-efficient (MFCC) and Mel-spectrogram images. These processes yield promising results in distinguishing faulty motors by their sound.

Keywords: Faulty motor detection, autoencoder, convolutional neural network (CNN), audio classification

1. Introduction

Audio data is one of the most common multimedia types used in almost every sector of modern life throughout the past decades. The availability of audio recording devices has also been increased and the usage of audio data can be found in security surveillance systems, health monitoring systems, and various autonomous systems along with the inseparable usage in daily human life [1], [2], [3], [4]. Successful utilization of audio resources depends on the efficiency of classification (process of identification of pre-defined label for audio signals [5]), transcription, and perception of underlying contents through a bunch of transforming algorithms and machine learning approaches carried out on various sources of audio data, as in, voices, music, environmental sound, traffic sound etc. [1]. Our research aimed at distinguishing the faulty motors from the normal ones by means of the soundwave processing followed by training with convolutional neural network so that it can be used in factories to reduce the requirement for human inspection, and in preventing hazards. As the adequacy of classification often gets drastically affected by noise [6], we also have incorporated an autoencoder based noise reduction model in parallel to the classification process. CNN has many proven outcomes in the field of audio classification [7], [8], [9]. Convolutional neural networks have been used widely for classification since it has been found competent for both image classification e.g. vehicle classification [9] and sound classification [10] on various fields namely environmental sound classification (ESC) [11], animal sound classification [10], [12], vehicle classification [7], [13], emotion detection [14], violence detection [15] etc.

In this paper we will be showing results of applying Lenet-5 on image transforms of the audio signal. Lenet-5 was developed on the context of classification of images especially text [16]. However, while we can transform sound data into various image representations it can be an auspicious tool for classification of audios as well. We found approaches with applying CNN for the task of Environmental sound classification [11] and deploying Lenet for acoustic scene detection [17]. Despite the time consumption issues with large scale datasets CNNs are well suited for classification tasks within a few parameters [18]. For our research we have used Lenet-5 for classifying audios by means of transforming audios into feature images. We have used 3 types of feature-

Md Shafayet Jamil, Praveen Nuwantha Gunaratne, Hiroki Tamura

Fig. 1 Overview of faulty motor classification

images (Mel-Frequency Cepstral Coefficient, Mel-spectrogram, and Spectrogram) from both the noise-free original and smoothed (with an Autoencoder) audio chunks. Finally, we fed the classifier with both (noisy and smoothed) type of feature images separately and compared the resultant accuracy for noise association and the above mentioned 3 types of feature images as well. The process yields promising results. The whole process is shown in Fig. 1.

2. Dataset and Preprocessing

Our dataset is comprised of 890 samples of motor sounds, divided into two classes, 645 normal sound and 245 samples of abnormal sound; both are of variable length audio file in the form of '.wav' file at 48000Hz. The audio data was acquired using electronic stethoscopes. In preprocessing we split 3 seconds long chunks from the audio samples, having a length of 3 * 48000; and the samples which's lengths are less than 144000 (3 * 48000) were zero padded. Finally, each datapoint is normalized in such a way that the underlying values of the data stream contain real numbers ranging from -1.0 to 1.0 only.

3. Methodology

For classifying the uniform length audios of normal and abnormal motor sound; derived from the preprocessing; we divided the task into 2 standalone pipelines. The first one involves smoothing with autoencoder and the second one doesn't incorporate any noise reduction measure. In the later sections we have discussed these steps.

3.1. Noise reduction

Autoencoders are used for noise reduction in various audio recognition experiments [19]. An autoencoder is a composition of an encoder and a decoder block having convolutional and transpositional layers respectively. It can reconstruct the input without noise through a series of compressions and then decompressions [20]. We optimized the autoencoder model described in [20] by accompanying 4 convolutional layers and 4 transposition or deconvolutional layers as shown in Fig. 2. Also, in the transposition layers hyperbolic tangent activation (tanh) is used instead of PRelU for keeping the negative elements [21] in the resultant audios and to preserve a near-zero mean [22]. The results showing the smoothing with both types of activations are shown in Fig. 3 and Fig. 4 by superimposing the smoothed wave (yellow) onto the original one (blue).


```
Input
Encoder
Conv1D (128, k =7)
Conv1D (32, k =5)
Conv1D (16, k =3)
Conv1D (8, k=3)
Decoder
Transpose (8, k =3)
Transpose (16, k=3)
Transpose (32, k=5)
Transpose (128, k=7)
Output
```

Fig. 2 Autoencoder Model

3.2. Feature extraction

The selection of features has significant influence on classification accuracies [4]. That is why we fed both the noisy and smoothened audios through a transformation process which yield 3 types of feature images- a. Spectrogram, b. Mel-spectrogram, c. Mel-frequency cepstral coefficient (MFCC). We used a Python-based library called Librosa [23] for these processing.

a. Spectrogram
Transforming one dimensional audio signal into matrix like representation makes them suitable for training with neural networks [4]. Spectrograms are such visual depictions of frequency-wise strength of an audio which can improve classification results [24]. A sample spectrogram of a 3-second-long normal motor sound is shown in Fig 5 (a).

b. Mel-spectrogram
Mel scaled spectrograms can represent the signal analogous to the human auditory system. The formula for converting f Hz into m mel is $m = 2595 \log_{10}(1 + \frac{f}{700})$ [4]. The mel-spectrogram of a 3 second sound of abnormal motor looks like Fig. 5 (b).

c. Mel-frequency cepstral coefficient (MFCC)
MFCC incorporates Discrete Cosine Transform and has a compressed representation signal [4] so it is quite useful for training audio features. We have put an MFCC image of a smoothened normal motor sound (3 second) in Fig. 5 (c).

3.3. Classification

Image classification networks also have good results on sound data [24]. So, we trained these feature images with a CNN based model Lenet-5 (shown in Fig. 6) [16] which was proposed for classifying handwritten characters with requiring low preprocessing. Our implementation of Lenet-5 has an identical structure. It has 2 convolutional layers working as feature extractors followed by 3 fully connected layers working as classifier, as it appears in Fig. 7.

Fig. 3 Smoothing with autoencoder (tanh activation)

Fig. 4 Smoothing with autoencoder (PRelU activation)

a)

b)

c)

Fig. 5. Feature image samples a) Spectrogram b) Mel-spectrogram c) MFCC

Fig. 6 Generic architecture of Lenet-5 [16]

```
Feature extractor
C1: Conv2d (3, 32, k=5)
S2: AvgPool2d (k=2)
C3: Conv2d (32, 64, k=5)
S4: AvgPool2d (k=2)
Classifier
C5: Convolution (128)
Gaussian Con.: Linear (84)
Output: Linear (2)
```

Fig. 7 Structure of the implemented Lenet-5 model.

Fig. 8 Loss function of the autoencoder

4. Results

The results derived from the various experiments can be divided into 4 parts. Firstly, the performance of the autoencoder in terms of loss function (refers to Fig. 8). We have trained the autoencoder model up to 1.6% of loss using the metrics of mean squared errors (MSE), but the autoencoder model with 2.3% loss yield the best result during classification. Secondly, the straight-forward performance comparison between MFCC, Mel-spectrogram, spectrogram image-based classification of the model. Thirdly, comparison between the performance with and without using autoencoder smoothing, results of second and third steps are summarized in Table. 1. Fourthly, reviewing the impact of different dataset combinations, amounts data, and biases as shown in Table. 2. The best accuracies (%) found in these three steps are shown in Table 1 and Table 2. To put it in a nutshell, we can deduce that at least 97% accuracy could be expected from this approach in classifying 2 class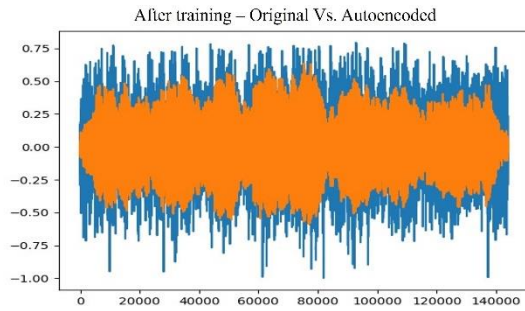es of motors (normal and abnormal) where a totally unknown type of motor sound is not present in the testing data. And as for noise reduction, it is not evident to have positive influence on classification accuracy with the current configuration of the model of the autoencoder that has been used for smoothing.

5. Conclusion

Sound classification with the help of CNN could be a feasible solution to various realistic problems especially in automation of maintenance and monitoring scenarios in large scale factories. It can reduce the need for manpower and improve Mean time to response (MTTR) by providing automated detection of faulty machines in industries. The scope of this paper could also be easily extended to other fields where a small number of classes are needed to be classified with high precision.

Table 1. Performance with and without autoencoder

Feature	With Autoencoder	Without Autoencoder
Spectrogram	98.39871898	98.71897518
Mel-spectrogram	99.93327996	99.90659194
MFCC	99.82652789	99.91638471

Table 2. Performance by dividing motors into two groups (motor 1-50 and motor 51-93)

Feature	Motor1–50 normal: Motor 51-93 abnormal	Motor51–93 normal: Motor 1-50 abnormal	Motor1–50 normal: Motor1-50 abnormal	Motor 51–93 normal: Motor 51-93 abnormal
Spectrogram	97.001	97.26	98.86	98.29
Mel-spectrogram	99.7	99.72	99.95	99.88
MFCC	99.9	99.9	99.95	99.99

References

1. L. Gao, K. Xu, H. Wang, and Y. Peng, "Multi-representation knowledge distillation for audio classification," Multimedia Tools and Applications, vol. 81, no. 4, pp. 5089–5112, Jan. 2022, doi: https://doi.org/10.1007/s11042-021-11610-8.
2. C. Clavel, T. Ehrette, and G. Richard, "Events Detection for an Audio-Based Surveillance System," IEEE Xplore, Jul. 01, 2005. https://ieeexplore.ieee.org/document/1521669
3. Y.-T. Peng, C. Lin, M. Sun, and K.-L. Tsai, "Healthcare audio event classification using Hidden Markov Models and Hierarchical Hidden Markov Models," Jun. 2009, doi: https://doi.org/10.1109/icme.2009.5202720.
4. M. Turab, T. Kumar, M. Bendechache, and T. Saber, "Investigating Multi-feature Selection and Ensembling for Audio Classification," International Journal of Artificial Intelligence & Applications, vol. 13, no. 3, pp. 69–84, May 2022, doi: https://doi.org/10.5121/ijaia.2022.13306.
5. Tuomas Virtanen, M. D. Plumbley, and D. Ellis, "Computational Analysis of Sound Scenes and Events." Springer, 2017, https://link.springer.com/book/10.1007/978-3-319-63450-0
6. J. Meyer, L. Dentel, and F. Meunier, "Speech Recognition in Natural Background Noise," PLoS ONE, vol. 8, no. 11, Nov. 2013, doi: https://doi.org/10.1371/journal.pone.0079279

7. K. W. Cheng et al., "Spectrogram-based classification on vehicles with modified loud exhausts via convolutional neural networks," Applied Acoustics, vol. 205, p. 109254, Mar. 2023, doi: https://doi.org/10.1016/j.apacoust.2023.109254.

8. C. Yang, X. Gan, A. Peng, and X. Yuan, "ResNet Based on Multi-Feature Attention Mechanism for Sound Classification in Noisy Environments," Sustainability, vol. 15, no. 14, p. 10762, Jan. 2023, doi: https://doi.org/10.3390/su151410762.

9. M. A. Butt et al., "Convolutional Neural Network Based Vehicle Classification in Adverse Illuminous Conditions for Intelligent Transportation Systems," Complexity, vol. 2021, pp. 1–11, Feb. 2021, doi: https://doi.org/10.1155/2021/6644861.

10. F. Merchan, A. Guerra, H. Poveda, H. M. Guzmán, and J. E. Sanchez-Galan, "Bioacoustic Classification of Antillean Manatee Vocalization Spectrograms Using Deep Convolutional Neural Networks," Applied Sciences, vol. 10, no. 9, p. 3286, May 2020, doi: https://doi.org/10.3390/app10093286.

11. Z. Mushtaq and S.-F. Su, "Environmental sound classification using a regularized deep convolutional neural network with data augmentation," Applied Acoustics, vol. 167, p. 107389, Oct. 2020, doi: https://doi.org/10.1016/j.apacoust.2020.107389.

12. Yin, Y., Tu, D., Shen, W., & Bao, J. , "Recognition of sick pig cough sounds based on convolutional neural network in field situations," Information Processing in Agriculture, Nov. 2020, doi: https://doi.org/10.1016/j.inpa.2020.11.001.

13. M. A. Hedeya, A. H. Eid, and R. F. Abdel-Kader, "A Super-Learner Ensemble of Deep Networks for Vehicle-Type Classification," IEEE Access, vol. 8, pp. 98266–98280, 2020, doi: https://doi.org/10.1109/access.2020.2997286.

14. R. Cai, L. Lu, H.-J. Zhang, and L. Cai, "Highlight sound effects detection in audio stream," Jan. 2003, doi: https://doi.org/10.1109/icme.2003.1221242.

15. S. Pfeiffer, S. Fischer, and W. Effelsberg, "Automatic audio content analysis," Proceedings of the fourth ACM international conference on Multimedia - MULTIMEDIA '96, 1996, doi: https://doi.org/10.1145/244130.244139.

16. Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proceedings of the IEEE, vol. 86, no. 11, pp. 2278–2324, 1998, doi: https://doi.org/10.1109/5.726791

17. Venkatesh Duppada and Sushant Hiray, "Ensemble Of Deep Neural Networks For Acoustic Scene Classification," arXiv (Cornell University), Aug. 2017, doi: https://doi.org/10.48550/arxiv.1708.05826.

18. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Communications of the ACM, vol. 60, no. 6, pp. 84–90, May 2012, doi: https://doi.org/10.1145/3065386.

19. S. Alharbi et al., "Automatic Speech Recognition: Systematic Literature Review," IEEE Access, vol. 9, pp. 131858–131876, 2021, doi: https://doi.org/10.1109/access.2021.3112535.

20. K. Bajaj, D. K. Singh, and Mohd. A. Ansari, "Autoencoders Based Deep Learner for Image Denoising," Procedia Computer Science, vol. 171, pp. 1535–1541, 2020, doi: https://doi.org/10.1016/j.procs.2020.04.164.

21. Swalpa Kumar Roy, S. Manna, Shiv Ram Dubey, and B. B. Chaudhuri, "LiSHT: Non-Parametric Linearly Scaled Hyperbolic Tangent Activation Function for Neural Networks," arXiv (Cornell University), Dec. 2018, doi: https://doi.org/10.48550/arxiv.1901.05894.

22. S. R. Dubey, S. K. Singh, and B. B. Chaudhuri, "Activation functions in deep learning: A comprehensive survey and benchmark," Neurocomputing, vol. 503, pp. 92–108, Sep. 2022, doi: https://doi.org/10.48550/arXiv.2109.14545.

23. B. McFee et al., "librosa: Audio and Music Signal Analysis in Python," Proceedings of the 14th Python in Science Conference, 2015, doi: https://doi.org/10.25080/majora-7b98e3ed-003.

24. Boddapati, V., Petef, A., Rasmusson, J., & Lundberg, L., "Classifying environmental sounds using image recognition networks," Procedia Computer Science, vol. 112, pp. 2048–2056, Jan. 2017, doi: https://doi.org/10.1016/j.procs.2017.08.250.

Authors Introduction

Mr. Md Shafayet Jamil

He received his bachelor's degree in science in 2016 from the Institute of Information Technology, Jahangirnagar University, Bangladesh. He is currently a master's student in University of Miyazaki, Japan.

Mr. Praveen Nuwantha Gunaratne

He received his Bachelor's degree in Engineering in 2018 from the Faculty of Engineering, University of Moratuwa, Sri Lanka. He is currently a Doctoral student in University of Miyazaki, Japan

Prof. Hiroki Tamura

He received the B.E. and M.E. degree from Miyazaki University in 1998 and 2000, respectively. From 2000 to 2001, he was an Engineer in Asahi Kasei Corporation, Japan. In 2001, he joined Toyama University, Toyama, Japan, where he was a Technical Official in the Department of Intellectual Information Systems. In 2006, he joined Miyazaki University, Miyazaki, Japan, where he was an Assistant Professor in the Department of Electrical and Electronic Engineering. Since 2015, he is currently a Professor in the Department of Environmental Robotics. His main research interests are Neural Networks and Optimization Problems. In recent years, he has had interest in Biomedical Signal Processing using Soft Computing.

Design and Development of a Flexible Active Ankle Joint Orthosis for Locomotion Assistance

Praveen Nuwantha Gunaratne
Interdisciplinary Graduate School of Agriculture and Engineering, University of Miyazaki, 1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan

Hiroki Tamura
Faculty of Engineering, University of Miyazaki, 1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan
Email: ti20060@student.miyazaki-u.ac.jp, htamura@cc.miyazaki-u.ac.jp

Abstract

Active ankle joint orthosis is an artificial wearable device that is closely fitted to the human ankle and combines human intelligence and intentions with the powered robot joints. However, most of the existing designs have only focused on one degree of freedom (DOF) of the ankle joint: namely plantarflexion-dorsiflexion. Alternatively, a modular power-assist flexible active ankle orthosis is proposed, addressing anthropomorphic architecture of design, and supporting multiple DOFs of ankle for locomotion assistance of the physically weak individuals. The proposed ankle joint orthosis was designed with the aid of CAD tools followed by fabrication of a working model. Electrically powered systems with gear transmissions were adapted to support the principal motions during normal Gait. In parallel, the design of the control algorithm was carried out based on the EMG signals. The future advancements have been focused on developing a novel control method that provides sufficient flexibility to assist a wide variety of lower-limb motions.

Keywords: Active ankle orthoses, power-assist robots, anthropomorphic design, joint axes mapping, rehabilitation

1. Introduction

Every human being would prefer to spend their entire lifetime as an independent individual without becoming dependent on others for locomotion. According to statistics around the globe, the elderly population exceeds 10% of the total population and conversely, the number of people suffering from neurological or muscular disorders is on the rise. As such, the medical industry has shown a growing interest in using exoskeleton robots for restoring patients' mobility and to alleviate health care issues arising from locomotion difficulties.

With the continuous development in biomedical engineering field over the past few decades, the robotic orthosis devices have been used for many different applications. The active orthosis/ robotic exoskeleton is an artificial wearable device that is powered by the means of electric motors, pneumatics, hydraulics, or linear actuation methods. In the current market, active orthoses are classified as upper extremity systems, lower-extremity systems, and full-body systems. The lower extremity exoskeletons (LEE) are designed to support anatomical functions of different joints of the lower extremity. In general, hip, knee and ankle joints have seven degrees of freedom (DOF) per limb. The LEEs are primarily developed for three main types of applications such as gait rehabilitation (i.e. helping patients with mobility disorders in the rehabilitation of musculoskeletal strength, motor control, and gait), human locomotion assistance (i.e. targeted at paralyzed patients who have lost motor and sensor function in their lower limbs) and human power augmentation (i.e. enhance human strength and endurance during locomotion and enable individuals to perform tasks that they cannot easily perform by themselves).

The ankle joint complex is concerned, encompassing the lower leg and foot, serves as the crucial link facilitating interaction between the lower limb and the ground. Especially the essential aspect for activities like walking and daily tasks. Despite experiencing significant compressive and shear forces during walking, the ankle's robust structure provides a high level of stability. The key ankle joint complex movements include sagittal plane, plantarflexion-dorsiflexion, transverse plane, abduction-adduction, and frontal plane, inversion-eversion [1].

While commonly referred to as the 'ankle joint,' the foot's movement is facilitated by several articulations. The ankle joint complex comprises the subtalar joint (talocalcaneal), tibiotalar joint (talocrural), and transverse-tarsal joint (talocalcaneonavicular) [1]. The subtalar joint's geometry permits ankle inversion and eversion, with a primary role in these movements. Although other motions are possible, the bulk of foot inversion and eversion occurs here. The talocrural joint acts as a hinge, mainly contributing to plantar- and

dorsiflexion. The transverse tarsal joint is considered part of the functional unit with the subtalar joint, sharing a common axis, and contributing to foot inversion-eversion. The combining motions at the subtalar and talocrural joints results in three-dimensional supination and pronation [2], [3].

While some suggest the talocrural joint is multi-axial due to internal rotation during dorsiflexion and external rotation during plantarflexion, evidence indicates it may be uniaxial. The observed simultaneous motions are attributed to its oblique axis. Similar to the talocrural joint, the subtalar joint has an oblique axis, contributing to multiple motions during plantar- and dorsiflexion, leading to pronation and supination [4].

In the past, there was a traditional belief that dorsiflexion and plantarflexion were exclusively associated with the motion of the talocrural joint, while inversion and eversion were thought to occur solely at the subtalar joint. However, recent perspectives have rejected the idea of completely segregating these motions to each joint. Although the majority of plantar and dorsiflexion is still attributed to the talocrural joint, there is now acknowledgment that a few degrees of these motions also involve the subtalar joint [5].

Despite some of the research efforts aimed at creating active ankle orthoses to aid human movements, challenges persist in effectively supporting natural human motions. These challenges revolve around the goal of enabling limbs to move freely and comfortably, particularly in relation to the corresponding oblique axes.

2. Design Considerations

Designing a device that will effectively interact with the human dynamic conditions is a difficult task to accomplish, especially at the distal end of limbs. Gait analysis can be used as an objective tool for quantifying motion of lower limb joints and forces that act upon these joints. However, gait analysis cannot separate it joint wise due to the major limitation of accurately measuring talus motion using skin-mounted markers. Fig. 1 depicts example gait analysis data of the ankle joint complex kinematics [6]. During a normal gait cycle, the stance phase begins at heel strike, where the dorsiflexors are eccentrically contracting to lower the foot to the ground based on the sagittal motion of the ankle. Then the ankle moves from plantarflexion to dorsiflexion during which the shank rotates forward around the ankle and after that the foot rotates around the forefoot phase continuing until maximum plantarflexion (approximately 14°) being achieved at toe-off. In the swing phase of walking, the ankle undergoes dorsiflexion, allowing the foot to lift off the ground and then transitioning to a slight plantarflexion upon heel strike. Simultaneously, there is motion at the subtalar joint, involving around 15° of

inversion and eversion to complement this flexion movement [7].

As illustrated in Fig. 2, the anatomical axes of talocrural and subtalar joints are placed in an oblique sense that cause foot to move across all three planes allowing pronation and supination to occur during walking [6], [8]. Hence, constraining motions to a single plane can lead to abnormal joint movements, poor muscle recruitment and increasing overall energetic cost. Therefore, it is essential to facilitate multiple DOF to improve exoskeleton performance by active or passive means.

Fig. 1. The outputs from gait analysis representing ankle complex rotation in sagittal and frontal planes [6].

Fig. 2. The axis of rotation at the talocrural joint and subtalar joint - oblique axes of rotation (red) are shown in different views; a) talocrural joint; b) subtalar joint.

Fig. 3. CAD model of the proposed ankle joint orthosis; a) conceptual design; b) CAD model.

Fig. 4. The axis of rotation and osteo-kinematics at the plantar-/dorsiflexion assembly a) component axis (red) [8]; b) developed plantar-/dorsiflexion assembly.

Fig. 5. The axis of rotation and osteo-kinematics at the inv-/ eversion assembly a) component axis(red) [8]; b) developed inv-/ eversion assembly.

3. Mechanical Design

This flexible active ankle joint orthosis is a multipurpose robotic ankle exoskeleton that can perform rehabilitation exercises as well as locomotion assistance. It is worn as a bilateral system and supports all three DOF of the ankle.

The two significant DOF, namely, plantarflexion-dorsiflexion and inversion-eversion are externally powered using electrical actuators and remaining DOF, namely, internal- external rotation is passively supported. In addition, with the intention of maximizing user compatibility and performance, the mechanisms proposed in the mechanical design have compatible joint axes for each DOF. The CAD model of the flexible active ankle joint orthosis is shown in Fig. 3.

3.1. Plantarflexion/ dorsiflexion assembly

The movement of the ankle predominantly takes place in the sagittal plane, where plantarflexion and dorsiflexion are mainly observed at the talocrural joint. Various research studies have suggested a range of motion (ROM) in the sagittal plane, typically falling between 65° and 75°. This encompasses a span from 10° to 20° of dorsiflexion to 40°–55° of plantarflexion [9]. Here the plantarflexion-dorsiflexion axis of the suggested ankle joint brace is accurately aligned with the talocrural axis of the natural joint. Fig. 4 illustrates the plantar-/dorsiflexion assembly with the component axis.

3.2. Inversion/ eversion assembly

In the frontal plane, the complete range of motion is around 35°, consisting of 23° of inversion and 12° of eversion [10]. The axis of inversion-eversion in the flexible active ankle joint brace is accurately aligned with the subtalar axis of the natural joint. The assembly indicated in Fig. 5 has been mechanically formed and positioned normal to the subtalar axis of ankle.

3.3. Internal/ external rotation assembly

The internal/external rotation mechanism is crafted and adjusted to closely align with the ankle's axis of rotation. A stabilizer system is employed to maintain the assembly in its neutral position.

4. Prototype Development

The prototype incorporates tri-planar motion with oblique axes. The spacing between the two motion axes (sagittal and frontal plane motion) is adaptable to accommodate the wearer, making it one of the most human-like structures among the initial design concepts. Following multiple manufacturing processes at the production facility, the final prototype was successfully assembled. Fig. 6 illustrates the conclusive form of the active ankle orthosis device, while Table 1 provides detailed specifications outlining the design features of the developed prototype.

Fig. 6. Prototype of the proposed mechanical structure

Table 1. Design specifications of the flexible active ankle joint orthosis

Characteristic	Plantar-Dorsiflexion	Inversion-Eversion
ROM (Degree)	55/20	35/25
Actuation method	Motor driven	Motor driven
Transmission	Bevel and spur gear system	Spur gear system
Max. torque output (Nm)	22.3	2.3

5. Conclusion

The distinctive feature of the design for the flexible active ankle joint brace lies in its capability to meet the functional and ergonomic demands of the natural ankle joint. This is accomplished through precise alignment of the orthosis axes of rotation with the oblique axes of rotation found in the talocrural and subtalar joints of the ankle. The innovative mechanism, inclusive of the drive units, is situated anteriorly to the shank and foot segments. The proposed active ankle brace is easily adaptable for performing ankle rehabilitation exercises, serving as a robotic therapy module to aid paralyzed individuals in regaining or enhancing their functional mobility, or assisting those with physical weakness in their motion. Future work will involve creating an EMG-based controller for the flexible active ankle joint orthosis. This aims to enable real-time control by accurately anticipating the patient's intended motion through precise predictive capabilities.

References

1. H. Gray. Gray's anatomy: with original illustrations by Henry Carter. Arcturus Publishing, 2009.
2. Cailliet R: Foot and Ankle Pain. Philadelphia: FA Davis Co, 1968
3. "Biomechanics of the ankle joint and clinical outcomes of total ankle replacement," *Journal of the Mechanical Behavior of Biomedical Materials*, vol. 1, no. 4, pp. 276–294, Oct. 2008, doi: https://doi.org/10.1016/j.jmbbm.2008.01.005.
4. A. Lundberg, O. Svensson, G. Nemeth, and G. Selvik, "The axis of rotation of the ankle joint," *The Journal of Bone and Joint Surgery*. British volume, vol. 71-B, no. 1, pp. 94–99, Jan. 1989, doi: https://doi.org/10.1302/0301-620x.71b1.2915016.
5. J. H. Calhoun, F. Li, B. R. Ledbetter, and S. F. Viegas, "A Comprehensive Study of Pressure Distribution in the Ankle Joint with Inversion and Eversion," *Foot & Ankle International*, vol. 15, no. 3, pp. 125–133, Mar. 1994, doi: https://doi.org/10.1177/107110079401500307.
6. C. L. Brockett and G. J. Chapman, "Biomechanics of the ankle," *Orthopaedics and Trauma*, vol. 30, no. 3, pp. 232–238, Jun. 2016, doi: https://doi.org/10.1016/j.mporth.2016.04.015.
7. M. Nordin, *Basic biomechanics of the musculoskeletal system by Margareta Nordin Victor H. Frankel*. Philadelphia: Lippincott Williams And Wililins, 2001.
8. "Ankle and Foot," *Clinical Gate*, Mar. 18, 2015. https://clinicalgate.com/ankle-and-foot/ (accessed Nov. 16, 2023).
9. S. K. Grimston, B. M. Nigg, D. A. Hanley, and J. R. Engsberg, "Differences in Ankle Joint Complex Range of Motion as a Function of Age," *Foot & Ankle*, vol. 14, no. 4, pp. 215–222, May 1993, doi: https://doi.org/10.1177/107110079301400407.
10. R. N. Stauffer, "Total Ankle Joint Replacement," *Archives of Surgery*, vol. 112, no. 9, p. 1105, Sep. 1977, doi: https://doi.org/10.1001/archsurg.1977.01370090087018.

Authors Introduction

Mr. Praveen Nuwantha Gunaratne

He received his Bachelor's degree in Engineering in 2018 from the Faculty of Engineering, University of Moratuwa, Sri Lanka. He is currently a Doctoral student in University of Miyazaki, Japan

Prof. Hiroki Tamura

He received the B.E. and M.E. degree from Miyazaki University in 1998 and 2000, respectively. From 2000 to 2001, he was an Engineer in Asahi Kasei Corporation, Japan. In 2001, he joined Toyama University, Toyama, Japan, where he was a Technical Official in the Department of Intellectual Information Systems. In 2006, he joined Miyazaki University, Miyazaki, Japan, where he was an Assistant Professor in the Department of Electrical and Electronic Engineering. Since 2015, he is currently a Professor in the Department of Environmental Robotics. His main research interests are Neural Networks and Optimization Problems. In recent years, he has had an interest in Biomedical Signal Processing using Soft Computing.

A study on the Real-Time Biomechanical Analysis of Lumbar Burden Utilizing Stereoscope Cameras

Taufik Hidayat Soesilo

Graduate School of Engineering, University of Miyazaki, 1-1 Gakuen Kibanadai-Nishi, Miyazaki-shi, Miyazaki, 889-2155, Japan

Praveen Nuwantha Gunaratne

Interdisciplinary Graduate School of Agriculture and Engineering, University of Miyazaki, 1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan

Hiroki Tamura

Faculty of Engineering, University of Miyazaki,1-1 Gakuen Kibanadai-Nishi, Miyazaki-shi, Miyazaki, 889-2155, Japan
E-mail: ti22068@student.miyazaki-u.ac.jp, ti20060@student.miyazaki-u.ac.jp, htamura@cc.miyazaki-u.ac.jp

Abstract

Lumbar region is susceptible to strain and stress due to various physical activities and occupational task. To study about the lumbar burden, researchers have developed many tools but most of existing design can only use static imaging or doesn't provide real-time update. The proposed system records and examines the change of body posture in real time by utilizing the capabilities of stereoscope cameras. By using MediaPipe algorithms, 2D key-points representing body joint can be extracted from images. Afterward, using the Direct Linear Transform (DLT) the corresponding 3D key-points can be calculated using obtained 2D key-points. With the 3D key-points the body angles can be computed and used to calculate the wight of the lumbar burden using JACK calculation. Finally, the proposed system reached to its aim to study about lumbar burden and adjust the person needs in real-time.

Keywords: Lumbar burden, real-time biomechanical analysis, computer vision

1. Introduction

The lumbar region of the spine, commonly referred to as the lower back, is susceptible to strain and stress due to various physical activities and occupational tasks. Around 577.0 million people worldwide, or approx. 7.5% of the total population, were projected to have low back pain at any given time in 2017, and that number is continuing to rise [1]. This incident occurs due to doing the same activities that affect the lumbar region over a long period of time. Examples of activities that cause this include agricultural work, nursing, construction, etc. Therefore, this needs to be monitored and corrected before it gets exaggerated. One of the consequences that can occur due to ignoring pain in the lumbar region is disability [2], [3].

Therefore, many researchers are trying to monitor and improve body movements with various approaches. Among these methods, the approach that is often to be seen is using smartphones and Kinect sensors [4], [5], [6], [7], [8]. The approaches have already taken their respective advantages. However, when it is applied on real scenarios, such approaches demand several additional features. For example, smartphone-based approach, it employs gyro and accelerator sensors embedded on the smartphone to measure changes in body inclination [4]. This change in body inclination will then be used to estimate the lumbar burden that occurs on the body and classify the movements the body makes [5]. On the other hand, the Kinect sensor [6-8] uses a camera to capture human images and a depth sensor to measure the depth of key-points captured by the camera. This data will then be used to obtain 3-dimensional body key points. Afterward, the 3D key-points are being used to estimate the body posture, angles, and the biomechanical of the human body. Each approach has its own advantages and disadvantages. for example, smartphones can be used in all conditions. However, based on the measurement criteria, several devices need to be attached to the body and cannot be monitored in real-time. As for Kinect sensor, the sensor can find out various information on parts of the body at once, but Kinect sensor cannot be used under sunlight.

The proposed system is designed for real-time biomechanical analysis of the body, considering the use in different environmental conditions. This research employs a camera stereoscope to analyze the body and carry out comprehensive calculations in order to identify the lumbar burden in real time.

Taufik Hidayat Soesilo, Praveen Nuwantha Gunaratne, Hiroki Tamura

allows the system to effectively detect movements within a 10-meter range while minimizing noise interference.

2.2. *Data Acquisition*

Data collection is carried out according to the flowchart depicts Fig.2. In order to carry out biomechanical analysis, the system needs to obtain 3D key-points so that monitoring of the desired body parts can be obtained [9]. Initially, the camera captures an image of the subject positioned in front of it. Then the obtained image undergoes processing via the MediaPipe algorithm to extract 2D key-points. Given that the images are derived from two distinct perspectives through two cameras, a pair of key-point perspectives right and left cameras, illustrated in Fig.3. To achieve 3D key-points, some calculations are conducted using DLT method. However, calibration is needed to derive the translation and rotation parameters of the camera in order to carry out DLT calculation. The obtained 3D key-points serve as foundational data for extracting biomechanical insights from the body. Afterwards, in order to compute lumbar burden, it's necessary to deepen the understanding of body posture, one of which is through the body angle. This angle is obtained through vector calculations utilizing the known 3D key-points coordinates. This process involves presenting the 3D key-points within a three-dimensional environment, as visualized in Fig.4.

Fig. 1. Camera setup

Fig. 2. System Flowchart

Fig. 3. 2D key-points from left and right cameras

2. Methodology

2.1. System setup

The study employed a stereoscope system comprising two cameras to investigate the activities. The cameras were strategically placed on a 1-meter-high camera support with a 1-meter separation, as depicted in Fig. 1. Distance between the two cameras was carefully addressed, as overly close placement could constrain the view, leading to more noise and less precision. Use of stereoscopic approach creates a three-dimensional image, and it leads to a comprehensive understanding of complex movement patterns involved in walking and the factors that influence them. Further, this configuration

Fig. 4. 3D key-points

2.3. *Lumbar burden calculation*

The estimation of lumbar burden involves the computation of body posture angles, as illustrated in Fig.5. Calculations are performed according to formulas Eq. 1, 2, 3, 4, 5, 6.

$$A = (x_1, y_1, z_1) \tag{1}$$

$$B = (x_2, y_2, z_2) \tag{2}$$

$$C = (x_3, y_3, z_3) \tag{3}$$

$$a = (x_a, y_a, z_a) = ((x_1 - x_2), (y_1 - y_2), (z_1 - z_2)) \tag{4}$$

$$b = (x_b, y_b, z_b) = ((x_3 - x_2), (y_3 - y_2), (z_3 - z_2)) \tag{5}$$

$$\theta = \cos^{-1}\left(\frac{(x_a x_b + y_a y_b + z_a z_b)}{\left(\sqrt{x_a^2 + y_a^2 + z_a^2} \cdot \sqrt{x_b^2 + y_b^2 + z_b^2}\right)}\right) \tag{6}$$

This angle calculation can then be used to carry out biomechanical analysis by applying JACK calculations. The JACK computation demands data on body posture (such as body angle), subject-specific information such as height and weight, and the weight of the object intended for lifting. Fig.6 represents the coordinates involved, and Table 1 provides a detailed description of the corresponding angles denoted by its abbreviations.

In this study, the height of the participated subject was 160 cm with a body weight of 50 kg, with different load settings of objects, ranging from 0 kg (absence of a load) to 2.5 kg and 5 kg. The proposed system records the computed results in an Excel file, facilitating comprehensive post-analysis.

3. Results and Discussion

In order to know the effect of load on lumbar burden, a comparison is needed by providing different loads. In recent research, the load given was divided into 3 loads, namely 0 kg or no load, 2.5 kg load, and 5 kg load. The movement carried out by the subject is the same movement as shown in Fig 7. but the system simulates the target of lifting a different load. The results of the lumbar burden analysis can be seen in the Fig. 8.

Based on the results, it can be observed that the system can analyze the lumbar burden according to the movements made by the subject. Moreover, the load that occurs in the lumbar region is directly proportional with the load that being lifted. The maximum load that occurs in the lumbar region when the subject was at the lowest point and tries to lift the weight; 2098 N for 0 kg, 2843 N for 2.5 kg, and 3612 N for 5 kg (Table 2). Further, minimum, and average values can also be found in Table 2.

Fig. 5 Vector coordinates

Fig. 6 Respective vector coordinates

Table 1. The coordinate details.

TA	Trunk anteversion angle
RUA	Right upper arm angle
LUA	Left upper arm angle
RF	Right forearm angle
LF	Left forearm angle
RT	Right thigh angle
LT	Left thigh angle
RLL	Right lower leg angle
LLL	Left lower leg angle

Fig. 7 JACK angle for Lumbar Burden calculation

Fig.8. Lumbar burden analysis

Table 2. Lumbar burden Max, Min, and Average

Load (kg)	Lumbar Burden (N)		
	Max	Min	Average
0	2098	790	1655
2.5	2843	1299	2327
5	3612	1770	2974

4. Conclusion

Based on the results it can be concluded that the proposed system can carry out biomechanical analysis to determine the lumbar burden effectively. The procedure was executed in real-time, also enabling monitoring of the system within any occupational settings. Further, the system demonstrated a robust performance under sunlight, showing its adaptability to diverse environmental conditions.

For future work, the observed results will undergo a validation process using an established system to assess the efficacy of the developed system. Furthermore, as a system constraint, the current study is limited to the detection of a single person. Therefore, in order to conduct real-time analysis of lumbar burden for multiple persons, the system requires further development.

References

1. Wu A, March L, Zheng X, Huang J, Wang X, Zhao J, Blyth FM, Smith E, Buchbinder R, Hoy D. Global low back pain prevalence and years lived with disability from 1990 to 2017: estimates from the Global Burden of Disease Study 2017. Ann Trans Med 2020; 8(6): 299-313.
2. Global Health Group Data Exchange http://ghdx.healthdata.org/gbd-results-tool accessed Nov 20, 2023.
3. Hartvingsen J, Hancock MJ, Kongsted A, Louw Q, Ferreira ML, Genevay S, Hoy D, Karppinen J, Glenn Pransky, Sieper J, Smeets RJ, Underwood M. What low back pain is and why we need to pay attention. Lancet 2015; 386: 2145-2191.
4. N.D. Nath, R. Akhavian, A.H. Behzadan, Ergonomic analysis of construction worker's body postures using wearable mobile sensors, Appl. Ergon. 62 (2017), 107–117.
5. H. Tamura, K. Sakurai, K. Tanno, Y. Fuse. A Study on the Lumbar Burden Evaluation of Work using One Smartphone. Journal of Robotics, Networking and Artificial Life, Vol. 5(3); December (2018), pp. 173–179.
6. Liu, P.-L.; Chang, C.-C.; Li, L.; Xu, X. A Simple Method to Optimally Select Upper-Limb Joint Angle Trajectories from Two Kinect Sensors during the Twisting Task for Posture Analysis. Sensors 2022, 22, 7662. https://doi.org/10.3390/s22197662.
7. Singh, L.P., Kumar, P. & Lohan, S.K. (2023) Development of a real‐time work‐related postural risk assessment system of farm workers using a sensor‐based artificial

intelligence approach. Journal of Field Robotics, 1‑14. https://doi.org/10.1002/rob.22215.

8. Christopher B., Oliver B., Pietro M., Tobias H., Laura J., Alexander M., Marco T., Verena N., Using real-time feedback of L5/S1 compression force based on markerless optical motion capture to improve the lifting technique in manual materials handling, https://doi.org/10.1016/j.ergon.2022.103350.

9. Temuge B. Real time 3D body pose estimation using MediaPipe. Github. https://github.com/TemugeB/bodypose3d accessed Nov 14, 2022.

Authors Introduction

Mr. Taufik Hidayat Soesilo

He received his Bachelor's degree in Engineering in 2021 from the Faculty of Engineering, University of Brawijaya in Indonesia. He is currently a master student in University of Miyazaki, Japan

Mr. Praveen Nuwantha Gunaratne

He received his bachelor's degree in engineering in 2018 from the Faculty of Engineering, University of Moratuwa, Sri Lanka. He is currently a doctoral student in University of Miyazaki, Japan.

Prof. Hiroki Tamura

He received the B.E. and M.E. degree from Miyazaki University in 1998 and 2000, respectively. From 2000 to 2001, he was an Engineer in Asahi Kasei Corporation, Japan. In 2001, he joined Toyama University, Toyama, Japan, where he was a Technical Official in the Department of Intellectual Information Systems. In 2006, he joined Miyazaki University, Miyazaki, Japan, where he was an Assistant Professor in the Department of Electrical and Electronic Engineering. Since 2015, he is currently a Professor in the Department of Environmental Robotics. His main research interests are Neural Networks and Optimization Problems. In recent years, he has had interest in Biomedical Signal Processing using Soft Computing.

Verification of Determination Possibility using Convolutional Autoencoder for Machine Tool Abnormality Detection

Yuta Sumoto
Faculty of Engineering, University of Miyazaki, 1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan

Praveen Nuwantha Gunaratne
Interdisciplinary Graduate School of Agriculture and Engineering, University of Miyazaki, 1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan

Hiroki Tamura
Faculty of Engineering, University of Miyazaki, 1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan
Email: hi20024@student.miyazaki-u.ac.jp, ti20060@student.miyazaki-u.ac.jp, htamura@cc.miyazaki-u.ac.jp

Abstract

The purpose of this research is to clarify the cause of failure and to improve the accuracy of abnormality detection by predicting the noise added to the machine tool from the results of the convolutional autoencoder. Data obtained from an acceleration sensor mounted on the machine tool are reconstructed using a convolutional autoencoder, and the average absolute error is calculated. The maximum value of the average absolute error of the training data is used as the threshold value for abnormality detection. Multiple simulated data with different amplitude values based on a composite sine wave with a specific frequency and random numbers within a specified amplitude value were verified. In this paper, the characteristics of each type of noise and the parameters of the optimal model were examined from error rate and error distribution.

Keywords: machine tools, deep learning, anomaly detection, convolutional autoencoder, noises

1. Introduction

In factories, machine failures can be categorized as sudden breakdowns and deteriorating faults [1]. Preventing failures proactively during maintenance is exceptionally challenging, and excessive maintenance or machine breakdowns can result in significant losses for companies. Examples include production downtime leading to manufacturing time losses, increased labor costs due to stoppages, and the generation of defective products. The purpose of this study is to reduce losses and enhance productivity by providing maintenance timing suggestions and early detection of abnormalities in machining equipment.

In collaboration with businesses, this research involved installing acceleration sensors on machining equipment used in semiconductor manufacturing. The data acquired from the sensors is in the form of continuous time-series data. Since explicit abnormal judgments are significantly fewer than normal data, it is necessary to perform anomaly detection without explicit labels for normal and abnormal data. Consequently, anomaly detection was conducted using a convolutional autoencoder [2].

Initially, the machine with the attached sensor experienced a breakdown, and there were no prospects for its reprocessing. Subsequently, the sensor was transferred to a similar type of machining equipment. After the transfer, only normal data was obtained from the factory's machines. As abnormal data obtained before the transfer exhibited specific frequency components, simulated data resembling those characteristics was generated. Parameters for multiple models were created based on different numbers of training data, and abnormal detection was verified by introducing specific noise to the training data [3] to assess how much noise is required to detect anomalies. Moreover, considering the possibility of identifying malfunctioning machine components by pinpointing the noise, we investigated the distribution of reconstruction errors.

2. Proposed method

The proposed method consists of two steps: Learning and parameter saving, and Testing validation.

2.1. Learning and parameter saving

Initially, a convolutional autoencoder was created by referencing the Keras tutorial [4] for parameter

Yuta Sumoto, Praveen Nuwantha Gunaratne, Hiroki Tamura

Fig. 1. Flowchart of learning and parameter saving

inheritance (Fig. 1). The training data was standardized using the mean and standard deviation of the data. A sequence with a timestep of 100 was generated using a sequential model, and training was conducted. A callback was implemented to halt training if the validation loss did not improve for five consecutive epochs. The hyperparameters were set as follows: epochs = 5, batch size = 4096, and validation split = 0.1.

After training, the parameters were saved, and the model was used to reconstruct the training data. The absolute mean error between the actual data and the reconstructed data for each sequence was calculated, and the maximum value was set as the threshold for testing.

2.2. Testing validation

The process involves extracting one CSV file at a time from the dataset. The training data is then standardized using the mean and standard deviation of the respective file. Test data is reconstructed using the specified model parameters. For each sequence, the absolute mean error between the test data and the reconstructed data is calculated. Instances where this error exceeds the predefined threshold are considered anomalies. (Fig. 2)

3. Experiment

The sampling frequency of the experimental data was 500Hz, and each file contained 600,000 data points (equivalent to 20 minutes) with columns for time (s), x

(m/s²), y (m/s²), and z (m/s²). For this study, only the time (s) and x (m/s²) columns were utilized. To assess the accuracy of each parameter, twelve CSV files were utilized for training. The process depicted in Figure 1 was repeated twelve times, and the resulting parameters were saved. The model parameters were denoted as model_1, model_2, ..., model_n corresponding to the number of training data (n). With an epoch count of 5 per training, a total of 60 training iterations were conducted as the process was repeated 12 times. (Fig. 3).

Analyzing the frequency components of normal and abnormal data obtained from the machine before the transfer revealed that normal data covers a wide frequency range, whereas abnormal data is concentrated in the high-frequency band. The results of the FFT (Fast Fourier Transform) applied to normal and abnormal data are shown in Fig. 4.

Based on these results, a composite wave of sine waves with specific frequencies was added to the initial training data to create simulated abnormal data. The added composite waves consisted of low frequencies at 40Hz and 90Hz, and high frequencies at 410Hz and 460Hz.

Fig. 2. Flowchart of testing validation

Fig. 3. Loss function

Fig. 4. Frequency analysis results of normal and abnormal data

Additionally, simulated data was generated by adding random numbers within specified maximum and minimum values. The specified absolute values for amplitude and random numbers ranged from 0.1 to a maximum of 0.8.

3.1. Testing validation

Test the three types of simulated data using twelve parameters (model_n) with varying amounts of training data. Based on the results, examine the optimal amount of training for anomaly detection.

3.2. Testing validation

Detect the three types of simulated data using the best parameters. Based on the results, evaluate how well anomalies are detected when various levels of noise are added and investigate the threshold for anomaly detection.

3.3. Testing validation

When the three types exhibit the same error rate, illustrate the distribution of the mean absolute errors between the actual data and the reconstructed data for each type. Compare the mean absolute errors when the

error is approximately 10% and investigate whether there are any distinctive features in the distribution due to the added noise.

4. Results
4.1. Parameter Comparison

The order of accuracy in detecting anomalies for abnormal data is as follows:
- Low-frequency components: model_3, model_5, model_11, model_2 (Fig. 5)
- Random noise: model_3, model_11, model_12, model_5 (Fig. 6)
- High-frequency components: model_11, model_12, model_10, model_3 (Fig. 7)

For data with added low-frequency components and random noise, the parameters of Model_3, trained using three datasets, exhibited the most effective anomaly detection. Regarding data with added high-frequency component noise, Model_11, trained using eleven datasets, demonstrated the most proficient anomaly detection.

Fig. 5. Parameter comparison (Low-Frequency component)

Fig. 6. Parameter Comparison (Random Data)

Yuta Sumoto, Praveen Nuwantha Gunaratne, Hiroki Tamura

Fig. 7. Parameter Comparison (High-Frequency Component)

Fig. 8. Comparison of Model_3

4.2. Simulated Data Comparison

Based on the results from 4.1, a comparison of the three types of simulated data was performed using model_3 (Fig. 8). For all three types, errors reached 10% when the additional amplitude and specified absolute values ranged from 0.3 to 0.4. Additionally, it was observed that errors sharply increased for values equal to or greater than 0.4.

4.3. Noise Identification

When the error is approximately 10%, the distribution of actual data and reconstructed data for the three types is shown. Additionally, when the error is approximately 10%, the amplitude of the synthesized wave for low-frequency is 0.325, for high-frequency is 0.35, and the specified absolute value for random data is 0.3 (Fig. 9, Fig. 10 and Fig. 11). Differences in items such as mean, median, standard deviation, minimum, 25th percentile, 75th percentile, and maximum, based on the obtained basic statistical information, are on the order of one-thousandth. It can be observed from the distribution that when the anomaly rate is the same, there is little difference.

5. Conclusion

In the experimental data, step count, and hyperparameters employed in this study, the parameters of model_3 and model_11 demonstrated high accuracy in anomaly detection. From this, it was found that under the current conditions, abnormalities could be detected with high accuracy if three or 11 or more pieces of normal data, each with a total of 600,000 pieces of data, were trained.

From experiments using parameters trained three times, it was observed that, irrespective of whether the noise is periodic, high-frequency, or low-frequency, anomalies were detected when noise with an amplitude value or specified absolute value exceeding 0.3 was introduced. While it was not possible to identify the characteristics of the noise from the distribution of reconstruction errors, it is possible to estimate the amplitude of the added noise from the error percentages.

Currently, data is being measured in the factory and saved in specific files. Processing is set to analyze the second-to-last CSV file every 20 minutes. If anomalies are detected above 10%, the number 1 is added to the array; if below 10%, the number 0 is added. An alert is triggered when anomalies are detected continuously three times, and the array is output daily.

Fig. 9. Low Frequency- Additional Amplitude 0.325

Fig. 10. High Frequency- Additional Amplitude 0.35

Fig. 11. Random- Specified Absolute Value 0.3

References

1. J. Guo, Y. Miki, Y. Fujita, K. Kiritoshi, and K. Ito, "Anomaly Detection of Seismogram with Autoencoder." Accessed: Dec. 10, 2023. [Online]. Available: https://www.jstage.jst.go.jp/article/pjsai/JSAI2020/0/JSAI2020_4L2GS1303/_pdf/-char/ja

2. S. Yan, H. Shao, Y. Xiao, B. Liu, and J. Wan, "Hybrid robust convolutional autoencoder for unsupervised anomaly detection of machine tools under noises," *Robotics and Computer-Integrated Manufacturing*, vol. 79, p. 102441, Feb. 2023, doi: https://doi.org/10.1016/j.rcim.2022.102441.

3. T. Mine et al., "Anomaly detection of mechanical equipment with autoencoder," 2016. Accessed: Dec. 10, 2023. [Online]. Available: https://www.jstage.jst.go.jp/article/fss/35/0/35_506/_pdf/-char/ja

4. K. Team, "Keras documentation: Timeseries anomaly detection using an Autoencoder," keras.io. https://keras.io/examples/timeseries/timeseries_anomaly_detection/

Authors Introduction

Mr. Yuta Sumoto

Yuta Sumoto born in 2001. He is currently studying in department of environmental robotics, and will receive the B.Eng from University of Miyazaki in 2022. His current research is Abnormality detection using auto encoder.

Mr. Praveen Nuwantha Gunaratne

He received his Bachelor's degree in Engineering in 2018 from the Faculty of Engineering, University of Moratuwa, Sri Lanka. He is currently a Doctoral student in University of Miyazaki, Japan

Prof. Hiroki Tamura

He received the B.E. and M.E. degree from Miyazaki University in 1998 and 2000, respectively. From 2000 to 2001, he was an Engineer in Asahi Kasei Corporation, Japan. In 2001, he joined Toyama University, Toyama, Japan, where he was a Technical Official in the Department of Intellectual Information Systems. In 2006, he joined Miyazaki University, Miyazaki, Japan, where he was an Assistant Professor in the Department of Electrical and Electronic Engineering. Since 2015, he is currently a Professor in the Department of Environmental Robotics. His main research interests are Neural Networks and Optimization Problems. In recent years, he has had interest in Biomedical Signal Processing using Soft Computing.

Basic Research for High-speed Heart Sound Determination using AI

Riku Nakashima
Faculty of Engineering, University of Miyazaki,
1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan

Praveen Nuwantha Gunaratne
Interdisciplinary Graduate School of Agriculture and Engineering, University of Miyazaki, 1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan

Hiroki Tamura
Faculty of Engineering, University of Miyazaki,
1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan

Email: hi20031@student.miyazaki-u.ac.jp, ti20060@student.miyazaki-u.ac.jp, htamura@cc.miyazaki-u.ac.jp

Abstract

In recent years, many researchers have focused on enhancing the current situation of home healthcare systems. In this paper, we developed a system that uses AI to quickly determine whether auscultation sounds are normal or not. In our analysis, heart sounds were imaged, and their abnormalities were identified using machine learning tools. The proposed approach uses a Yolo_v7 model-based anomaly detection and the Wavelet Transform was employed to analyze the acquired heart sound data. Proposed system resulted with an 90% accuracy rate or over on the basis of 5 seconds of heartbeat data, in various recording environments.

Keywords: Auscultation sounds, abnormality identification, YOLO_v7, Wavelet Transform

1. Introduction

In today's society, the demand for home medical care is rapidly increasing due to the aging population. However, home medical care requires a physician to visit the patient in person to make a diagnosis, which poses problems such as securing time for physicians. A possible solution to this problem would be to make medical equipment readily accessible to ordinary households, enabling individuals to personally use and evaluate the effectiveness of the equipment. However, as proficiency in medical knowledge is crucial for operating medical equipment, the current situation hinders its widespread utilization. To identify an approach suitable for the general public, eliminating the necessity for a visit to a physician or medical expertise, our team has attempted to create a system capable of swiftly assessing the normality of heart sound data obtained through auscultation.

In this paper, machine learning tools were employed to analyze the heart sound data. Heart sound data from diverse recording environments was utilized in this paper. The acquired heart sound data was transformed into images through wavelet transforms, and the yolov7 model was employed for abnormality detection. Additionally, a *k*-fold approach was applied to a limited dataset to enhance accuracy. The findings revealed a 90% accuracy rate based on 5 seconds of heartbeat data. This paper compares Yolo model with different durations of phonocardiography data and elucidates strategies for achieving higher confidence scores.

2. Methodology

Initially the Speech data, illustrating frequency features of heart sounds, went through wavelet transformation at various intervals to generate images. Afterwards, the transformed images were utilized to train multiple Yolo models, facilitating a comparative analysis to identify the superior model. Here, the assessment was conducted based on confidence scores assigned to the normal and abnormal classes. For the Yolo training data set, the images were divided into three parts: training data, valid data, and test data. Table 1 illustrates the learning environment used in this analysis, and the Fig. 1 depicts the experimental procedure in brief.

3. Experiment

In the section 3.1, the ratio of training data to validation data and test data was set at 7:2:1 in the Yolo model. The subsequent phases, starting from section 3.2, the ratio was adjusted to 6:2:2. The offline datasets used were heart sound data auscultated by Sono-Support clinical

Riku Nakashima, Praveen Nuwantha Gunaratne, Hiroki Tamura

technologists (dataA – 10 healthy subjects and 40 sick subjects), heart sound data from former laboratory members (dataC - 23 healthy subjects), and heart sound data of few laboratory members (jibun - 2 healthy subjects). The data from physioNet (dataB~dataB_F), comprising 2558 healthy subjects, 665 sick subjects was used as online data [1].

Table 1. Learning environment of the YOLO model

Learning environment	
Memory	32GB
CPU	Intel(R) Core(TM) i7 – 10700 CPU@2.90GHz
GPU	NVIDIA GeForce RTX 3070
CUDA	11.2

Fig. 1. Experimental Procedure

Fig. 2, 3, and 4 illustrate the images generated by the wavelet transform used in this paper. The vertical axis represents frequency, the horizontal axis signifies a specific time width of the experimental data, and the color bar indicates amplitude. The wavelet transform is output as an image after being processed using the Morlet function with MATLAB. A one second shift was used to increase the number of data by shifting the image by few data. Due to the variation in the number of seconds in heartbeat, the number of images produced varies as well. The *k*-fold method involves creating *k* splits of the training data images and training them *k* times (*k*=5 in this case). The splits were randomly divided into 8 for the training data and 2 for the validation data. Those with a confidence score of less than 0.1 between normal and abnormal were excluded due to their classification falling within gray zones as lacking in reliability. The formulas for calculating the percentage of correct answers Eq. (1) and the percentage of target hits Eq. (2) are presented below.

$$Percentage\ of\ correct\ answers = \frac{Number\ of\ correct\ images}{Total\ number\ of\ images} \times 100 \quad (1)$$

$$Percentage\ of\ target\ hits = \frac{true\ positive}{true\ positive + false\ negative} \times 100 \quad (2)$$

3.1. Experiment 1

In experiment 1, multiple Yolo models (Yolo_v2, Yolo_v5, and Yolo_v7) were employed to jointly train normal and abnormal heart sound images and assess the superiority of each model.[2], [3].

3.2. Experiment 2

Multiple Yolo models with favorable confidence scores were employed, and machine learning was conducted on each of them for varying durations (3, 5, and 10 seconds), comparing their performance in terms of the percentage of correct responses. However, a 10-second wavelet transform could not be generated due to insufficient seconds in imaging dataB_B. Fig. 3, 4 was generated by performing a wavelet transform on a single speech sample, cut out for a fixed time width of 1 second intervals and converting it to an image.

Fig. 2. Wavelet transform 3 sec.

Fig. 3. Wavelet transform 5 sec.

3.3. *Experiment 3*

In Experiment 3, the accuracy of the 5-second wavelet transforms, identified as the best on real data in Section 3.2. was enhanced by employing the *k*-fold method. This method involved covering a small data set and increasing the number of training patterns.

Fig. 4. Wavelet transform 10 sec.

4. Results
4.1. Results of Experiment 1

The correctness of machine learning responses for each Yolo was determined based on the agreement between the label and the test judgment. A comparison of the mean confidence scores of the correct answers revealed the superiority of Yolo_v7, as illustrated in Table 2.

Table 2. Confidence score results of each YOLO

	Yolov2	Yolov5	Yolov7
Confidence Score	0.957	0.973	0.997

4.2. Results of Experiment 2

Machine learning with Yolo_v7 was carried forward, as it came out as the best among the other Yolo models based on the results of Section 4.1. Consequently, the correct response rates were as follows: 3 seconds, 90.6% for normal and 89.2% for abnormal (as shown in Table 3); 5 seconds, 86.8% for normal and 90.6% for abnormal (Table 4); and in 10 seconds, 68.7% for normal and 94.4% for abnormal (Table 5). These three results demonstrate that the 3-second cutout yields the highest average correct response rate. However, upon comparing with actual data (dataA), the 5-second wavelet transform exhibits the best results. While deciding which result to adopt presents a doubt, move forwarded with k-fold using the 5-second wavelet transform, as it yields the best results on real data.

4.3. Results of Experiment 3

The outcomes obtained through k-fold training on a 5-second wavelet transform revealed accuracy rates of 90.3% on normal and 93.6% on abnormal, which are over 90%.

5. Conclusion

Yolo_v7 enhanced the accuracy to 90% in diverse recording environments, achieved by optimizing a

limited dataset and enhancing learning patterns through k-fold techniques. Future research will focus on strategies to enhance accuracy by expanding the dataset. Further, considering the satisfactory average accuracy of the 3-second wavelet transform, experiment with k-fold on the 3-second version will be continued.

Table. 3. Wavelet transform 3-second correct response rate

3s	data A	data B	data B_B	data B_C	data B_D	data B_E	data B_F	data C	jibun	Aver age
Nor mal	84.1	67.6	80.6	89.8	100	100	100	100	100	92.4
Abno rmal	80	89.4	95.2	100	100	99.3	100			94.8

Table. 4. Wavelet transform 5-second correct response rate

5s	data A	data B	data B_B	data B_C	data B_D	data B_E	data B_F	data C	jibun	aver age
Nor mal	84.2	93.5	91.9	100	35.9	100	75.4	100	100	86.8
abno rmal	85.6	80	100	100	95	99.8	73.9			90.6

Table. 5. Wavelet transform 10-second correct response rate

10s	data A	data B	data B_C	data B_D	data B_E	data B_F	data C	jibun	Aver age
Nor mal	81.8	22	19	33	94	100	100	100	68.7
abno rmal	76	100	98.5	100	100	91.9			94.4

References

1. " Classification of Heart Sound Recordings: The PhysioNet/Computing in Cardiology Challenge 2016 v1.0.0," physionet.org. https://physionet.org/content/challenge-2016/1.0.0/
2. "Yolov7"(Japanese) Qiita, Oct. 10, 2022. https://qiita.com/johnrocky/items/a4c24e55acaf7dcccdd9 (accessed Dec. 01, 2023).
3. "YOLOv5"(Japanese) Qiita, Oct. 27, 2021. https://qiita.com/gotty21/items/37b52cec9542fd87bd5f (accessed Dec. 01, 2023)

Authors Introduction

Mr. Riku Nakashima

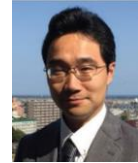He was born in 2001. He is currently studying indepartment of environmental robotics, and will receive the B.Eng from University of Miyazaki in 2024. His current research is Basic Research for High-speed Heart Sound Determination using AI.

Mr. Praveen Nuwantha Gunaratne

He received his Bachelor's degree in Engineering in 2018 from the Faculty of Engineering, University of Moratuwa, Sri Lanka. He is currently a Doctoral student in University of Miyazaki, Japan

Prof. Hiroki Tamura

He received the B.E. and M.E. degree from Miyazaki University in 1998 and 2000, respectively. From 2000 to 2001, he was an Engineer in Asahi Kasei Corporation, Japan. In 2001, he joined Toyama University, Toyama, Japan, where he was a Technical Official in the Department of Intellectual Information Systems. In 2006, he joined Miyazaki University, Miyazaki, Japan, where he was an Assistant Professor in the Department of Electrical and Electronic Engineering. Since 2015, he is currently a Professor in the Department of Environmental Robotics. His main research interests are Neural Networks and Optimization Problems. In recent years, he has had interest in Biomedical Signal Processing using Soft Computing.

Prediction of High-Energy Electron Flux at Geosynchronous Orbit using a Neural Network Technique

Ami Iwabu

Kyushu Institute of Technology, 1-1 Sensui-cho, Tobata-ku, Kitakyushu-shi,Fukuoka, 804-8550, Japan

Kentaro Kitamura

Kyushu Institute of Technology, 1-1 Sensui-cho, Tobata-ku, Kitakyushu-shi,Fukuoka, 804-8550, Japan
E-mail: iwabu.ami593@mail.kyutech.jp, kitamura.kentaro375@mail.kyutech.jp
https://www.kyutech.ac.jp

Abstract

The radiation belt, where the high-energy particles are predominant in near earth space from the low earth orbit (LEO) to the geostationary orbit (GEO), sometimes causes satellite malfunction. Therefore, the objective of this study is to predict the high-energy electron flux at GEO with the energy above 2 MeV after 24 hours with higher accuracy for the safety satellite operation in terms of the space weather science. In this study the various kinds of solar wind data from satellite observations and ground geomagnetic observation data in 1999 were used for the Recurrent Neural Network (RNN). Prediction results were evaluated by the prediction efficiency, which is derived from both predicted and actual variation data. As a result, the prediction using combined data of solar wind and geomagnetic data shows highest prediction efficiency of 0.72.

Keywords: Neural network, Spacecraft, High-energy electron flux

1. Introduction

The region in space where high-energy particles are concentrated is called the radiation belt, and the electrons that exist there are called radiation belt electrons. The radiation belts consist of the inner and outer belts. In particular, the outer belt sometimes extends into geostationary orbit, where many satellites are operated [1], [2], [3]. High-energy electrons with energies above MeV are known to cause satellite malfunctions and failures [4], [5].

So far, several studies have been conducted to predict fluctuations in high-energy electron flux by using only ground-based observation data as input parameters or combining ground-based observation data with solar wind data observed in space [6]. Fukada *et al.* [7] used only ground-based AE index and D_{st} index data during magnetic storms as input data and made predictions using a neural network. In their study, predictions were made from 2 to 12 hours later, and the prediction efficiency was as high as 0.71. However, since only the case of magnetic storms was used as training data, it was not practical because it could not predict the starting point where the high-energy electron flux fluctuates significantly. Nakamura *et al.* [8] used seven solar wind data as input data: V_{sw}: solar wind velocity, Bx: solar direction component of solar wind magnetic field data, B_z: north-south component of solar wind magnetic field data, E: electron flux with energy above 2 [MeV], and AE index, D_{st} index, and UT: universal standard time as ground observation data, The same neural network was used to make predictions. In their study, data from 1999 to 2006 were used to predict the energetic electron flux 24 hours later by combining space-based and ground-based data, and the prediction efficiency was 0.61. There is room to further improve this prediction efficiency by the number of input data and the way they are combined.

In this study, with the aim of further improving the prediction accuracy, we use various combinations of ground-based observation data and solar wind data as input data to predict the energetic electron flux over the next 24 hours. The results are also evaluated using prediction efficiency (PE).

2. Dataset

In this study, solar wind velocity (V_{sw}), north-south component of the solar wind magnetic field (B_z), and high-energy (>2 MeV) electron flux (E), which are solar wind data, and AE index, AU index, AL index, D_{st} index, and Universal Time (UT), which are ground-based observation data, as input parameters. AE, AU and AL indices are used as a proxy of the Auroral activity, and

Dst index apparently represents the intensity of the Magnetic storms. We also used Akasofu ε parameter calculated by Eq. (1) [9], which efficiently represent a total energy transported from the solar wind into magnetosphere. V_{sw} and B_z were obtained from the solar wind probe ACE, which was acquired from Omni Web, and E was obtained from the geostationary satellite GOES10, which was acquired from the Space Weather Prediction Center (NOAA) of the U.S. National Oceanic and Atmospheric Administration. The AE, AU and AL indices, and D_{st} index, which are ground-based observation data, were obtained from the Kyoto University Geomagnetic World Data Analysis Center.

In this study, one year of data in 1999 was used, of which 80% was for training data and 20% was for test data. The respective parameters are shown in Fig. 2-1.

$$\text{Akasofu } \varepsilon = V_{sw} \times |B|^2 \times f(t) \times (7Re)^2 \quad (1)$$

Re：Earth radius

B：Solar wind magnetic field average

where $f(t) = \left\{ \sin\left(\frac{t}{2}\right) \right\}^4$, and the value of t is determined by Bz using the following equation.

$$B_z > 0 \quad t = \arctan \left| \frac{B_y}{B_z} \right|$$

$$B_z < 0 \quad t = \pi - \arctan \left| \frac{B_y}{B_z} \right|$$

Fig.2-1 Data for each parameter in 1999

Fig.2-1 shows that the temporal variations of solar wind parameters are basically correlated with that of the high-energy electron flux. However, in the short time-scales, the relationship is not necessarily a one-to-one correspondence, but rather a mixture of correlated and uncorrelated portions, indicating a complex relationship among various parameters.

3. Recurrent Neural Network (RNN)

Neural networks are systems that mimic the mechanisms of the human brain in order to perform processes such as recognition, memory, and judgment on a computer. There are two main types of neural networks: supervised learning and unsupervised learning, and supervised learning is used in this study. Supervised learning is a learning method in which an input signal is given and the output signal is repeatedly compared with a teacher signal, and the coupling loadings of each neuron are modified to reduce the error and adapt to a given problem [10].

In this study, a recurrent neural network (RNN) was employed to learn and predict the data. Here, RNN is a neural network that recursively repeats learning in each unit and is suitable for use on continuous data, as shown in Fig.3-1.

Fig.3-1 Structure of RNN

4. Results

In this study, predictions were made for various combinations of parameters as shown in Table 4-1, and the results were evaluated using the prediction efficiency (PE) shown in Eq. (2),

$$PE = 1 - \frac{MSE}{VAR} \quad (2)$$
$$MSE = \frac{1}{N} \sum_{i=1}^{N} (f_i - x_i)^2$$
$$VAR = \frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})^2$$
$$\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i$$

where x_i in each equation is the actual observed value and f_i is the predicted value.

Predictions were made five times for each epoch in each case. The results for each case are shown in Fig.4-1 to Fig.4-5. The horizontal axis is the number of epochs, the vertical axis is PE, and the average value of the five predictions in each epoch is shown as a dot.

Table 4-1 Input data for each forecast

Case 1	V_{sw}, E, B_z, D_{st} index, AE index
Case 2	V_{sw}, E, B_z, D_{st} index, AE index, UT
Case 3	V_{sw}, E, D_{st} index, AEindex, Akasofu ε
Case 4	V_{sw}, E, B_z, D_{st} index, AU index, AL index
Case 5	V_{sw}, E, B_z, D_{st} index, AU index,

Fig.4-1 PE for case 1

Fig.4-2 PE for case 2

Fig.4-3 PE for case 3

Fig.4-4 PE for case 4

Fig.4-5 PE for case 5

From Fig.4-1 to Fig.4-5, it seems that there is no correlation between the number of epochs and prediction efficiency between epochs 50 and 100.

Next, Table 4-2 shows the average prediction efficiencies obtained from the results for all epochs in each case.

Table 4-2 Comparison of average PE in each case

Case	1	2	3	4	5
Average PE	0.61	0.51	0.61	0.57	0.57

Table 4-2 indicates that Case 2, in which UT (the proxy of the satellite position) was added as input data to Case 1, shows the worst PE, while the case 1 shows the highest PE. On the other hand, Case 3, in which Akasofu ε (the total energy inflow from the solar wind) was added, had the same PE as Case 1.

Case 1 with an epoch number of 70, PE shows the highest value of 0.72. The prediction results for this case are shown in Fig.4-6 and Fig.4-7.

Fig.4-6 Comparison of predicted and observed data at the highest PE

Ami Iwabu, Kentaro Kitamura

Fig.4-7 Comparison of predicted and observed data at the highest PE

In Fig.4-6 and Fig.4-7 show the comparison of the temporal variations for both predicted and observed high-energy electron flux. The horizontal axis shows the date and the vertical axis shows the energetic electron flux. The blue line shows the actual observed data, and the orange line shows the predicted data. The more the orange line, the predicted data, overlaps the blue line, the observed data, the higher the prediction efficiency and accuracy.

Fig.4-6 shows that the forecast results are largely consistent for large changes. Fig.4-7 shows that the predicted values are lower than the actual observed values in many areas where the value of E is small.

5. Discussion

In Case 2, in which UT (the positional information of the satellite) was added as input data to Case 1, resulted PE is lower than that in Case 1. On the other hand, PE in Case 3 is almost same as that in Case 1, while Case 3 is added one parameter (the Akasofu ε which is the total number of energy inflows from the solar wind) to Case 1 same as Case 2. The difference of the added parameters causes significant difference in consequent result. This suggests that the spatial information might reduce the PE, rather than the increasing of number of input parameters. Especially, using one parameter that include the integrated several information can efficiently improve PE, rather than superposing several parameters that has individual information. On the other hand, the input parameters of solar wind data might have potential to improve the PE compared to that of ground data.

6. Conclusion

In each case, there is no correlation between the number of epochs and the PE, which means that the PE is not dependent on the number of epochs between 50 and 100, and there is no need to set the number of epochs above 100.

From the comparison of the average prediction efficiency for each case, the accuracy of prediction is higher when the parameters to be used as input data have integral information rather than when the number of input data is increased.

From the prediction results, the trend of the change in high-energy electron flux is captured. In addition, in the short time scale, we find that the prediction shows a little underestimation at the time of decreasing of E, but during E has an energy of 10^4 or more (alert level for the satellite malfunction), the prediction is highly accurate. In conclusion, the current result suggests that the RNN method proposed in this study is able to predict the periods considered important for the safe operation of the satellite.

Acknowledgements

References

1. Van Allen, J. A. G. H. Ludwing, E.C. Ray, and C. E. McIlwain (1958), Observations of high intensity radiation by satellites 1958 Alpha and Gamma, Jet Propulsion, 28, 588-592.
2. Van Allen, J. A., and L. A. Frank (1959), Radiation around the Earth to a radial distance of 107,400 km, Nature, 183, 430.
3. Baker DN, McPherron RL, Cayton TE, *et al*. (1990) Linear prediction filter analysis of relativistic electron properties at 6.6 Rea. J Geophys Res 95:15 133-15 140.
4. Lohmeyer, W. Q., and K. Cahoy (2013), Space weather radiation effects on geostationary satellite solid-state power amplifiers, Space Weather, 11, 476–488, doi:10.1002/swe.20071.
5. Baker, D. N. (2000), The occurrence of operational anomalies in space-craft and their relationship to space weather, IEEE Trans. Plasma Sci.,28(6), 2007–2016, doi:10.1109/27.902228.
6. Koga S, Koshiishi H, Matsumoto H, *et al*. (2006) Prediction of electron flux variations at radiation belt (in Japanese). JAXA-SP-05, p 24.
7. Fukata M, Taguchi S, Okuzawa T, and Obara T (2002) Neural network prediction of relativistic electrons in geosynchronous orbit during the storm recovery phase: effects of recurring. Ann Geophys 20:947-951.
8. Nakamura Y, Kitamura K, Tokunaga M, *et al*. (2009) Prediction of electron flux at Geosynchronous Obit by using the Neural Network (in Japanese). JAXA-SP-09, p 6.
9. Akasofu, S.-I. (1981), Energy coupling between the solar wind and the magnetosphere, Space Sci. Rev., 28, 121.
10. Yoshitomi Y (2002), Neural network (in Japanese), Asakura shoten.

Authors Introduction

Ms. Ami Iwabu

She received her Bachelor's degree in Engineering in 2023 from the Faculty of Engineering, Kyushu Institute of technology in Japan. She is currently a master student in Kyushu Institute of Technology, Japan

Dr. Kentaro Kitamura

He is a Professor of Kyushu Institute of Technology, and director of Laboratory of Lean Satellite Enterprises and In-orbit Experiments (LaSEINE). He received his BS and MS degrees in science from Kyushu University, in 1996 and 1998, respectively. Then he received his Ph.D. degrees in science (Earth and Planetary Science) from Kyushu University in 2001.

Image Gradient-based Monocular Visual-Inertial Odometry

Tae Ihn Kim
Robotics Lab, Hyundai Motor Company,
Uiwang-si, Gyeonggi-do, 16082, Republic of Korea

Jae Hyung Jung
Dept. of Aerospace Engineering, Seoul National University / ASRI
Seoul, 08826, Republic of Korea

Chan Gook Park
Dept. of Aerospace Engineering, Seoul National University / ASRI
Seoul, 08826, Republic of Korea
E-mail: natigerya@snu.ac.kr, lastflowers@snu.ac.kr, chanpark@snu.ac.kr

Abstract

This paper presents image gradient-based monocular visual-inertial odometry (VIO) algorithm, using image gradient measurements, robust to illumination change. We expand the measurements from the reprojected feature locations on the image coordinates to the corresponding image gradients. The iterated EKF and low-pass pyramid are adapted to reduce the linearization error in the multi-state constraint Kalman filter (MSCKF) measurement update process. We verify that our proposed algorithm outperforms both conventional indirect and direct MSCKF-based VIO algorithms by evaluating the pose estimation performance.

Keywords: Illumination Change, Image Gradient, Iterated Extended Kalman Filter, Low-pass Pyramid, Visual-Inertial Odometry

1. Introduction

An accurate ego-motion estimation of a vehicle in a global navigation satellite system (GNSS) denied environment has been a challenging task in robotics and autonomous driving for several decades [1]. To tackle this issue, the ego-motion estimation algorithm using camera information, so-called Visual Odometry (VO), has been studied [2], [3], [4]. VO incrementally estimates a relative 6-DOF pose between consecutive images.

The most representative fusion of camera and alternative sensors is visual-inertial odometry (VIO), which additionally uses IMU; accelerometers, and gyroscopes. Despite various VIO algorithm studies, several challenging environments exist for ego-motion estimation, still being tackled. One representative environment is where the photometric consistency assumption is violated due to the illumination change.

The main contribution of this paper is threefold. First, inspired by the studies on image gradient-based VO algorithms [5]. we adapted the novel measurement, image gradient, to the VIO framework. Image gradient is known to be robust to illumination change. We utilize image gradients in the horizontal and vertical image directions as measurements of the multi-state constraint Kalman filter (MSCKF)-based VIO algorithm. Second, unlike the conventional MSCKF-based VIO algorithm,

we apply iterated extended Kalman filter (EKF) and low-pass pyramid to minimize linearization error. Lastly, we evaluate our proposed algorithm at both simulation and real-world datasets and demonstrate that our proposed algorithm outperforms both conventional indirect and direct methods in the environment where illumination changes.

2. Preliminaries

In this paper, the 3D pose is described at the current IMU-affixed frame, $\{B_k\}$, at time k with respect to a global frame, $\{G\}$. The global frame is the initial body frame aligned with gravity. Unlike EKF, MSCKF considers multiple images and utilizes the geometric constraint to estimate the ego-motion of the vehicle. Therefore, MSCKF includes the previous camera poses in the state. $\{C_i\}$ is a camera frame of the ith element in the sliding window of MSCKF.

The mathematical expression is described as follows. For example, $p_{C_l f_i}^{C_l}$ is the position of camera frame described as the translation from the lth camera frame, C_l to the ith 3D feature position, f_i with respect to C_l.

3. State Representation

The MSCKF state, X_{MSCKF} is formulated as a combination vector of the IMU state, X_{IMU}, and the previous camera state, X_{cam}, in the sliding window.

$$X_{MSCKF} = \begin{bmatrix} X_{IMU}{}^T & X_{camera}{}^T \end{bmatrix}^T \in \mathbb{R}^{(16+7N)\times 1} \tag{1}$$

$$X_{IMU} = \begin{bmatrix} q_{GB}{}^T & p_{GB}^{G}{}^T & v_{GB}^{G}{}^T & b_a{}^T & b_g{}^T \end{bmatrix}^T \in \mathbb{R}^{16\times 1} \tag{2}$$

$$X_{camera} = \begin{bmatrix} q_{GC_1}{}^T & p_{GC_1}^{G}{}^T & \dots & q_{GC_n}{}^T & p_{GC_n}^{G}{}^T \end{bmatrix}^T \in \mathbb{R}^{7N\times 1} \tag{3}$$

q_{GB} is the unit quaternion describing the rotation from body frame $\{B\}$ to the global frame $\{G\}$ (or the body attitude with respect to $\{G\}$, p_{GB}^{G} and v_{GB}^{G} are the body position and velocity with respect to $\{G\}$, respectively, b_a and b_g are accelerometer and gyroscope bias, respectively, and q_{GC_l} and $p_{GC_l}^{G}$ are the l th camera attitude and position, respectively.

True position, velocity, accelerometer and gyroscope bias are expressed as $X = \hat{X} + \tilde{X}$, where $\hat{X}$ and $\tilde{X}$ are estimated and error state, respectively, and true attitude as $q = \delta q \otimes \hat{q}$, where $\otimes$ denotes quaternion multiplication. Then, the quaternion error is derived as follows.

$$\delta q = q \otimes \hat{q}^{-1} \approx \begin{bmatrix} 1 \\ \frac{1}{2}\delta\theta \end{bmatrix} \tag{4}$$

Therefore, the error state is defined as below.

$$\tilde{X}_{MSCKF} = \begin{bmatrix} \tilde{X}_{IMU}^T & \tilde{X}_{cam}^T \end{bmatrix}^T \in \mathbb{R}^{(15+6N)\times 1} \tag{5}$$

$$\tilde{X}_{IMU} = \begin{bmatrix} \delta\theta_B^{G}{}^T & \tilde{p}_{GB}^{G}{}^T & \tilde{v}_{GB}^{G}{}^T & \tilde{b}_a{}^T & \tilde{b}_g{}^T \end{bmatrix} \in \mathbb{R}^{15\times 1} \tag{6}$$

$$\tilde{X}_{cam} = \begin{bmatrix} \delta\theta_{C_1}^{G}{}^T & \tilde{p}_{GC_1}^{G}{}^T & \dots & \delta\theta_{C_N}^{G}{}^T & \tilde{p}_{GC_N}^{G}{}^T \end{bmatrix} \in \mathbb{R}^{6N\times 1} \tag{7}$$

4. Measurement Model

In our proposed algorithm, the measurement of the conventional MSCKF, which is the reprojected feature location on the image coordinate, is extended to the image gradients along the horizontal and vertical direction of the image, $z_{gradient}$,

$$z_{gradient,i,l} = \begin{bmatrix} \nabla_x I_{i,l}(\pi(p_{C_l f_i}^{C_l})) \\ \nabla_y I_{i,l}(\pi(p_{C_l f_i}^{C_l})) \end{bmatrix} + n_{gradient,i,l} \tag{8}$$

where $p_{C_l f_i}^{C_l} = h(p_{Gf_i}^{G}, \theta_{C_l}^{G}, p_{GC_l}^{G})$. $z_{gradient,i,l}$ is the measurement of the ith feature on the lth image. $\pi(\cdot)$ is a reprojected feature location, $I_{i,l}(\pi(\cdot))$ is an intensity at the reprojected feature location, $\nabla_x I_{i,l}(\pi(\cdot))$ and $\nabla_y I_{i,l}(\pi(\cdot))$ are gradients along horizontal and vertical directions of image. $n_{gradient,i,l}$ is the noise of the ith feature on the lth image.

5. Measurement Update

The measurement update of our proposed algorithm proceeds when either a feature is lost from tracking or the length of the feature track exceeds the size of the sliding window, which are the same measurement update conditions of the conventional MSCKF-based VIO algorithm [6], [7]. Since the novel measurement, image gradients along the horizontal and vertical direction of the

image are utilized, the residual is expanded to the difference between the anchor and the lth image gradient along two directions as,

$$
\begin{aligned}
r_{i,l} &= z_{i,l} - \hat{z}_{i,l} \\
&= \begin{bmatrix} \nabla_x I_{i,A} - \nabla_x I_{i,l}(\pi(\hat{p}_{C_l f_i}^{C_l})) \\ \nabla_y I_{i,A} - \nabla_y I_{i,l}(\pi(\hat{p}_{C_l f_i}^{C_l})) \end{bmatrix} \\
&\approx H_{X_{i,l}} \tilde{X}_k + H_{f_{i,l}} \tilde{p}_{Gf_i}^{G} + n_{i,l}
\end{aligned}
\tag{9}
$$

where $\hat{p}_{C_l f_i}^{C_l} = h(\hat{p}_{Gf_i}^{G}, \hat{\theta}_{C_l}^{G}, \hat{p}_{GC_l}^{G})$. $\nabla_x I_{i,A}$ and $\nabla_y I_{i,A}$ are the image gradients along the horizontal and vertical direction of the anchor image. The anchor frame is defined as the camera image where the ith feature is captured in the sliding window. $r_{i,l}$, $z_{i,l}$ and $\hat{z}_{i,l}$ are the residuals, measured gradients and estimated gradients of the ith feature on the lth image, respectively. $H_{X_{i,l}}$ and H_{f_i} are state and feature position Jacobian matrix, respectively. $\tilde{X}_k$ is the error state at step k and $\tilde{p}_{Gf_i}^{G}$ is the error feature position.

The residuals of the ith feature on all image frames are accumulated to form a block vector of residuals of the ith feature. Since the measurement is extended from the image feature projection location on the image coordinates to image gradients, Jacobian matrices of state, $H_{X_{C_l}}$, and feature position, $H_{p_{Gf_i}^{G}}$, are newly derived as below.

$$
\begin{aligned}
H_{x_{C_l}} &= \frac{\nabla^2 I_l}{\partial(u_d, v_d)} \frac{\partial(u_d, v_d)}{\partial p_{C_l f_i}^{C_l}} \frac{\partial p_{C_l f_i}^{C_l}}{\partial X_{C_l}}, \\
H_{p_{Gf_i}^{G}} &= \frac{\nabla^2 I_l}{\partial(u_d, v_d)} \frac{\partial(u_d, v_d)}{\partial p_{Gf_i}^{G}}, \\
\frac{\nabla^2 I_l}{\partial(u_d, v_d)} &= \begin{bmatrix} \frac{\nabla_x(\nabla_x I_l)}{\partial(u_d)} & \frac{\nabla_x(\nabla_x I_l)}{\partial(v_d)} \\ \frac{\nabla_x(\nabla_y I_l)}{\partial(u_d)} & \frac{\nabla_y(\nabla_y I_l)}{\partial(v_d)} \end{bmatrix}
\end{aligned}
\tag{10}
$$

∇_x and ∇_y are gradients along horizontal and vertical direction of the image, respectively. u_d and v_d are undistorted image pixels along horizontal and vertical direction of image, respectively. X_{C_l} is the lth camera pose. $\frac{\nabla^2 I_l}{\partial(u_d, v_d)}$ implies the gradients of the gradient image. Then, the residual is projected to a left null-space to eliminate the terms related to the 3D feature position. The block vectors of the residuals of the ith feature are accumulated to form a block vector of all residuals and proceed QR decomposition to reduce computational complexity.

Unlike EKF, which utilizes prior estimate as a linearization point, iterated EKF iteratively computes posterior estimate, which is more accurate since the measurement is reflected. The MSCKF measurement update equation is summarized as follows.

$$
\begin{aligned}
\delta X_k^+ &= K_k(y_k - h(\hat{X}_k^-)) \\
K_k &= P_k^- T_{H_k}{}^T (T_{H_k} P_k^- T_{H_k}{}^T + Q_{1_k}{}^T R_{0_k} Q_{1_k})^{-1} \\
r_{n_k} &= y_k - h(\hat{X}_k^-) = Q_{1_k}{}^T r_{0_k}
\end{aligned}
\tag{11}
$$

δX_k^+ is a posterior estimated state, which is the correction term with respect to the error state from $\hat{X}_k^+ = \hat{X}_k^- \boxplus \delta X_k^+$. $\boxplus$ is the operator defining the error state as the estimated error state and the correction term. $h(\hat{X}_k^-)$ is the measurement equation of the system, K_k is the

Kalman gain, and P_k^- is the prior covariance of the estimated error. T_{H_k} is the upper triangular matrix computed from the QR decomposition. Q_{1_k} is the unitary matrix computed from the QR decomposition, whose columns form bases for the range of H_X. y_k is the measurement, and R_{0_k} is the covariance matrix of the measurement.

The EKF-based MSCKF uses the prior estimate state as the linearization point. Therefore, the measurement is not reflected on the linearization point. However, the iterated EKF-based MSCKF first computes the posterior estimate state using the prior estimate state to reflect the measurement on the linearization point and iteratively computes the posterior estimate state to reduce the linearization point. Therefore, the adaptation of the iterated EKF to MSCKF is made through modification of equation 11 to

$$\delta X_{k,i+1}^+ = K_{k,i}(y_k - h(\hat{X}_{k,i}^+) + T_{H_{k,i}}(\delta X_{k,i}^+))$$
$$K_{k,i} = P_k^- T_{H_{k,i}}^T (T_{H_{k,i}} P_k^- T_{H_{k,i}}^T + Q_{1_{k,i}}^T R_{0,k,i} Q_{1_{k,i}})^{-1} \quad (12)$$
$$r_{n_{k,i}} = y_k - h(\hat{X}_{k,i}^+) + T_{H_{k,i}}(\delta X_{k,i}^+) = Q_{1_k}^T r_{0_k} + T_{H_{k,i}}(\delta X_{k,i}^+)$$

$\delta X_{k,i+1}^+$ is the correction term at the $i+1$th iteration and $h(\hat{X}_{k,i}^+)$ is the measurement equation of system, where posterior is reflected. $K_{k,i}$, $T_{H_{k,i}}$, and $Q_{1_{k,i}}$ are the Kalman gain, the upper triangular matrix, and the unitary matrix, respectively, at the ith iteration. The combination of the iterated EKF and the low-pass pyramid results in a reduction of the residual, which implies the improvement of linearization error due to the high non-linearity of the image gradients.

Fig. 1. Attitude and position RMSE

6. Experimental Results

The proposed algorithm is evaluated at a real-world dataset to verify that it is more robust to the illumination change than the feature-based and intensity-based methods. EuRoC real-world dataset is collected from an on-board hex-rotor helicopter. This dataset provides stereo images, IMU data, and 6-DOF ground-truth. Illumination change due to the camera's automatic exposure time is included in the image. The performance of our proposed algorithm is analyzed on five different

sequences collected in the same indoor environment, and MAV travels 94.86 m in 134.8 s on average.

The experiment result is summarized in Fig. 1. It reports that the average 3D position RMSE of the feature-based method is less than that of the intensity-based method. This result implies that the illumination change strongly influences the image intensity, while the reprojected feature position is less influenced. The feature-based method is less influenced because the extracted feature position's accuracy depends on motion-blur, not the illumination change. Average 3D position RMSE of the intensity-based method is reduced by utilizing the image gradient

7. Conclusion

In this paper, we have proposed the utilization of image gradients as measurements in the monocular MSCKF-based VIO algorithm with an adaptation of iterated EKF with the low-pass pyramid. The novel measurement is used to tackle the illumination change environment, where the photometric-consistency assumption is violated. Unlike the conventional method, which utilizes the image feature coordinates, our proposed algorithm utilizes the image gradients along the horizontal and vertical direction of the image with the adaptation of iterated EKF and low-pass pyramid. The performance of the proposed algorithm is analyzed by comparing it to that of feature-based and intensity-based methods at the real-world dataset. Our experimental result reports that our proposed algorithm's accuracy of both 3D attitude and position estimation outperforms both feature-based and intensity-based algorithms.

References

1. M. Tang, Z. Chen, "Robot Tracking in SLAM with Masreliez-Martin Unscented Kalman Filter", International Journal of Control, Automation and Systems, vol. 18, no. 9, pp. 2315-2325, 2020.
2. F. Zhang, D. Clarke and A. Knoll, "Visual Odometry based on a Bernoulli Filter", International Journal of Control, Automation and Systems, vol. 13, no. 3, pp. 530-538, 2015.
3. W. Jeong, J. Moon, and B. Lee, "Error improvement in visual odometry using super-resolution", International Journal of Control, Automation and Systems, vol. 18, no. 2, pp. 322-329, 2020.
4. L. Wang, S. Xia, H. Xi, S. Li, and L. Wang, "Robust Visual Inertial Odometry Estimation Based on Adaptive Interactive Multiple Model Algorithm", International Journal of Control, Automation and Systems, vol. 20, no. 10, pp. 3335-3346, 2022.
5. J. Quenzel, R. A. Rosu, T. Labe, C. Stachniss and S. Behnke, "Beyond Photometric Consistency: Gradient-based Dissimilarity for Improving Visual Odometry and Stereo Matching", 2020 IEEE International Conference on Robotics and Automation (ICRA), pp. 272-278, 2020.

6. M. Li and A. I. Mourikis, "High-precision, Consistent EKF-based Visual-Inertial Odometry", The International Journal of Robotics Research, vol. 32, no. 6, pp. 690-711, May 2013.
7. A. I. Mourikis and S. I. Roumeliotis, "A Multi-State Constraint Kalman Filter for Vision-aided Inertial Navigation", 2007 IEEE International Conference on Robotics and Automation (ICRA), pp. 3565-3572, 2007.

Authors Introduction

Mr. Tae Ihn Kim

He received the M.S. degree in the Department of Aerospace Engineering from Seoul National University, Republic of Korea, in 2021. He received the B.S. degree in Mechanical and Aerospace Engineering from Seoul National University, Republic of Korea, in 2019. His research interests include Visual Odometry and Visual-Inertial Odometry.

Mr. Jae Hyung Jung

He received the M.S., and Ph.D. degrees in the Department of Aerospace Engineering from Seoul National University, Republic of Korea. He received the B.S. degree in the Department of Aerospace Engineering from Pusan National University, Republic of Korea, in 2017. His research interests include visual odometry and vision-aided inertial navigation for mobile robots.

Dr. Chan Gook Park

He received the B.S., M.S., and Ph.D. degrees in control and instrumentation engineering from the Seoul National University, Seoul, South Korea, in 1985, 1987, and 1993, respectively. He was as a Post-Doctoral Fellow with Prof. J. L. Speyer about peak seeking control for formation flight at the University of California at Los Angeles, Los Angeles, CA, USA, in 1998. From 1994 to 2003, he was an Associate Professor with Kwangwoon University, Seoul. In 2003, he joined the Faculty with the School of Mechanical and Aerospace Engineering, National University, Seoul, where he is currently a professor. In 2009, He was a Visiting Scholar with the Department of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA, USA. He served as a chair of IEEE AES Korea Chapter until 2009. His current research topics include advanced filtering techniques, high precision inertial navigation system (INS), visual-inertial odometry (VIO), INS/GNSS/IBN integration, and smartphone-based/foot-mounted pedestrian dead reckoning (PDR) systems.

A Method for Embedding Multiple Photographic Images in a Photographic Image

Naoki Kouno
Department of Information Systems, University of Nagasaki, Japan

Kanya Goto
Graduate School of Regional Design and Creation, University of Nagasaki, Japan

Toru Hiraoka
Department of Information Systems, University of Nagasaki, Japan
E-mail: hiraoka@sun.ac.jp

Abstract

A method has been proposed for embedding another photographic image (image B) in a photographic image (image A). An image (image A') is generated by embedding information of image B in image A, and then an image (image B') is restored by extracting information from image A'. The conventional methods can only embed one photographic image in one photographic image. Therefore, we extend the conventional method and propose a method for embedding multiple photographic images in one photographic image. In our method, as more images B are embedding in image A, the image quality of image A' deteriorates, but the image quality of images B' dose not deteriorate. To verify the effectiveness of our method, experiments using various photographic images were performed. As a result of the experiments, the relationship between the number of images B and the image quality of image A' was clarified.

Keywords: Multiple photographic images, Embedding, Restoring, Digital watermark

1. Introduction

Digital watermarking [1] methods have been proposed that uses two photographic images of the same size with 256 gradations and embed another photographic image in one of the photographic images [2], [3]. The conventional methods [2], [3] embed one photographic image (image B) in one photographic image (image A). An image (image A') is generated by embedding information of image B in image A, and then an image (image B') is restored by extracting information from image A'. Images A and A' are similar, and images B and B' are similar. To make people unaware that image B is embedded in image A', the difference between images A and A' must be visually unrecognizable. In the conventional method [2], each pixel value of pixels in image B whose horizontal and vertical positions are odd numbers is embedded in 4 pixels in image A. At that time, the pixel values in image A are changed within ± 2, and then image A' is generated. In the conventional method [3], each pixel value of pixels in image B whose horizontal and vertical positions are the values of the tolerance sequence with first term 2 and tolerance 3 is embedded in 8 pixels in image A. At that time, the pixel values in image A are changed within ± 1, and then image A' is generated. It is known through experiments that the conventional method [3] reduces the difference between images A and A' about 60% compared to the conventional method [2]. On the other hand, the conventional method [3] is known to have a larger difference between images B and B' than the conventional method [2]. In the conventional methods [2], [3], it is important not to be aware that image B is embedded in image A', so we focus on the conventional method [3].

Since the conventional method [3] can only embed one photographic image in one photographic image, we extend the conventional method [3] and propose a method for embedding multiple photographic images in one photographic image. In our method, as more images B are embedding in image A, the image quality of image A' deteriorates, but the image quality of images B' dose not deteriorate. To verify the effectiveness of our method, experiments using various photographic images were performed. As a result of the experiments, the relationship between the number of images B and the image quality of image A' was clarified.

2. Our Method

Our method generates image A' by embedding images B_o ($o = 1, 2, \cdots, O$) in image A, where O is the number of images B_o to be embedded in image A, and then images B'_o are restored from image A'. A conceptual diagram of our method is shown in Fig. 1. The pixel values for spatial coordinates (i, j) ($i = 1, 2, \cdots, I; j = 1, 2, \cdots, J$) of images A, B, A' and B' are defined as $f_{A,i,j}$, $f_{B_o,i,j}$, $f_{A',i,j}$ and $f_{B'_o,i,j}$, respectively. Images A, B_o, A' and B'_o are the

Naoki Kouno, Kanya Goto, Toru Hiraoka

same size and have 256 gradations from 0 to 255. The methods for embedding images B_o in image A and restoring images B'_o from image A' are described below.

Fig. 1. Conceptual diagram of our method

2.1. *Embedding method*

Let the binary representations of $f_{B_o,k,l}$ be $b_{B_o,p,k,l}$ ($p = 1,2,\cdots,8$), and the following relationship (Eq. (1)) is established, where k ($= 2,5,8,\cdots$) and l ($= 2,5,8,\cdots$) are the values of the tolerance sequence with first term 2 and tolerance 3.

$$f_{B_o,k,l} = \sum_{p=1}^{8} 2^{8-p} b_{B_o,p,k,l} \tag{1}$$

Information is respectively embedded in $f_{A,k-1,l-1}$, $f_{A,k,l-1}$, $f_{A,k+1,l-1}$, $f_{A,k-1,l}$, $f_{A,k+1,l}$, $f_{A,k-1,l+1}$, $f_{A,k,l+1}$ and $f_{A,k+1,l+1}$ using the values $b_{B_o,1,k,l}$, $b_{B_o,2,k,l}$, $b_{B_o,3,k,l}$, $b_{B_o,4,k,l}$, $b_{B_o,5,k,l}$, $b_{B_o,6,k,l}$, $b_{B_o,7,k,l}$ and $b_{B_o,8,k,l}$, and then $f_{A',k-1,l-1}$, $f_{A',k,l-1}$, $f_{A',k+1,l-1}$, $f_{A',k-1,l}$, $f_{A',k+1,l}$, $f_{A',k-1,l+1}$, $f_{A',k,l+1}$ and $f_{A',k+1,l+1}$ are generated. Note that $f_{A',k,l}$ is the same value as $f_{A,k,l}$. Since $f_{A',k-1,l-1}$, $f_{A',k,l-1}$, $f_{A',k+1,l-1}$, $f_{A',k-1,l}$, $f_{A',k+1,l}$, $f_{A',k-1,l+1}$, $f_{A',k,l+1}$ and $f_{A',k+1,l+1}$ are calculated in the same procedure, only the case of $f_{A',k-1,l-1}$ is described below. The values $c_{A,k-1,l-1}$, which are integers greater than or equal to 0 and less than 2^O, are calculated from the pixel values $f_{A,k-1,l-1}$ by the following Eq. (2).

$$c_{A,k-1,l-1} = f_{A,k-1,l-1} \% 2^O \tag{2}$$

where % represents a remainder operation. The values $c_{B,k-1,l-1}$, which are integers greater than or equal to 0 and less than 2^O, are calculated from the values $b_{B_o,1,k,l}$ by the following Eq. (3).

$$c_{B,k-1,l-1} = \sum_{o=1}^{O} 2^{O-o} b_{B_o,1,k,l} \tag{3}$$

The values $f_{A',k-1,l-1}$ are calculated by the following Eq. (4).

$$f_{A',k-1,l-1} = f_{A,k-1,l-1} - c_{A,k-1,l-1} + c_{B,k-1,l-1} \tag{4}$$

If $f_{A',k-1,l-1}$ is smaller than 0, we must add 2^O to $f_{A',k-1,l-1}$. If $f_{A',k-1,l-1}$ is greater than 255, we must subtract 2^O from $f_{A',k-1,l-1}$.

When embedding images B_o in image A, our method changes the pixel values of image A within plus or minus $2^O - 1$, and then generates image A'. As more images B_o are embedding in image A, the image quality of image A' deteriorates.

2.2. *Restoring method*

Let the binary representations of $f_{A',i,j} \% 2^O$ be $b_{A',o,i,j}$, and the following relationship (Eq. (5)) is established.

$$f_{A',i,j} \% 2^O = \sum_{o=1}^{O} 2^{O-o} b_{A',o,i,j} \tag{5}$$

The pixel values $f_{B'_o,k,l}$ for spatial coordinates (k,l) are restored by the following Eq. (6).

$$\begin{aligned}
f_{B'_o,k,l} = \ &128 b_{A',o,k-1,l-1} + 64 b_{A',o,k,l-1} + \\
&32 b_{A',o,k+1,l-1} + 16 b_{A',o,k-1,l} + 8 b_{A',o,k+1,l} + \\
&4 b_{A',o,k-1,l+1} + 2 b_{A',o,k,l+1} + b_{A',o,k+1,l+1}
\end{aligned} \tag{6}$$

The pixel values $f_{B'_o,k',l'}$ for spatial coordinates (k',l') other than (k,l) are restored by the following Eq. (7) and Eq. (8).

$$f_{B'_o,k',l'} = \frac{\sum_{m=-2}^{2}\sum_{n=-2}^{2} f_{B'_o,k'+m,l'+n} d_{m,n}}{d_{m,n}} \tag{7}$$

$$d_{m,n} = \frac{1}{\sqrt{m^2+n^2}} \tag{8}$$

where m and n are the relative positions from the target pixel. Note that Eq. (7) is calculated using only the pixel values $f_{B'_o,k,l}$ obtained in Eq. (6).

3. Experiments

In this experiment, Lena image shown in Fig. 2 was used as image A, and 4 photographic images shown in Fig. 3 were used as images B_o. The size of all image A and images B_o was 256 * 256 pixels. Visual and quantitative evaluations were performed to verify the effectiveness of our method.

Fig. 2. Lena image (image A)

(a) Airplane   (b) Barbara   (c) Boat   (d)Cameraman
Fig. 3. Various photographic images (images B_o)

3.1. *Evaluation of embedded images*

Averages of absolute values of the differences between the pixel values of image A and images A' were calculated. Hereinafter, the averages are referred to as difference averages. The difference averages for 4 cases of embedding one image B_o are shown in Table 1. The difference averages for 6 cases of embedding 2 images B_o are shown in Table 2. For example, "(a)+(b)" in Table 1 means that 2 images B_o (a) and (b) of Fig. 3 are embedded. The difference averages for 4 cases of embedding 3 images B_o are shown in Table 3. The difference average for one case of embedding 4 images B_o is shown in Table 4. From Table 1 to 4, it was found that the larger the number of images B_o to be embedded in image A, the larger the difference averFiage.

Table 1. Difference averages between image A and images A' for 4 cases of embedding one image B_o

Symbols of images B_o	Difference averages
(a)	0.443
(b)	0.442
(c)	0.443
(d)	0.442

Table 2. Difference averages between image A and images A' for 6 cases of embedding 2 images B_o

Symbols of images B_o	Difference averages
(a) + (b)	1.093
(a) + (c)	1.100
(a) + (d)	1.092
(b) + (c)	1.099
(b) + (d)	1.096
(c) + (d)	1.127

Table 3. Difference averages between image A and images A' for 4 cases of embedding 3 images B_o

Symbols of images B_o	Difference averages
(a) + (b) + (c)	2.302
(a) + (b) + (d)	2.291
(a) + (c) + (d)	2.320
(b) + (c) + (d)	2.316

Table 4. Difference averages between image A and images A' for one case of embedding 4 images B_o

Symbols of images B_o	Difference averages
(a) + (b) + (c) + (d)	4.711

Next, the differences between image A and images A' were visually evaluated. Images A' for 4 cases of embedding one image B_o are shown in Fig. 4, Images A' for 6 cases of embedding 2 images B_o are shown in Fig. 5. Images A' for 4 cases of embedding 3 images B_o are shown in Fig. 6. Image A' for one case of embedding 4 images B_o is shown in Fig. 7. From Fig. 2 and 4 to 7, it was found that images A', in which one and 2 images B_o were embedded, were visually unrecognizable as different from image A. It was also found that images A',

in which one or 3 images B_o were embedded, were slightly recognizable as different from image A, and image A', in which one or 4 images B_o were embedded, were clearly recognizable as different from image A.

(a)　　　　(b)　　　　(c)　　　　(d)

Fig. 4. Images A' for 4 cases of embedding one image B_o

(a) + (b)　　(a) + (c)　　(a) + (d)　　(b) + (c)

(b) + (d)　　(c) + (d)

Fig. 5. Images A' for 6 cases of embedding 2 images B_o

(a) + (b) + (c) (a) + (b) + (d) (a) + (c) + (d) (b) + (c) + (d)

Fig. 6. Images A' for 4 cases of embedding 3 images B_o

(a) + (b) + (c) + (d)

Fig. 7. Images A' for one case of embedding 4 images B_o

3.2. *Evaluation of restored images*

Difference averages between images B_o and images B'_o restored from images A' were calculated. The difference averages in this case are the same value no matter how many images B_o are embedded. The difference averages are shown in Table 5. Additionally, images B'_o are shown in Fig. 8. Images B'_o do not change no matter how many images B_o are embedded. From Table 5, it was found that the difference averages were somewhat larger, ranging from 6 to 13 values. However, from Fig. 3 and 8, it was found that images B'_o are fully recognizable as same as images B_o, although images B'_o were disturbed at the edges.

Naoki Kouno, Kanya Goto, Toru Hiraoka

Table 5. Difference averages between images B_o and images B'_o

Symbols of images B_o	Difference averages
(a)	8.265
(b)	12.717
(c)	6.642
(d)	8.191

Fig. 8. Images B'_o

4. Conclusion

We proposed a method for embedding multiple photographic images (images B_o) in one photographic image (image A). An image (image A') was generated by embedding information of images B_o in image A, and then an image (images B'_o) was restored by extracting information from image A'. In our method, as more images B were embedding in image A, the image quality of image A' deteriorated, but the image quality of images images B'_o did not deteriorate. To verify the effectiveness of our method, experiments using various photographic images were performed. As a result of the experiments, it was found that images A', in which one and 2 images B_o were embedded, were visually unrecognizable as different from image A. Additionally, it was found that images A', in which one or 3 images B_o were embedded, were slightly recognizable as different from image A, and image A', in which one or 4 images B_o were embedded, were clearly recognizable as different from image A.

A future task is to be able to apply our method to photographic images of different sizes.

Acknowledgements

References

1. A. C. Pavan and M. T. Somashekara, An overview on research trends, challenges, applications and future direction in digital image watermarking, International Research Journal on Advanced Science Hub 5(1), 2023, pp. 2582-4376.
2. T. Hiraoka, A method for embedding another photographic image in a photographic image, ICIC Express Letter 14(4), 2020, pp. 311-317.
3. T. Hiraoka and K. Goto, A method to reduce changes in photographic images with embedded images, ICIC Express Letter.

Authors Introduction

Mr. Naoki Kouno

He is a student in the Department of Information Systems, University of Nagasaki. His research interests inculede image processing and data analysis.

Mr. Kanya Goto

He received B.Eco. degree from Keio University in 2006. He is currently a CEO in Movenext Corporation. His research interests include non-photorealistic rendering and electronic watermark.

Dr. Toru Hiraoka

He received B.Des., M.Des. and D.Eng. degrees from Kyushu Institute of Design in 1995, 1997 and 2005, respectively. He is currently a Professor in University of Nagasaki. His research interests include non-photorealistic rendering and disaster prevention.

Generation of Flowing-Line Images Using Vertical and Horizontal Smoothing Filters

Karin Kuroki
Department of Information Systems, University of Nagasaki, Japan

Toru Hiraoka
Department of Information Systems, University of Nagasaki, Japan
E-mail: hiraoka@sun.ac.jp

Abstract

We propose a non-photorealistic rendering method for automatically generating flowing-line images from photographic images. Flowing-line images consist of unidirectional lines with smooth curves. Our method is executed by an iterative calculation using vertical and horizontal smoothing filters. To verify the effectiveness of our method, we conducted an experiment using various photographic images to confirm that flowing-line patterns can be generated on the entire image. Additionally, we conducted an experiment to visually examine how flowing-line patterns generated by changing the values of the parameters in our method change.

Keywords: Non-photorealistic rendering, Flowing line, Vertical smoothing filter, Horizontal smoothing filter, Automatic generation

1. Introduction

Many studies [1],[2] have been conducted on non-photorealistic rendering using computer technology. Since non-photorealistic rendering targets all expressions other than realistic expressions by photorealistic rendering, there are various methods for dealing with various styles and uses. We focus on non-photorealistic rendering methods that use image processing to generate non-photorealistic images that are different from conventional art expressions such as oil paintings and watercolors. As non-photorealistic rendering methods of the unprecedented art expressions, many methods for generating labyrinth images [3], cell-like images [4], and moire-like images [5] from photographic images have been proposed. Labyrinth images are composed of equally spaced lines like a maze and are generated using minimum spanning trees, cell-like images are composed of cell-like patterns with cell membranes and cell nuclei and are generated using inverse iris filter, and moire-like images are a kind of op art and are generated using bilateral filter and unsharp musk.

We propose a non-photorealistic rendering method for automatically generating flowing-line images from photographic images. Flowing-line images are composed of unidirectional (here vertical) lines with smooth curves. The black lines are thickly represented at the edges and black areas of photographic images, and are represented along the edges in the nearly vertical direction. Our method is executed by an iterative calculation using vertical and horizontal smoothing filters. As images similar to flowing-line images, non-photorealistic rendering methods for generating ripple images has been proposed [6],[7]. Ripple images are composed of continuous lines with fluctuations and are generated using inverse Sobel filter [6] or intensity gradients [7]. Flowing-line images are represented by the lines closer to the binary value of black and white than ripple images, and flowing-line images and ripple images have different impressions. The conventional method [8] is executed by an iterative calculation using circular-sector-type smoothing filter and inverse filter [9]. Ripple images [8] are more similar to flowing-line images than ripple images [6],[7], but flowing-line patterns are slightly more linear than ripple patterns. In the conventional methods [6],[7],[8], ripple patterns cannot be generated in white and black areas. On the other hand, our method can generate flowing-line patterns in the white and black areas. Through an experiment using various photographic images, our method examines that flowing-line patterns can be generated on the entire image. Additionally, through an experiment that changes the values of the parameters in our method, how the generated flowing-line patterns change is examined.

2. Our method

Our method is largely executed in three steps. First, since flowing-line patterns are unlikely to occur in the white and black areas, gray scale transformation is performed on a photographic image. The gray scale transformation eliminates pixel values close to white and black, and eliminates areas where flowing-line patterns cannot be generated. Second, processing is performed

Karin Kuroki, Toru Hiraoka

using horizontal smoothing filter and inverse filter. Third, processing is performed using vertical smoothing filter. A flowing-line image is generated by iterating through the second and third steps. A flow chart of our method is shown in Fig. 1.

Fig. 1. Flow chart of our method

Details of the steps in Fig. 1. are explained below.

Step 0: Let the input pixel values on coordinates (i, j) of a gray-scale photographic image be $o_{i,j}$. The pixel values $o_{i,j}$ have value of U gradation from 0 to $U - 1$.

Step 1: Let the pixel values that have undergone gray scale transformation as shown in the following equation be $f_{i,j}$.

$$f_{i,j} = o_{i,j} \frac{U - 2D - 1}{U - 1} + D \qquad (1)$$

where D is a natural number. The larger the value of D, the fewer pixels that are close to white or black. When the value of D is 0, the pixel values $f_{i,j}$ have the same pixel values as the original photographic image.

Step 2: Let the pixel values of the image at the t-th iteration number be $f_{i,j}^{(t)}$, where $f_{i,j}^{(0)} = f_{i,j}$. Let the pixel values that horizontal smoothing filter is applied as shown in the following equation be $s_{1,i,j}^{(t)}$.

$$s_{1,i,j}^{(t)} = \frac{\sum_{k=-W_1}^{W_1} f_{i+k,j}^{(t-1)}}{2W_1 + 1} \qquad (2)$$

where W_1 is the window size and k is the positions in the window. Let the pixel values that inverse filter is applied as shown in the following equation be $g_{i,j}^{(t)}$.

$$g_{i,j}^{(t)} = f_{i,j}^{(t-1)} - s_{1,i,j}^{(t)} + f_{i,j} \qquad (3)$$

If $g_{i,j}^{(t)}$ is smaller than 0, then $g_{i,j}^{(t)}$ must be set to 0, and if $g_{i,j}^{(t)}$ is greater than $U - 1$, then $g_{i,j}^{(t)}$ must be set to $U - 1$.

Step 3: Let the pixel values that vertical smoothing filter is applied as shown in the following equation be $f_{i,j}^{(t)}$.

$$f_{i,j}^{(t)} = \frac{\sum_{l=-W_2}^{W_2} g_{i,j+l}^{(t)}}{2W_2 + 1} \qquad (4)$$

where W_2 is the window size and l is the positions in the window.

A flowing-line image is obtained after T times iteration of Steps 2 and 3.

3. Experiments

We conducted two experiments. First, we visually confirmed flowing-line patterns by changing the values of the parameters in our method using Lighthouse image shown in Fig. 2. Second, we applied our method to eight photographic images shown in Fig. 3. All photographic images used in the experiments were 512 * 512 pixels and 256 gradations.

Fig. 2. Lighthouse image

Fig. 3. Photographic images

3.1. *Experiment with changing parameter values*

Flowing-line images generated by changing the iteration number T were visually confirmed using Lighthouse image. The value of T was set to 5, 10, 20, and 40. The values of the other parameters D, W_1, and W_2 were set to 10, 6, and 3, respectively. The results of the experiment are shown in Fig. 4. As the value of T was larger, flowing-line patterns became clearer and were expressed finely.

Flowing-line images generated by changing the value of the parameter D were visually confirmed using Lighthouse image. The value of D was set to 0, 10, 20, and 30. The values of the other parameters T, W_1, and W_2 were set to 40, 6, and 3, respectively. The results of the experiment are shown in Fig. 5. As the value of D was larger, flowing-line patterns became clearer in the white area at the bottom of Lighthouse image. On the other hand, as the value of D was larger, it became more difficult to recall Lighthouse image.

Flowing-line images generated by changing the window size W_1 were visually confirmed using

Lighthouse image. The window size W_1 was set to 2, 4, 6, and 8. The values of the other parameters T, D, and W_2 were set to 40, 10, and 3, respectively. The results of the experiment are shown in Fig. 6. As the value of W_1 was larger, the interval of flowing-line patterns became wider.

Flowing-line images generated by changing the window size W_2 were visually confirmed using Lighthouse image. The window size W_2 was set to 1, 2, 3, and 4. The values of the other parameters T, D, and W_1

were set to 40, 10, and 6, respectively. The results of the experiment are shown in Fig. 7. As the value of W_2 was larger, flowing-line patterns became smoother. Furthermore, in Figs. 4 to 7, it was also found that flowing-line patterns can be generate in the white area at the bottom of Lighthouse image.

(a) $T = 5$ (b) $T = 10$

(c) $T = 20$ (d) $T = 40$

Fig. 4. Flowing-line images in the case of the iteration number $T = 5$, 10, 20, and 40

(a) $D = 0$ (b) $D = 10$

(c) $D = 20$ (d) $D = 30$

Fig. 5. Flowing-line images in the case of the parameter value $D = 0$, 10, 20, and 30

(a) $W_1 = 2$ (b) $W_1 = 4$

(c) $W_1 = 6$ (d) $W_1 = 8$

Fig. 6. Flowing-line images in the case of the window size $W_1 = 2$, 4, 6, and 8

(a) $W_2 = 1$ (b) $W_2 = 2$

(c) $W_2 = 3$ (d) $W_2 = 4$

Fig. 7. Flowing-line images in the case of the window size $W_2 = 1$, 2, 3, and 4

Karin Kuroki, Toru Hiraoka

3.2. *Experiment using eight photographic images*

Our method was applied to eight photographic images shown in Fig. 3. The values of the parameters T, D, W_1, and W_2 were set to 40, 10, 6, and 3, respectively. The results of the experiment are shown in Fig. 8. For all flowing-line images, flowing-line patterns were composed of unidirectional (here vertical) lines with smooth curves, and could be automatically generated on the entire image. Looking at more details, the black lines were thickly represented at the edges and black areas (clothes, hair, etc. in the lower left image) of photographic images, and were represented along the edges in the nearly vertical direction.

Fig. 8. Flowing-line images

4. Conclusion

We proposed a non-photorealistic rendering method for generating flowing-line images from gray-scale photographic images using vertical and horizontal smoothing filters. Through an experiment using nine photographic images, it was found that our method can automatically generate flowing-line patterns on the entire images. Additionally, through an experiment that the values of the parameters in our method were changed, it was found that the interval and smoothness of flowing-line patterns can be changed.

A subject for future study is to expand our method for application to color photographic images, videos, and three-dimensional data.

Acknowledgements

References

1. P. Haeberli. Paint by numbers, abstract image representations, ACM SIGGRAPH Computer Graphics 24(4), 1990, pp. 207-214.
2. J. Lansdown and S. Schofield, Expressive rendering, a review of nonphotorealistic techniques, IEEE Computer Graphics and Applications 15(3), 1995 pp. 29-37.
3. K. Inoue and K. Uraham, Halftoning with minimum spanning trees and its application to maze-like images, LSEVIER Computers & Graphics 33(5), 2009, pp. 638-647.
4. T. Hiraoka, M. Hirota, K. Inoue and K. Urahama, Generating cell-like color images by inverse iris filter, ICIC Express Letters 11(2), 2017, pp.399-404.
5. T. Hiraoka and K. Urahama, Generation of moire-picture-like color images by bilateral filter, IEICE Transactions on Information and Systems E96-D(8), 2013, pp. 1862-1866.
6. T. Hiraoka and K. Urahama, Generating ripple-pattern-like images by inverse Sobel filter, Journal of the Institute of Image Electronics Engineers of Japan 44(2), 2015, pp. 349-352.
7. T. Hiraoka, Generation of ripple images using intensity gradient, ICIC Express Letters 13(12), 2019, pp. 1097-1101.
8. T. Hiraoka, Generation of arbitrarily-oriented ripple images using circular-sector-type smoothing filter and inverse filter, Journal of Robotics, Networking and Artificial Life 6(4), 2020, pp. 213-216.
9. Z. Yu and K. Urahama, Iterative method for inverse nonlinear image processing, IEICE Transactions on Fundamentals E97-A(2), 2014, pp. 719-721.

Authors Introduction

Ms. Karin Kuroki

She is a student in the Department of Information Systems, University of Nagasaki. His research interests inculede image processing and non-photorealistic rendering.

Dr. Toru Hiraoka

He received B.Des., M.Des. and D.Eng. degrees from Kyushu Institute of Design in 1995, 1997 and 2005, respectively. He is currently a Professor in University of Nagasaki. His research interests include non-photorealistic rendering and disaster prevention.

Human Motion Recognition from Multiple Directions and Its Gait Cycles Analysis

Miki Ooba

Graduate School of Engineering, Kyushu Institute of Technology, 1-1 Sensui, Tobata-ku, Kitakyushu, 804-8550, Japan

Yui Tanjo

Faculty of Engineering, Kyushu Institute of Technology, 1-1 Sensui, Tobata-ku, Kitakyushu, 804-8550, Japan
Email: ooba.miki558@mail.kyutech.jp, tanjo@cntl.kyutech.ac.jp

Abstract

It is crucial for individuals to keep walking and stay healthy to prevent receiving nursing care. This paper proposes a method of recognizing walk motions and analyzing the gait cycle of a human focusing on his/her posture. We use 43 structural features defined from human joint coordinates obtained using OpenPose and 18 figural features from human domain images and their difference images. The feature vector containing these 61 features is used for the recognition of walk motion by Random Forest. In the experiment, we applied the method to recognizing six types of motions and analyzed the walk gait cycles of five persons, and obtained satisfactory results.

Keywords: OpenPose, images of human area, human walk motion, structural feature, figural feature, gait cycle

1. Introduction

According to the Ministry of Health, Labour and Welfare's 'Physical Activity Standards for Health Promotion' [1], increasing the amount of daily physical activity can reduce the risk of deterioration of life functions and other problems. Therefore, it is recommended to maintain a certain intensity of physical activity (walking or equivalent movements). The purpose of this research is to recognize human daily motions and analyze the information obtained from them.

Previous studies include posture and motion recognition methods that use a chest-mounted camera by the first person viewpoint called MY VISION [2][3], gait recognition by integrating the gait silhouette input into a CNN [4], and a method that evaluates movement based on the skeletal trajectory of the whole body using depth data [5]. There is a motion recognition method [6] using an extension of Motion History Image (MHI) called Triplet Motion Representation Images, with Histograms of Oriented Optical Flow. However, attaching a measurement device to a person may cause unnatural movements. It is necessary to consider an actual living environment.

In this paper, we propose a method of recognizing human posture and gait motion based on human joint positions and image features, as well as a method of analyzing gait motion based on the gait cycle, using a random forest classifier [7] based on a 61-dimensional feature vector.

Fig. 1. Figural and structural features. (a) Human domain image, (b) joint coordinates

2. Foreground Extraction

To recognize a motion from the shape features of a human domain image, the human domain is extracted from the image as the foreground. In this paper, the sequential background estimation method based on the Gaussian mixture model [8][9] is used to extract a human domain corresponding to background change. Then, the noise on the extracted human domain is removed by expansion and contraction processing and it is trimmed to a rectangle of a specified scale. Figural features are extracted from the human domain using the binary and frame difference images shown in Fig. 1(a). After having the foreground, figural features are extracted by the step in 3.1

3. Feature Extraction

We use figural features and structural features to recognize human motion and its gait cycles.

3.1 Figural Features

Our aim is to extract effective features from human posture images. Eight kinds of figural features used in the proposed method can be categorized as follows.

1) Image aspect ratio: The ratio of the height to the width of the human domain image.

2) Percentage of a white pixel area in a rectangle: The percentage of a white pixel area in the human domain image.

3) Hand swing (upper body width): The difference between the rightmost and the leftmost points at 60% of the height of the upper body of the human domain image.

4) Shoulder height: The *y* coordinate at which the number of white pixels per line exceeds a threshold when the image is scanned from the upper left to the lower right. The threshold was experimentally set at 40 pixels to distinguish between a stretch out the back posture and others.

5) Length of the contour line at the bottom of the human region image: The length of the contour line at the bottom 40% of the height of the person area image.

6) Body asymmetry: The center of gravity of the lower 40% portion of the person area, and the percentage of the area to the left of the center of gravity in the white pixel area.

7) Position of the center of gravity: The distance of the center of gravity of the head from the centerline of the image.

8) Number of white pixels of the difference image: The number of white pixels in the 25% height area from the bottom of the difference image.

Since features 2), 3) and 5) take periodic values, their minimum and maximum values are also extracted and used as features. In this way, 18-dimensional figural features are obtained from the image of the human domain.

3.2 Structural Features

We extract the following six structural features (43 dimensions) using human joint coordinates provided from OpenPose [10].

1) Knee, ankle, and elbow angles: The knee angle θ_{knee}[deg] is calculated using the inner product of the hip and knee vectors as follows;

$$\boldsymbol{a} = (x_{hip} - x_{knee}, y_{hip} - y_{knee}) \tag{1}$$

$$\boldsymbol{b} = (x_{ankle} - x_{knee}, y_{ankle} - y_{knee}) \tag{2}$$

$$\theta_{knee} = \frac{180}{\pi} \cos^{-1} \frac{\boldsymbol{a} \cdot \boldsymbol{b}}{\|\boldsymbol{a}\|\|\boldsymbol{b}\|} \tag{3}$$

where (x_{hip}, y_{hip}), (x_{knee}, y_{knee}), (x_{ankle}, y_{ankle}) are the coordinates of the hip, knee and ankle, respectively. Using the same method, the ankle angles θ_{ankle} (the inner product of toe and knee)and elbow angle θ_{elbow} (the inner product of wrist and shoulder) are also obtained.

2) Ankle-toe, Hip-ankle angle, and Body tilt: The angle between the ankle and toe relative to the horizontal line is defined as the ankle-toe angle θ_{instep} [deg] and is given by;

$$\theta_{instep} = \frac{180}{\pi} \tan^{-1} \frac{y_{ankle} - y_{toe}}{x_{ankle} - x_{toe}} \tag{4}$$

where (x_{toe}, y_{toe}) is the coordinates of the toe as shown in Fig. 1(b). Similarly, we determine the hip-ankle angle $\theta_{hip-ankle}$ (The angle between the vertical line of the hip-ankle and the horizontal line of the left and the right hip) and the body inclination $\theta_{bodytilt}$ (The angle between the neck and the vertical line of the waist center).

3) Steps, Hip-ankle distance, and Wrist-shoulder distance

(a) Scale transformation: Using the length of a person's *thighs* [cm], the parameter *scale* [cm/pixel], which indicates how many centimeters one pixel corresponds to in the current frame, is calculated by the following;

$$scale = \frac{thighs}{\sqrt{(x_{hip} - x_{knee})^2 + (y_{hip} - y_{knee})^2}} \tag{5}$$

(b) Step width, Hip-ankle distance, Wrist-shoulder distance: The step width is calculated using the Euclidean distance between the left heel and the right heel as follows;

$$\begin{aligned} step\_width &= scale \\ &\times \sqrt{(x_{Rheel} - x_{Lheel})^2 + (y_{Rheel} - y_{Lheel})^2} \end{aligned} \tag{6a}$$

where (x_{Rheel}, y_{Rheel}) and (x_{Lheel}, y_{Lheel}) are the coordinates of the right and the left heel, respectively. Concurrently, the Euclidean distance between hip-ankle and wrist-shoulder is calculated as follows;

$$*\_\# = scale \times \sqrt{(x_* - x_\#)^2 + (y_* - y_\#)^2} \tag{6b}$$

where * indicates hip (or wrist), and # indicates ankle (or shoulder). (x_*, y_*) and $(x_\#, y_\#)$ are the coordinates of hip and ankle or wrist and shoulder, respectively.

4) Walk speed: The feature $walk\_speed$ [cm/s] indicates how long the coordinates of the waist center (x_{c0}, y_{c0}) in the current frame have moved from (x_{cn}, y_{cn}) in the previous n frames, and is given by the following equation;

$$\begin{aligned} walk\_speed &= scale \\ &\times \frac{\sqrt{(x_{c0} - x_{cn})^2 + (y_{c0} - y_{cn})^2} \times frame\_rate}{n} \end{aligned} \quad (7)$$

where (x_{c0}, y_{c0}), (x_{cn}, y_{cn}) are the coordinates of the center of the waist between the current frame and previous n frames, whereas the *frame_rate* is a unit that indicates the number of the images makeup in one second in a video.

5) Direction of motion: The direction of motion in the horizontal (*x*-axis) and in the vertical (y-axis) is obtained by vectorizing the displacement of the coordinates of the waist center during n frames and multiplying it by *scale*.

$$\begin{aligned} x &= (x_{c0} - x_{cn}) \times scale \\ y &= (y_{c0} - y_{cn}) \times scale \end{aligned} \quad (8)$$

6) Difference of the foot joint heights and wrist height: The height of the knee, and heel $height_*$ [cm] is calculated by the difference between the heights of the right and the left $height_{-*}$, as follows;

$$height_{-*} = |(y_{R*} - y_{L*})| \times scale \quad (9)$$

where * indicates knee, or heel, (x_{R*}, y_{R*}) and (x_{L*}, y_{L*}) are the coordinates of the right and the left knee, or heel, respectively. Moreover, the height from floor to toe and the height from floor to wrist are calculated in the same way.

Features such as knee and ankle angles are calculated with the left and the right foot, respectively. The feature values 1), 3), and 6) are then repeated for similar values in each gait cycle (e.g., in the walking motion, a person repeatedly bends and extends his/her knees). Therefore, in addition to the features 1) to 6) above in the current frame, the minimum and maximum values of these features in the past n frames are included as features. This results in 43 dimensional features.

Finally, from the figural and structural features, we represent a human posture using a 18+43=61-dimensional feature vector.

4. Learning and Identification

The proposed method uses Random Forest [7] as a discriminator.

Table 1. The number of frames in the video.

Motion	The number of frames				
	A	B	C	D	E
Normal	2351	2101	1878	2124	2192
Forward leaning	3577	2830	2214	2290	2512
Fall	400	407	405	309	278
Help	394	327	511	375	565
Sit	1322	1039	1336	877	849
Sleep	1141	712	991	649	860
Total	9185	7416	7335	6624	7256

(a) walk   (b) fall and sleep
Fig. 2. Direction of the motion

5. Experiment

5.1. Experimental Method

One fixed camera is set up at a height of 90 cm from an indoor floor to simulate a real-life situation. A person appears in the video, and the whole body is taken a video. The experiment was conducted on five subjects (22-23 years old). They are referred to as A, B, C, D, and E. Table 1 shows the number of frames contained in each video.

Six types of motions: Normal walking, forward leaning walk, falling, asking for help (beckoning to the camera), sitting, and sleeping are captured on video. Fig. 2 shows the direction.

Gait cycle [11]: Gait cycle (seven periods) recognition is performed for normal walking, i.e. when walking from the right to the left, the target of the gait cycle is left foot, whereas when walking from the left to the right, the target of the gait cycle is right foot. So we simply call it left-right foot. We manually select seven periods of the normal walk and put the selected periods into seven classes for training. The seven classes (periods) are Loading Response (LR), Mid Stance (MSt), Terminal Stance (TSt), Pre swing (PSw), Initial Swing (ISw), Mid Swing (MSw), Terminal Swing (TSw). In the gait cycle, 58 of the 61-dimensional features mentioned in section 3 are used for training. However, posture Initial Contact at the start of the cycle is omitted.

5.2. Method of Evaluation

Leave-one-out cross-validation was applied to the experimental data. The percentage of correct frames to all frames is used for evaluation. In the gait cycle analysis, left-right balance is evaluated by determining the left-right difference in the percentage of each period to the full video.

Table 2. Result of motion recognition

Motion	Precision [%]					
	A	B	C	D	E	**Ave**
Normal walk	97.0	99.4	93.8	99.6	98.8	97.8
Forward leaning	91.5	98.4	89.6	98.9	95.6	94.7
Fall	81.3	75.2	87.7	87.7	78.4	82.0
Help	79.7	82.0	95.1	99.2	99.3	92.1
Sit	100	100	92.5	100	100	98.2
Sleep	92.3	96.1	99.0	93.1	90.6	94.2

Table 3. Result of gait cycle recognition

Phases	Periods (class)	Precision [%]					
		A	B	C	D	E	**Ave**
Stance phase	LR	45.0	61.4	63.9	76.8	67.4	62.4
	MSt	88.7	90.7	90.8	94.5	85.3	89.9
	TSt	97.4	93.6	79.0	98.0	81.2	89.6
	PSw	55.8	77.6	77.3	68.5	67.8	67.8
Swing phase	ISw	87.9	90.4	86.8	87.4	75.2	85.4
	MSw	84.7	83.0	81.6	80.4	82.5	82.6
	TSw	91.8	81.0	66.1	90.9	78.5	81.9

6. Results and Discussion

6.1. Results

Tables 2 and 3 show the motion and gait cycle recognition results, respectively. Note that, help and sit motion are recognized without considering the direction. Fig. 3 shows some of the images of the recognition results. The upper left of the image shows the motion and gait cycle recognition results. Table 4 shows the different ratios of each period of the gait cycle of walking from the right to the left (the gait cycle of left foot) and the left to the right (the gait cycle of right foot) of each subject.

6.2. Discussion

Motion Recognition: The overall average accuracy was 95.2%. As shown in Table 2, Fall was less accurate than other motions, and the variation among subjects was greater than others. In fact, there are individual differences in the way the subject falls and shifts diagonally. Collecting more training data is needed.

Gait cycle: The overall average accuracy was 82.4%. In many cases, the gait cycle class was recognized as the class before or after the class. This is because walking is a continuous motion and the before and after postures are similar. From Table 4, it is expected that subject A's Mst is shorter than that of the left, indicating a slight decrease in the muscle strength of the right leg. Since analysis of a subject gait cycle is dependent on the recognition rate, the recognition accuracy in each class needs to be improved.

7. Conclusion

In this paper, we proposed a motion recognition and gait cycle analysis method using 61 features representing human posture and Random Forest as a discriminator. It was applied to the recognition of 6 motions and the analysis of 7 periods of gait cycles. Further study is needed to obtain more accurate values, such as step width,

Fig. 3. Normal walk and the result of recognition

Table 4. The difference of left-right foot.

Phases	Periods (class)	left-right foot differences [%]				
		A	B	C	D	E
Stance phase	LR	0.3	-5.4	-2.4	2.0	2.5
	MSt	4.4	-0.3	0.3	0.8	-1.3
	TSt	-1.9	1.4	0.7	0.9	0.8
	PSw	-2.1	2.3	1.0	-1.1	1.1
Swing phase	ISw	0.6	-4.3	1.0	-3.7	0.4
	MSw	-0.4	1.4	0.1	2.7	0.7
	TSw	-0.8	4.8	-0.7	-1.7	-4.2

for human health, and to identify the cases where multiple movements and directions are mixed. If we can provide accurate and appropriate advices to individuals, we will be able to offer better health services.

References

1. Cabinet Office, "Physical Activity Standards for Health Promotion 2013", 2013.
2. J. K. Tan, T. Kurosaki, "Estimation of self-posture of a pedestrian using MY VISION based on depth and motion network", Journal of Robotics, Networking and Artificial Life, Vol.7, No.3, pp.152-155, 2020.
3. Z. Liu, J.K. Tan, "Analysis of human walking posture using a wearable camera", International Journal of Innovative Computing, Information and Control, Vol.19, No.3, pp.805-819, 2023.
4. H. Chao, K. Wang, Y. He, J. Zhang, "GaitSet: Cross-view gait recognition through utilizing gait as a deep set", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.44, No.7, pp.3467-3478, 2022.
5. B. B. Amor, A. Srivastava, P. Turaga, G. Coleman, "A framework for interpretable full-body kinematic description using geometric and functional analysis", IEEE Transactions on Biomedical Engineering, Vol.67, No.6, pp.1761-1774, 2022.
6. J. Cao, Y. Yamashita, J. K. Tan, "Human motion recognition using TMRIs with extended HOOF", Journal of Robotics, Networking and Artificial Life, Vol. 7, No. 4, pp. 231-235, 2021.
7. L. Breman, "Random forests", Machine Learning, Vol. 45, No. 1, pp. 5-32, 2001.
8. C. Stauffer, W. E. L. Grimson: "Adaptive background mixture models for real-time tracking", Proceedings of Conference on Computer Vision and Pattern Recognition, Vol. 2, pp. 246-252, 1999.
9. A. Shimada, D. Arita, R. Taniguchi, "Increment and decrement of Gaussians in adaptive mixture-of-Gaussian background models", MIRU2006, pp. 741-751, 2006.
10. Z. Cao, T. Simon, S. Wei, Y. Sheikh, "Realtime multi-person 2D pose estimation using part affinity fields.", Proc.

2017 IEEE Conference on Computer Vision and Pattern Recognition, No. 121, pp. 1302-1310, 2017.

11. S. Demura, "Fall prevention for community seniors from basic theory of falls to intervention practice", kyorin-shoin, pp.114-116, 2012. (in Japanese)

Authors Introduction

Ms. Miki Ooba

She received B.E. degree in Intelligent Control Engineering Course, Faculty of Engineering, Kyushu Institute of Technology. Her research interests include motion analysis of elderly people in order to improve their QOL.

Dr. Yui Tanjo

Dr. Tanjo is currently a professor with the Department of Mechanical and Control Engineering, Kyushu Institute of Technology. Her current research interests include ego-motion analysis by MY VISION, three-dimensional shape/motion recovery, human detection, and its motion analysis from video. She was awarded SICE Kyushu Branch Young Author's Award in 1999, the AROB Young Author's Award in 2004, the Young Author's Award from IPSJ of Kyushu Branch in 2004, and the BMFSA Best Paper Award in 2008, 2010, 2013 and 2015. She is a member of IEEE, The Information Processing Society, The Institute of Electronics, Information and Communication Engineers of Japan.

A Method of Improving the QOL of the People with Visual Impairment by MY VISION

Shun Kitazumi

Graduate School of Engineering, Kyushu Institute of Technology, 1-1 Sensuicho, Tobata-ku, Kitakyushu, 804-8550, Japan

Yui Tanjo

Faculty of Engineering of Engineering, Kyushu Institute of Technology, 1-1 Sensuicho, Tobata-ku, Kitakyushu, 804-8550, Japan
Email: kitazumi.shun215@mail.kyutech.jp, tanjo@cntl.kyutech.ac.jp

Abstract

Visually impaired people face several difficulties in indoor activities, such as spending excessive time in locating objects. This paper proposes a method for assisting object acquisition by detecting desired objects and guiding users to them. The method requests a user to express the object he/she wants to acquire verbally and utilizes speech recognition to detect the specified object. Subsequently, the system guides in voice the user's hand to the location of the desired object. The performance of the method is experimentally shown. The method contributes to enhancing the comfort of indoor activities of visually impaired and, in this way, improves their quality of life.

Keywords: MY VISION, Visually impaired, QOL, Object acquisition, Network diagram

1. Introduction

Many people around the world suffer from visually impaired vision, a condition that interferes with their daily lives due to persistent vision loss. According to a report [1] published by the World Health Organization (WHO) in 2019, the number of people with visual difficulties worldwide is estimated to be at least 2.2 billion, and there are concerns that the number of people affected will increase further due to population growth and aging.

In response to this worldwide social problem of visibility difficulties, a great deal of research and development has been conducted to support the daily lives of people with visibility difficulties. For example, there has been the development of a pedestrian crossing navigation system using smart glasses to enable people with vision difficulties to safely walk across pedestrian crossings alone [2][3][4][5], and the development of a system to assist people in getting on and off public transportation [6][7].

As in the above studies, much of the research on assisting people with vision difficulties tends to focus on supporting outdoor activities. However, there are many problems that occur during indoor activities, such as spending excessive time for acquiring an object when looking for it at home. Therefore, the purpose of this study is to improve the quality of life (QOL) of people with visual acuity difficulties by focusing on a proposal for an object acquisition support method, which plays a part in supporting indoor activities for people with visual acuity difficulties.

Currently, an application has been released that uses a smartphone to capture a registered object with a camera and read it aloud to the user[8]. This application is very useful for identifying objects that are similar in shape, but since it assumes that the object's location can be accurately identified and captured by the camera, it does not work unless the visually impaired person knows the object's location. In addition, as a study that addresses the guidance of object acquisition paths, a system has been proposed by J.P. Docto *et al.* [9], in which a smart glove is worn by a visually impaired person and guides him or her to the object utilizing the built-in camera and sensors. This system does not require text input or touch operation and is easy for visually impaired people to handle. However, in practical use, there are issues related to practicality, such as the time and effort required to put on and take off the gloves, the weight of the gloves when they are worn, and the resistance caused by the wiring.

In this study, we propose a system that provides voice guidance on a route to the object requested based on MY VISION (a Magic eYe of a Visually Impaired for Safety and Independent actiON), thereby reducing the burden of visually impaired people and making the system more intuitive and practical. Specifically, the system can provide real-time guidance using a notebook PC and a webcam, regardless of location, and enables it by voice without text input or button operation. This is expected to solve the practical limitations of smartphone applications and smart gloves that have been pointed out in previous research. Therefore, the system proposed in this study can

Shun Kitazumi, Yui Tanjo

be an important step toward improving the quality of life of people with vision difficulties during indoor activities by supporting smooth acquisition of the target object when they are looking for something at home.

2. Methodology

2.1. Voice recognition

In this study, when a user has an object that he/she wants to acquire, the user is asked to tell the object loud (speech transmission) to the proposed system, and the target object's name is recognized by analyzing the speech. By asking the user to say the trigger word before the object name, the noun that follows the trigger word is known as the target object's name. In the system, "MY VISION" is used as the trigger word. For example, if the user wants to acquire a cell phone, he/she just say, "MY VISION, find a cell phone". The system then guides the user's hand to the location of the desired object using voice guidance.

2.2. *Object detection*

In the proposed system, we use YOLOv7 (You Only Look Once version 7) [10] to detect a target object in real time. Note that in the system, objects that are present on the desk and easily portable (e.g., remote controls, cell phones, cups) are considered as a class of objects to be detected. In addition, by combining YOLOv7 and DeepSORT (Deep Learning Simple Online and Realtime Tracking) [11], object features are extracted on each frame, and each object is assigned a unique ID for tracking. Fig. 1 shows the object detection results.

2.3. *Route guidance to a target object*

If a target object is detected, the direction of the object location (referred to as a route hereafter) is guided by voice guidance according to the clock position. We use MediaPipe Hands [12] to estimate the user's hand area and align the user's hand with the target object. When the target object is in the 9 o'clock direction (left direction), the system announces "move to 9 o'clock direction" with voice.

The proposed method not only focuses on the hand and the target object, but also employs proposed a network diagram to represent the positional relationship among the surrounding objects based on the object detection results. In the network diagram, nodes are defined as the center coordinates of the object, and the edges are the distances in pixel between objects. The network diagram is introduced because the system can find the target object position smoothly, and also can find the position of the target object more precisely, when the target is very close to the other objects.

Fig. 1 Object detection results. (The target object is drawn in a blue bounding box and the user's hand is in a green box.)

Moreover, if an object to be detected is not detected in the current frame, then the system estimates the undetected position by referring to the previous network diagram.

The first condition for selecting the network diagram to be referred to is when the number of detected objects is three or more and the number of detected objects is the largest (Eq. (1)), and the second condition is when the network diagram consists of the same type of objects for a certain period (Eq. (2)).

$$\begin{cases} obj\_cnt_{F(t)} \geq 3 \\ obj\_cnt_{F(t)} \geq max\_obj\_cnt \end{cases} \quad (1)$$

Here, $F(t)$, $obj\_cnt_{F(t)}$ and max_object_cnt are current frame, the number of objects detected in the current frame and the maximum number of objects detected in each frame, respectively.

$$S = \begin{cases} S + 1 & \text{if } \forall E(F(t)) = \forall E(F(t-1)) \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

$$S \geq Th \Longrightarrow \text{SetReferenceModel}(\forall E(F(t)))$$

Here, S, $E(F(t))$, and Th are the variable for state continuity, node in $F(t)$ and threshold, respectively.

The object location estimation method first calculates the amount of object movement common to the current frame and the reference network diagram, respectively, and derives the average amount of movement (Eq. (3)). The position of the undetected object is then estimated based on the coordinates of the model and the average displacement (Eq. (4)). The meanings of the letters in the equations are shown in Table 2.

$$\begin{cases} \bar{x} = \dfrac{1}{n} \sum_{i=1}^{n} (x_{t,i} - x_{model,i}) \\ \bar{y} = \dfrac{1}{n} \sum_{i=1}^{n} (y_{t,i} - y_{model,i}) \\ (i = 1, \cdots, n) \end{cases} \quad (3)$$

$$\begin{cases} x_{est,k} = x_{model,k} + \bar{x} \\ y_{est,k} = y_{model,k} + \bar{y} \end{cases} \quad (4)$$

Table 2. The meanings of Eq. (3) and Eq. (4)

$\bar{x}, \bar{y}$	Average displacement in the x or y-axis direction
n	Number of common nodes between current frame and reference model
$x_{t,i}, y_{t,i}$	x or y coordinates of the common node in the current frame
$x_{model,i}, y_{model,i}$	x or y coordinate of the common node in the model
$x_{est,k}, y_{est,k}$	x or y coordinate estimates for missing nodes
$x_{model,k}, y_{model,k}$	x or y coordinates of the node in the model corresponding to $x_{est,k}$ or $y_{est,k}$

The first condition for starting object location estimation is when a reference network diagram exists, and the second condition is when the number of detected objects is two or more and there are two or more objects in common with the reference network diagram. Fig. 2 shows the results of object location estimation using the above condition.

In the proposed system, a monocular RGB camera is used to align the hand with the target object without using a depth camera. The condition for determining object acquisition is achieved by employing the bounding box positional relationship between the target object and the hand region, and by judging the change in hand orientation that occurs when the hand gestures come closer to the camera after grasping the object. By adding the latter condition, it is expected to prevent the error judgment of acquisition as successful simply because the bounding boxes of the two objects overlap, even when the depths are different, and the objects are not touched. Furthermore, the system also incorporates a process that checks whether or not the object closest to the hand matches the target object by referring to the network diagram and announces when the subject has acquired the wrong object. When the above three conditions are satisfied, the target object is judged to have been acquired, and the acquisition is communicated by voice as completed. Fig. 3 shows an example that the object acquisition decision was fulfilled.

(a) Network diagram of the current frame

(b) Network diagram of the current frame

(c) Network diagram of the reference model

(d) Network diagram reflecting estimation results

(e) Before reflecting location estimation

(f) After reflecting location estimation

Fig. 2 Result of object location estimation
(When the target object is estimated, it is surrounded by a red bounding box.)

Shun Kitazumi, Yui Tanjo

Fig. 3 Image on completion of the object acquisition that satisfies the three conditions

Fig. 4 Arrangement where other objects are present in the guidance path to the target object (target object: remote controller)

3. Experimental Result

Five objects were selected as candidate objects for acquisition: a book, a cell phone, a cup, a plastic bottle, and a remote controller. An experiment was conducted to select an acquisition target from these objects and to make a robot guide a blindfolded user along a path to the target object. As shown in Fig. 4, we placed other objects on the guided path to the target object to verify if it is possible to approach the target object again after announcing that the wrong object has been grabbed. The target object position was changed in the directions of 9, 10, 12, 2, and 3 o'clock with respect to the hand position to ensure that there was no dependence on the object placement. The evaluation index was *Accuracy* which represents the percentage of correct answers acquired. A total of 25 acquisition experiments were conducted with 5 choices among 5 objects times 5 object positions, resulting in an Accuracy of 92%.

4. Conclusion

In this paper, we proposed a system that provides voice guidance on the path to object acquisition based on MY VISION and contributes to improving the quality of life during indoor activities of the people with vision difficulties. In addition, by utilizing the information obtained from Mediapipe Hands and network diagrams, we developed a practical and intuitive system that provides all information and guidance by voice, as well as countermeasures in case that object tracking is interrupted.

References

1. WHO: ""World report on vision", https://www.who.int/publications/i/item/9789241516570, 2019, (Accessed 2023-12-04).
2. H. Son, D. Krishnagiri, V. Jeganathan, J. Weiland: "Crosswalk guidance system for the blind", 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), pp.3327-3330, 2020.
3. J. K. Tan, T. Ishimine, S. Arimasu: "Walk Environment Analysis Using MY VISION: Toward a Navigation System Providing Visual Assistance", International Journal of Innovative Computing, Information and Control, Vol. 15, No.3, pp.861-871, 2019.
4. J. K. Tan, K. Kitagawa: "A Method of Navigating a Visually Impaired Person Using MY VISION", Journal of Robotics, Networking and Artificial Life, Vol.9, No. 1, pp. 25-30, 2022.
5. T. Kumano, J. K. Tan: "Traffic signs and signals detection employing the my vision system for a visually impaired person", ICIC Express Letters, Part B: Applications pp. 385-391, 2015.
6. S. M. Cruz, L. A. M. Hernández, et al.: "An Outdoor Navigation Assistance System for Visually Impaired People in Public Transportation", IEEE Access, vol.9, pp.130767-130777, 2021.
7. J. K. Tan, Y. Hamasaki, Y. Zhou, I. Kazuma: "A Method of Identifying a Public Bus Route Number Employing MY VISION", Journal of Robotics, Networking and Artificial Life Vol.8, No.3, pp.224-228, 2021.
8. INTEC Inc. : https://www.intec.co.jp/news/2020/0318_1.html, (Accessed 2023-12-05).
9. J. P. Docto, A. I. Labininay, J. F. Villaverde: "Third eye hand glove object detection for visually impaired using You Only Look Once (YOLO)v4-tiny algorithm", 2022 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET), pp.1-6, 2022.
10. Y. Wang, A. Bochkovskiy, H. Liao: ""YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for realtime object detectors", arXiv preprint arXiv:2207.02696, 2022.
11. N. Wojke, A. Bewley, D. Paulus, "Simple online and realtime tracking with a deep association metric", 2017 IEEE International Conference on Image Processing (ICIP), pp. 3645–3649, 2017.
12. MediaPipe, https://developers.google.com/mediapipe, (Accessed 2023-12-05).

Authors Introduction

Mr. Shun Kitazumi

He received his Bachelor's degree in Engineering in 2022 from the Faculty of Engineering, Kyushu Institute of technology in Japan. He is currently a master student in Kyushu Institute of Technology, Japan

Dr. Yui Tanjo

Dr. Tanjo is currently a professor with the Department of Mechanical and Control Engineering, Kyushu Institute of Technology. Her current research interests include ego-motion analysis by MY VISION, three-dimensional shape/motion recovery, human detection, and its motion analysis from video. She was awarded SICE Kyushu Branch Young Author's Award in 1999, the AROB Young Author's Award in 2004, the Young Author's Award from IPSJ of Kyushu Branch in 2004, and the BMFSA Best Paper Award in 2008, 2010, 2013 and 2015. She is a member of IEEE, The Information Processing Society, The Institute of Electronics, Information and Communication Engineers of Japan.

Human Behavior Segmentation and Recognition Using a Single-camera

Jing Cao

Graduate School of Engineering, Kyushu Institute of Technology, 1-1 Sensui, Tobata-ku, Kitakyushu, 804-8550, Japan

Yui Tanjo

Faculty of Engineering, Kyushu Institute of Technology, 1-1 Sensui, Tobata-ku, Kitakyushu, 804-8550, Japan
Email: cao.jing644@mail.kyutech.jp, tanjo@cntl.kyutech.ac.jp

Abstract

In recent years, elderly people living alone account for a large proportion of the elderly population, and the issue of safety has also been a matter of great concern for the public. Considering the importance of monitoring the behavior and activities of the elderly and detecting abnormal movements, this paper proposes a method that can segment human behavior into each action and identify the action from the videos taken by a single camera. It uses features that can represent the shape of the human area in the depth direction, as well as the features such as motion direction and speed. The performance and effectiveness of the method are verified by experiments.

Keywords: Behavior segmentation, Motion recognition, Optical flows, TMRIs, Ex-HOOF, MHI

1. Introduction

Nowadays, the world's population is aging. The number and proportion of older people in the population is growing in every country in the world. Population aging will become one of the most significant social changes in the 21st century. By 2030, 1 in 6 people in the world will be 60 years or older (16%), and the share of people aged 60 and above will increase from 1 billion in 2020 to 1.4 billion. By 2050, the number of people aged 80 and over is expected to triple between 2020 and 2050, to 426 million [1]. Especially in Japan, 30% of the population is over 60 years old. Some social problems, such as solitary death which more than 50% of people worry about, are also intensifying [2]. Moreover, 67.1% of the respondents felt that the domestic security situation in Japan has deteriorated in the past 10 years [3]. Considering these problems, it is necessary to develop a system that can detect the behavior of the elderly and detect abnormal behaviors such as crime and theft, so as to realize a safe and secure society.

In related research on behavior recognition, X. Yang et al. [4] proposed a behavior segmentation and recognition method based on CSI, but it is not aimed at the segmentation of continuous human behavior. W. Xing et al. [5] proposed a behavior segmentation method based on posture histograms and adjusted sliding windows. This method requires learning the features of various postures. In related research on action recognition using computer vision, the conventional methods include the Flow Vectors method [6] and the MHI (Motion History Image) method [7]. However, these methods using a

single camera are only suitable for the motion on the plane perpendicular to the optical axis of a camera, and the motion in the direction of the camera optical axis (toward or away from the camera) is not dealt with. Therefore, an extended 3D-MHI [8] and a reverse MHI method [9] have been proposed. However, [8] has the problem of high computational cost of creating 3D images, and, with [9], the recognition rate needs to be improved.

In this study, we propose a new method to describe the motion in the depth direction, making it possible to realize the segmentation and recognition of behaviors containing these motions. We segment human behavior by extracting features from the human area and then select key frames of the segmented actions. The key frames can be described through TMRIs (Triplet Motion Representation Images) [10]. The shape features of TMRIs, the features of optical flow and the changes of motion are also extracted to analyze and identify the motion of the frame. Finally, the recognition results of these frames are counted, and the final result of behavior recognition is obtained.

2. Methodology
2.1. *Behavior segmentation*

When human behavior consists of several types of motions, it is necessary to separate the behavior into respective motions to analyze the behavior. In this paper, we propose an automatic behavior segmentation method that does not require prior learning or training. The feature points are set at equal distances on the contour

line of the human area. Then feature points are tracked through the LK method [11] to obtain the optical flow, and RANSAC [12] is applied to find the true value by removing the influence of outliers. Changes in the direction of human movement can be distinguished from changes in the trajectory of the center of gravity coordinates and the average value of optical flow. Therefore, we calculate the point where a sudden change in trajectory occurs in the *x*-axis direction of the center of gravity coordinates and the boundary point between plus and minus of the average value of optical flow separately.

In Equation (1) and (2), if the change of the current frame exceeds a certain threshold, division is performed from the current frame and the number of the previous frame is used as the division point.

$$num_{tmp} = \begin{cases} num_{tmp} + 1 & if\ th1 < \frac{X\_dif_{sum}}{num_{tmp}} < th2 \\ 1 & otherwise \end{cases} \quad (1)$$

$$X\_dif_{sum} = \frac{\sum_{n=new}^{new-num_{tmp}}(x_n - x_{n-1})}{num_{tmp}} \quad (2)$$

where, *th1* and *th2* are thresholds. num_{tmp} is a parameter that determines whether or not to divide from the current frame. X_dif_{sum} is the average rate of change of the center of gravity's *x*-axis coordinate, and *new* is the number of the current frames.

When the direction of human motion changes, it can also respond to the average of optical flow. The intersection of the average optical flow m_{ave} and the *x*-axis is determined to be the division point. m_{ave} is defined by the following formula.

$$m_{ave} = \frac{\sum_{k=1}^{N} m_k}{N} \quad (3)$$

where, m_k represents the value of the *k*-th optical flow, and *N* denotes the total number of optical flows.

Then the division points are determined by matching the points of the two features after performing data smoothing within the threshold. The video can be divided using the obtained division points. After that, key frames are selected in each divided video, as shown in Fig. 1. In the figure, the orange points are the division points detected by features, which separate the behavior into several motions. The green points are the points that divide the motion into four equal parts, are also extracted as key frames.

In the end, by identifying key frames and determining the classification of the actions they represent through majority voting, the analysis of behavior composition is completed.

Fig.1 The definition of the ground truth.

2.2. *Motion recognition*

To capture the motion in the depth dimension, we employ TMRIs to represent actions. We identify and classify actions by leveraging their shape features, extracting directional velocities from the movements, and observing changes in the human region.

TMRIs is an extension of the conventional MHI. It is designed to express human motion, including motion along the camera optical axis, using three types of motion history images: *newness*, *density*, and *depth*. Among them, *newness* represents the MHI, *density* illustrates the appearance frequency of the foreground in the past τ frames, and *depth* indicates the depth information obtained from the FoE detection results. The shape features of TMRIs, is defined as V^{TMRIs}. Ex-HOOF (Extended Histogram of Oriented Optical Flow) [10] is used to extract the speed and directional information of motions that are not readily visible in TMRIs features. The feature of Ex-HOOF is denoted as $V^{Ex-HOOF}$.

Additionally, to represent more complex and detailed motion features, we incorporate changes in the centroid and area (AC) of the human region. The feature can be defined as $V^{AC} = (F_A, F_C)$. Among them, the area feature $F_A = (Area_\tau^{ave}, Area_\tau^{sd})$, where $Area_\tau^{ave}$ is the average value, and $Area_\tau^{sd}$ denotes the standard deviation of the change of the foreground area. The formulas are as follows;

$$Area_\tau^{ave} = \frac{1}{\tau}\sum_{i=0}^{\tau} Area_{comp}^{t-i} \quad (4)$$

$$Area_\tau^{sd} = \sqrt{\frac{1}{\tau}\sum_{i=0}^{\tau}(Area_{comp}^{t-i} - Area_\tau^{ave})^2} \quad (5)$$

Here, $Area_{comp}^t$ is the changes of areas between the frames, given by

$$Area_{comp}^t = \frac{(Area_t - Area_{t-p})}{Area_t} \quad (6)$$

a. Change in the trajectory of the center of gravity coordinate in the *x*-axis

b. Average change in optical flow in the *x*-axis direction

c. Average change in optical flow in the *y*-axis direction

Fig.2. Features used for behavior segmentation.

Fig.3 The definition of the ground truth.

The centroid feature $F_C = (Cx_\tau^{ave}, Cy_\tau^{ave}, Cx_\tau^{sd}, Cy_\tau^{sd})$. C_τ^{ave} means the average value and C_τ^{sd} indicates the standard deviation of the change in the coordinate of the center of gravity. They are defined by

$$Cx_\tau^{ave} = \frac{1}{\tau} \sum_{i=0}^{\tau} (Cx_t - Cx_{(t-i-p)}) \qquad (7)$$

$$Cy_\tau^{ave} = \frac{1}{\tau} \sum_{i=0}^{\tau} (Cy_t - Cy_{(t-i-p)}) \qquad (8)$$

$$Cx_\tau^{sd} = \sqrt{\frac{1}{\tau} \sum_{i=0}^{\tau} ((Cx_t - Cx_{(t-i-p)}) - Cx_\tau^{ave})^2} \qquad (9)$$

$$Cy_\tau^{sd} = \sqrt{\frac{1}{\tau} \sum_{i=0}^{\tau} ((Cy_t - Cy_{(t-i-p)}) - Cy_\tau^{ave})^2} \qquad (10)$$

Finally, by integrating the above features, we get a feature vector V, as shown below.

$$V = (V^{TMRIs}, V^{Ex-HOOF}, V^{AC}) \qquad (11)$$

Action recognition is conducted using the *k*-nearest neighbors (*k*-NN) method in the proposed method.

3. Results and Discussion

3.1. *Accuracy of behavior segmentation*

In the behavior segmentation experiment, three people (labelled as 1,2,3) perform behavior A, B, C and D. The specific motion sequences are as follows: A: Walking left rear → Walking right → Walking front, B: Walking right rear → Walking left → Walking front, C: Walking rear → Walking left → Walking right front, D: Walking rear → Walking right → Walking left front. In the segmentation experiments, we verify the accuracy and recall of segmentation point selection based on the trajectory of the centroid in the *x*-axis and the average direction of optical flow. The example of the features of behavior A is shown in Fig. 2. The recall rate (*Recall*) is defined by

$$Recall = \frac{P_T}{P_{GT}} \times 100[\%] \qquad (12)$$

Here, P_{ALL} is the total number of calculated segmentation points, P_T is the number of segmentation points that were accurately segmented, and P_{GT} is the number of ground truth segmentation points.

In the experiment, a motion between two motions, such as turning around or standing, is called a transition action. If the division point is within a transition action, this point is considered a correct division point. If one behavior consists of four consecutive motions, the definition of the ground truth division points is shown in Fig. 3.

Table 1 shows the results of behavior segmentation, yielding an average recall of 91.7%.

3.2. *Recognition rate of the motions*

In the motion recognition experiment, four people performed 12 kinds of motions, and features were created for these motions. In the experiment, we use leave-one-out cross-validation to evaluate the accuracy of motion recognition. The recognition rate R is defined by

$$R = \frac{N_T}{N_{ALL}} \times 100[\%] \qquad (13)$$

Jing Cao, Yui Tanjo

Here, N_T is the total number of correctly recognized features and N_{ALL} represents the total number of features.

The recognition rates obtained using TMRIs, EX-HOOF, TMRIs + EX-HOOF, and the proposed method (TMRIs + EX-HOOF + AC) are shown in Table 2.

As shown in the table, the proposed method achieved an average recognition rate of 97.25%. This is because TMRIs represent depth information determined by FoE detection results. In addition, Ex-HOOF includes movement speed and direction information, and the detailed feature AC expresses changes in the area of the motion region and the center of gravity. Therefore, in the final proposed method, the average recognition rate for daily activities is 99.1%, and the average recognition rate for falling activities is 89.84%. Furthermore, the average recognition accuracy for motions that include movement in the depth direction is 96.60%, and the average recognition accuracy for other motions is 94.85%. The average recognition accuracy of the proposed method using TMRIs + EX-HOOF + AC is about 17.41% higher than the method using only TMRIs.

4. Conclusion

In this paper, we proposed a human behavior segmentation and recognition method, which can handle the behavior including movements in the depth direction. For each behavior, the segmentation method proposed in the paper effectively identifies division points, enabling the recognition of segmented actions. The average recall rate for behavior segmentation has reached 91.69%. Furthermore, in action recognition, the proposed method achieved an average recognition rate of 97.25%. This strongly validates the effectiveness of the proposed approach.

In the future, our focus will be on achieving a high-speed and fully automated process of behavior segmentation through recognition. Additionally, despite the current good recall rate, there is still a need to enhance the accuracy of the automatic segmentation method to reduce the subsequent workload of recognition and segmentation adjustments.

Table 1. The result of behavior segmentation

Per-son	Beha-vior	Trajectory	Optical flow	Integration	Positive number	Recall [%]
1	A	23	9	6	3	100.0
	B	33	20	9	4	100.0
	C	46	17	10	4	100.0
	D	27	17	7	4	100.0
2	A	21	11	5	3	75.0
	B	17	10	4	3	75.0
	C	28	9	5	4	100.0
	D	30	19	6	4	100.0
3	A	28	18	6	3	75.0
	B	20	16	6	4	100.0
	C	27	10	7	4	100.0
	D	28	14	6	3	75.0
Average						91.7

Table 2. The recognition rate of the motions

Motion	TMRIs	Ex-HOOF	TMRIs + Ex-HOOF	TMRIs + Ex-HOOF + AC
Walk left	87.5	44.4	93.1	99.4
Walk right	70.3	93.1	99.4	100.0
Walk front	78.1	75.3	99.4	99.7
Walk back	73.1	97.8	100.0	100.0
Walk left front	83.8	96.9	96.3	99.1
Walk right front	87.8	72.8	94.7	100.0
Walk left rear	83.4	92.2	96.6	98.8
Walk right rear	83.1	94.4	95.3	95.9
Fall left	91.9	63.8	91.9	91.9
Fall right	79.4	71.3	75.6	88.1
Fall front	61.9	44.4	80.6	89.4
Fall rear	69.4	63.1	86.3	90.0
Average	79.8	78.8	94.2	97.3

References

1. United Nations, Department of Economic and Social Affairs, Population Division, World Population Prospects 2022. (Online Edition).
2. Cabinet Office, 2022 White Paper on Aging Society (Overall version), 2022, pp.2-6. (in Japanese)
3. Cabinet Office, Outline of the 2022 Public Opinion Poll on Security, 2022, pp.24. (in Japanese)
4. X. Yang, J. Cheng, X. Tang et al., CSI-based human behavior segmentation and recognition using commodity Wi-Fi. J Wireless Com Network 2023(46), 2023.
5. W. Xing, W. Wang et al., A Novel Method for Automated Human Behavior Segmentation. Apr. 2016.
6. E.L. Andrade, R.B. Fisher, and S. Blunsden, Detection of emergency events in crowded scenes, Proceedings of IEEE International Symposium on Imaging for Crime Detection and Prevention, Hong Kong, China, 2006, pp.528-533.
7. A. Bobick, J. Davis, The recognition of human movement using temporal templates, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.23(3), 2001, pp.257-267.
8. Y. Yamashita, J.K. Tan, and S. Ishikawa, Human motion description and recognition under arbitrary motion direction, Proceedings of SICE Annual Conference 2017, Kanazawa, Japan, 2017, pp.110-115.
9. J.K. Tan, S. Okae, Y. Yamashita, and Y. Ono, A method of describing a self-occlusive motion – A reverse motion history image, International Journal of Biomedical Soft Computing and Human Sciences, vol.24(1), 2019, pp.1-7.
10. J. Cao, Y. Yamashita, J.K. Tan, Human motion recognition using TMRIs with extended HOOF, Journal of Robotics, Networking and Artificial Life, vol.7(4), 2021, pp.231-235.
11. B.D. Lucas, T. Kanade, An Iterative Image Registration Technique with an Application to Stereo Vision, Proc. of Int. Joint Conf. on Artificial Intelligence, 1981, pp.674-679.
12. M. A. Fischer, R. C. Bolles, Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography, 1981, pp.381-395.

Authors Introduction

Ms. Jing Cao

She received B.E. from Republic of China and M.E. from the Graduate School of Engineering, Kyushu Institute of Technology, Japan in 2020. She is acquiring the D.E. in the same University. Her research interest includes computer vision, machine learning and motion recognition.

Dr. Yui Tanjo

She is currently a professor with the Department of Mechanical and Control Engineering, Kyushu Institute of Technology. Her current research interests include ego-motion analysis by MY VISION, three-dimensional shape/motion recovery, human detection, and its motion analysis from video. She was awarded SICE Kyushu Branch Young Author's Award in 1999, the AROB Young Author's Award in 2004, the Young Author's Award from IPSJ of Kyushu Branch in 2004, and the BMFSA Best Paper Award in 2008, 2010, 2013 and 2015. She is a member of IEEE, The Information Processing Society, The Institute of Electronics, Information and Communication Engineers of Japan.

Online Classroom Student Engagement Analysis using Enhanced YOLOv5

Shuai Wang
*Institute of Computer Science and Digital Innovation, UCSI University, No.1, Jalan Menara Gading,
Cheras, Kuala Lumpur, Malaysia*
Abdul Samad Shibghatullah
*Institute of Computer Science and Digital Innovation, UCSI University, No.1, Jalan Menara Gading,
Cheras, Kuala Lumpur, Malaysia*
Javid Iqbal
*Institute of Computer Science and Digital Innovation, UCSI University, No.1, Jalan Menara Gading,
Cheras, Kuala Lumpur, Malaysia*
Kay Hooi Keoy
*UCSI Graduate Business School, UCSI University, No.1, Jalan Menara Gading,
Cheras, Kuala Lumpur, Malaysia*
*E-mail: 1002268166@ucsiuniversity.edu.my,abdulsamad@ucsiuniversity.edu.my,javid@ucsiuniversity.edu.my,
keoykh@ucsiuniversity.edu.my
www.ucsiuniversity.edu.my*

Abstract

The surge in online education has underscored the pressing issue of cyberbullying in virtual classrooms. This paper introduces an inventive method for early cyberbullying detection by analyzing students' engagement and emotional responses in online classrooms. The proposed *SFER-YOLOv5* model, integrating Student Facial Expression Recognition with an enhanced YOLOv5 object detection model, incorporates transformative optimizations. These include Soft NMS for Non-Maximum Suppression, the integration of a Channel Attention (CA) module within the YOLOv5 architecture, and the use of Enhanced Intersection over Union (EIOU) as the bounding box regression loss function. This approach identifies diminished engagement and emotional irregularities, offering a proactive framework for mitigating cyberbullying in online classrooms.

Keywords: Cyberbullying Mitigation, Emotional Analysis, Facial Expression Recognition, Online Classroom.

1. Introduction

With the rapid advancement of information technology and the global rise of online education, online teaching platforms have become an integral part of modern education. The sudden outbreak of the COVID-19 pandemic in 2020 has had a significant impact on our learning, work, and daily lives. Traditional face-to-face teaching and communication between students and teachers have become a major challenge. Many universities have chosen online teaching as an alternative. However, accompanying this advancement are a host of challenges, one of which is the issue of cyberbullying in the digital realm. Cyberbullying, also known as online harassment or cyber aggression, refers to the malicious use of digital technology to target others with acts of aggression, insults, threats, and more. Particularly in online classroom environments, both students and educators may find themselves vulnerable to cyberbullying, which can severely impact their emotional well-being, psychological state, and overall learning experience.

In the online classroom environment, students and teachers communicate through screens, and teachers can only judge students' listening status based on their facial expressions. Analyzing changes in students' facial expressions can help teachers better understand their listening status and make timely adjustments to the

Shuai Wang, Abdul Samad Shibghatullah, Javid Iqbal Thirupattur, Kay Hooi Keoy

teaching mode [1]. In 1971, American scholar Ekman et al. [2] conducted extensive experiments on facial expressions, categorizing them into six primary emotions: happiness, surprise, fear, sadness, disgust, and anger. Students' emotions in class can be categorized as positive, negative, or neutral. Positive emotions include happiness and surprise. When students show positive emotions, it indicates their willingness to accept the knowledge taught in class, and they are actively engaged in listening and thinking. Negative emotions include sadness, anger, fear, and disgust, which indicate that students have a dislike or lack of attentiveness towards the knowledge taught by the teacher. This can be an indication of their lack of concentration in class. Neutral emotions suggest an average level of student engagement in listening during class. Therefore, analyzing students' facial expressions in online classrooms holds great significance.

Faced with this challenge, many educational institutions and online teaching platforms are turning their attention to how to effectively detect and address cyberbullying in virtual classrooms. Facial expression recognition technology has garnered significant attention due to its potential applications in emotion analysis and affective state monitoring. Currently, there is limited research on facial expression recognition based on YOLOv5. This paper proposes a facial expression recognition method based on an improved YOLOv5 model, aiming to achieve timely detection and intervention of cyberbullying incidents in online classrooms. The original YOLOv5 algorithm achieved an overall recognition accuracy of 73.1% and 83.4% on the Fer2013 dataset and a self-constructed dataset for the three expressions involved in the study: happiness, sadness, and neutral. Through improvements in NMS, and the addition of an attention mechanism module, the improved YOLOv5 algorithm further enhances the recognition accuracy on the Fer2013 dataset and the self-constructed dataset.

2. Related work

2.1. *Online classroom cyberbullying*

Online education has become increasingly prevalent in recent years, especially due to the COVID-19 pandemic, which has posed new challenges and opportunities for educators and students. One of the major challenges is cyberbullying [3], [4], which refers to various forms of

online harassment, such as verbal abuse, threats, spreading false information, and exclusion from online discussions, that occur within virtual classroom environments. Cyberbullying can have detrimental effects on students' learning outcomes, psychological well-being, and social relationships [5]. Therefore, it is imperative to develop effective strategies and mechanisms to detect and prevent cyberbullying incidents in online classrooms. Previous research on this topic has adopted various techniques and perspectives, such as behavior analysis, natural language processing, social network analysis, machine learning, and emotion recognition, to identify and intervene in potential bullying situations that students may face in virtual classrooms. Moreover, some studies have highlighted the role of educational interventions and awareness-raising among educators and students about cyberbullying, aiming to foster a safer and more positive online learning environment [6].

2.2. *YOLOv5 and object detection*

In May 2020, UltralyticsLLC released YOLOv5 [7], a lightweight model with an image inference speed of up to 0.007 second, developed using python. It can process 140 FPS and can meet the real-time requirement for video sequences. The four network structures Yolov5s\5m\5l\5x are basically same in principle and content, controlled respectively by the parameters of width_multiple and depth_multiple in width and depth. The size and accuracy of the four versions of the model increase sequentially. In practical applications, the appropriate size of the model can be selected based on different scenarios. The YOLOv5s model has small depth and fewer parameters, making it more applicable to real-time tasks in facial expression recognition due to its faster inference speed compared to the other three versions. Since the release of YOLOv5, its versions have been updated and iterated, and this paper is based on the version 6.0 for improvement.

In summary, prior research has highlighted the urgency of addressing online classroom cyberbullying and the potential of facial expression recognition technology for emotion analysis. Additionally, the advancements in object detection methods, particularly YOLOv5, provide a solid foundation for integrating object detection into cyberbullying detection frameworks. This study builds upon these foundations to propose a novel model *SFER-YOLOv5* for detecting cyberbullying incidents in online

classrooms using an enhanced YOLOv5 model for facial expression recognition.

3. Methodology

Currently, there is limited research on facial expression recognition based on YOLOv5, and this method presents a unique opportunity for enhancing online classroom engagement and security. The original YOLOv5 algorithm achieved an overall recognition accuracy of 72% and 83.1% for the three facial expressions of happiness, sadness, and neutrality on the Fer2013 dataset and the self-built dataset. Through improvements in the non-maximum suppression (NMS) algorithm, loss function, and the addition of an attention mechanism module, the improved YOLOv5 algorithm achieved further improvement in recognition accuracy on the Fer2013 dataset and the self-built dataset.

3.1. *Data collection and processing*

As students' understanding of knowledge can be broadly categorized into three levels: mastery, confusion, and non-mastery, in order to gain a more accurate understanding of the facial expression feedback corresponding to these three levels, we conducted a survey in the form of a questionnaire to investigate the relationship between mastery levels and facial expression feedback. The survey involved higher vocational schools, and undergraduate institutions, with a total of 520 questionnaires collected (male: 287, female: 233). Seven basic expressions were selected, including anger, surprise, disgust, happiness, fear, sadness, and neutral.

This study used the Fer2013 dataset and a self-constructed dataset of student facial expressions in the classroom. The Fer2013 dataset contains some non-facial images and mislabeled images. To improve the experimental accuracy, the dataset was carefully selected, cleaned, and filtered, resulting in a total of 11,000 images. The self-constructed dataset of student facial expressions in the classroom was collected through video monitoring, and it consists of 668 images. The images were annotated by extracting frames from the video footage.

3.2. *Enhanced YOLOv5*

3.2.1. *Improved Non-Maximum Suppression (NMS)*

Both NMS and Soft NMS utilize predicted classification confidence as a measure, where higher confidence indicates more accurate localization. In this

study, the improved Soft NMS is used instead of the original NMS, effectively improving the performance of the detection.

3.2.2. *Coordinate attention module*

In this study, we introduce the Coordinate Attention (CA) mechanism, which incorporates positional information into channel attention. Experimental results reveal that incorporating the CA attention mechanism module into the network backbone improves detection accuracy to some extent, particularly in classroom settings with numerous students or when students in the back rows are at a considerable distance from the camera. This enhancement is attributed to the CA module (Fig. 1), which reinforces channel features in the feature map, enabling the network to acquire more detailed and effective information.

Fig. 1 The improved network structure.

3.2.3. *Improved Loss function*

We have improved the initial loss function by incorporating the Efficient-IOU (EIOU) to compute width and height losses, replacing the aspect ratio calculation based on the CIOU foundation. Furthermore, Focal Loss has been introduced concurrently to mitigate the challenge posed by sample imbalance.

4. Experiments and results

The experimental platform in this study was a 64-bit Windows 10 operating system with a Gen Intel Core i5-11400H CPU and NVIDIA RTX 3050 graphics card. The deep learning framework used was PyTorch, with a development environment consisting of PyTorch 1.8 and Python 3.7. The number of training epochs was set to 300, and the dataset was divided into training, validation, and

Shuai Wang, Abdul Samad Shibghatullah, Javid Iqbal Thirupattur, Kay Hooi Keoy

testing sets following a 6:2:2 ratio. To effectively evaluate the experimental results, a comparative experiment was conducted on the self-constructed dataset and the Fer2013 dataset before and after algorithm improvement. The evaluation criteria were detection accuracy (P) and mean average precision (mAP@0.5).

Table 1 presents the experimental outcomes on our self-constructed dataset. Fig.2 illustrates the detection accuracy (*P*) curve of the self-constructed dataset, comparing YOLOv5 with SFER-YOLOv5.

Table 1. Experimental results on self-constructed dataset.

Expression	YOLOv5		SFER-YOLOv5	
	P	*mAP@0.5*	*P*	*mAP@0.5*
happy	0.816	0.853	0.851	0.887
sad	0.845	0.867	0.889	0.909
neutral	0.842	0.857	0.879	0.871
all	0.834	0.859	**0.873**	**0.889**

Fig.2 P curve on the self-constructed dataset.

Table 2 presents the experimental results on the Fer-2013 dataset. Fig. 3 illustrates the *P* curve of the Fer-2013 dataset.

Table 2. Experimental results on FER-2013 dataset.

Expression	YOLOv5		SFER-YOLOv5	
	P	*mAP@0.5*	*P*	*mAP@0.5*
happy	0.863	0.918	0.876	0.911
sad	0.636	0.594	0.669	0.626
neutral	0.694	0.799	0.722	0.515
all	0.731	0.770	**0.756**	**0.784**

Fig.3 P curve on the FER-2013 dataset.

5. Conclusion

This study addresses the challenges of low real-time detection rates and poor timeliness in recognizing students' facial expressions in complex classroom environments while simultaneously focusing on mitigating and preventing cyberbullying incidents. Our *SFER-YOLOv5* model improves the non-maximum suppression (NMS) by replacing it with Soft NMS, enhances the feature extraction capability by adding the Coordinate Attention (CA) mechanism module, and improves the representation ability of target boxes by using EIoU instead of CIoU. Our experimental results are highly encouraging, demonstrating a 3.9% increase in detection accuracy (P) and a 3.0% increase in mAP@0.5 on the self-constructed dataset. Beyond facial expressions, future research will integrate poses and gestures for a comprehensive analysis of online classroom dynamics. This enhances our cyberbullying prevention by understanding students' engagement and emotions more comprehensively.

References

1. J. Bao, X. Tao, Y. Zhou, "An Emotion Recognition Method Based on Eye Movement and Audiovisual Features in MOOC Learning Environment", IEEE Transactions on Computational Social Systems, 2022.
2. P. Ekman, W.V. Friesen, "Constants across cultures in the face and emotion", Journal of personality and social psychology, 17.2 (1971), 124.
3. N. Shakee, R. K. Dwivedi, "A survey on detection of cyberbullying in social media using machine learning techniques", Intelligent Communication Technologies and Virtual Mobile Networks: Proceedings of ICICV 2022, Singapore: Springer Nature, pp. 323-340, 2022.

4. A. Bozyiğit, S. Utku, E. Nasibov, "Cyberbullying detection: Utilizing social media features", Expert Systems with Applications, 179: 115001, 2021.
5. R. Kumar, A. Bhat, "A study of machine learning-based models for detection, control, and mitigation of cyberbullying in online social media", International Journal of Information Security, 21(6), 1409-1431, 2022.
6. G.W. Giumetti, R.M. Kowalski, "Cyberbullying via social media and well-being", Current Opinion in Psychology, 45, 101314, 2022.
7. X. Zhu, et al, "TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios", Proceedings of the IEEE/CVF international conference on computer vision, pp. 2778-2788, 2021.

Authors Introduction

Mr. Shuai Wang

He is currently pursuing the Ph.D. degree in Computer Science programme with the Institute of Computer Science and Digital Innovation, UCSI University, Kuala Lumpur, Malaysia. He is also working in the Department of Physics and Electronic Engineering, Yuncheng University, Yuncheng, Shanxi, China. His current research interests include computer vision, machine learning and multi-modal learning.

Dr. Abdul Samad Bin Shibghatullah

He received the Ph.D. degree in computer science from the Brunel University of Uxbridge, United Kingdom. He is currently an Associate Professor at Institute of Computer Science & Digital Innovation, UCSI University, Kuala Lumpur, Malaysia. His current research interests include optimization, modelling and scheduling, Robotics.

Dr. Javid Iqbal Thirupattur

He works as an Assistant professor at UCSI University, Malaysia. He holds a Ph.D. degree in Information and Communication Technology from the National Energy University Malaysia. His research interests include Multimedia, Augmented Reality, and Virtual Reality.

Dr. Keoy Kay Hooi

He received the Ph.D. degree at Sheffield Hallam University, United Kingdom in 2006. He is currently an Associate Professor, UCSI University, Kuala Lumpur, Malaysia. His current research interests include machine Learning, block chain technology & management information system.

A Method of Recognizing Body Movements Based on a Self-viewpoint Video

Iichirou Moribe

Graduate School of Engineering, Kyushu Institute of Technology, 1-1 Sensui, Tobata-ku, Kitakyushu, 804-8550, Japan

Yui Tanjo

Faculty of Engineering, Kyushu Institute of Technology, 1-1 Sensui, Tobata-ku, Kitakyushu, 804-8550, Japan

Email: tanjo@.cntl.kyutech.ac.jp

Abstract

The most critical human sensory function resides in vision. This paper focuses on utilizing visual information, specifically self-perspective footage, to identify individual movements. Existing researches require third-party filming to recognize human body movements and states. The proposed method, on the other hand, simply attaches a camera to the human head and enables the recognition of the subject's actions. Consequently, it becomes easier to monitor daily movements of a human and gather his/her data on body kinetics. This approach would be beneficial in scenarios involving individuals engaging in risky behavior or, during a certain emergency, providing valuable assistance.

Keywords: My VISION, Posture estimation, Optical flow, HSV conversion

1. Introduction

Vision is the most crucial function among human sensory organs. Human intakes a vast amount of information through vision. Therefore, self-perspective footage can potentially serve as the primary source of information from an individual. In fact, it has been evident that visual information is more potent in perceiving human posture and bodily movements compared to non-visual information [1]. Therefore, it was considered possible to recognize the filmmaker's physical state by analyzing self-perspective footage.

Maintaining physical health has become more commonplace, exemplified by the widespread use of smartwatches. Smartwatches offer various functions such as recording heart rate, blood oxygen levels, and sleep patterns. By utilizing these features to log daily physical conditions, it becomes possible to promptly recognize deviations from normal bodily states.

Hence, this study aims to develop a method for analyzing and recording an individual's activities from a self-perspective video. Similar research includes studies on self-posture estimation [2], methods utilizing single-eye images based on HOG features [3], and methods employing deep learning [4]. However, these methods focus on action estimation or recognition rather than understanding a person's activity. Additionally, the techniques using HOG-based single-eye images or employing deep learning require capturing individuals externally, which is entirely different from the self-perspective video approach in this paper.

2. Methodology

This section describes feature extraction methods. The features to be extracted are the values of the hue and the norm of the optical flow on an image.

2.1. Detection of area similarity

This subsection describes the method for detecting area similarity. By dividing the image into multiple blocks and comparing the H (Hue) histogram features of specific blocks, area similarity is determined.

2.1.1 HSV Conversion

First, the RGB color image, which serves as the input image, is converted from the RGB color space to the HSV color space. The RGB color space represents a color space with red, green, and blue as coordinate axes, while the HSV color space represents a color space with hue, saturation, and value as coordinate axes.

2.1.2 Creation of H(Hue) histogram

In the proposed method, a histogram based on the Hue (H) is constructed. The values of S (Saturation) and V (Value) are not taken into account. The H histogram, as depicted in Fig. 1, represents the hue values on the horizontal axis and the frequency count on the vertical

Iichirou Moribe, Yui Tanjo

axis. Although the original range of hue values is 1 to 360, for expedited processing, it is scaled to 1 to 180.

Fig. 1. H (Hue) Histogram

2.1.3 *Calculation of similarity*

The H histogram is used to determine the similarity between two images. Here, the similarity is calculated using the following formula. The meanings of the characters in Eq. (1) are given in Table 1.

$$S(A, B) = \frac{\sum_{i=0}^{N} \min(h_i^A, \ h_i^B)}{\sum_{i=0}^{N} h_i^A} \qquad (1)$$

Table 1. Meanings of the characters in Eq. (1)

N	Number of pixels in the image
h_i^x	The ith element of histogram x
$\min(x_A, x_B)$	Minimum value of x_A and x_B

2.2 *Derivation of optical flow*

This subsection describes the derivation of optical flow. The optical flow is the movement vector of the camera, or the subject obtained by mapping feature points in two consecutive frame images extracted from the video.

2.2.1 *Feature point extraction and description*

With respect to two consecutive frame images, the feature points in the first frame image are focused on. After extracting the feature points, the feature values of each feature point are described, and the feature points are matched in the two frame images. The Shi-Tomasi corner detection method is used for feature point extraction and feature point description [5]. The image from which the feature points were extracted is shown in Fig. 2.

Fig. 2. Input image with feature point extraction

2.2.2 *Feature point matching*

The Lucas-Kanade method is used for matching feature points [6]. The Lucas-Kanade method is one of the leading methods for deriving optical flow and is computationally less expensive than the method that finds all pixels in the image by searching for them.

2.2.3 *Removal of outliers*

The optical flow is obtained by the feature point matching, but the optical flow obtained from the actual video contains many outliers. Outliers are false flows obtained by matching different feature points and need to be removed for accurate analysis. Outliers are removed by applying RANSAC to modelling with homography matrices.

An example of the optical flow obtained by the above process is shown in Fig. 3. Note that the green points in the figure represent the end points of the optical flow.

Fig. 3. Derived optical flow.

2.3 *Normal state recognition*

In this subsection, the method of normal state recognition is first described. Next, the methods for classifying states are described.

2.3.1 *Feature extraction methods*

The input image is separated into nine blocks and each block image is assigned a number as shown in Fig. 4. Focusing on blocks 2, 5 and 8, the three block images are HSV-transformed and H histograms are created. The H histograms are compared to obtain the similarity between block 8 and block 2, and between block 8 and block 5, respectively. The similarity between block 8 and block 2 and between block 8 and block 5 is calculated by comparing the H histograms.

Block 1	Block 2	Block 3
Block 4	Block 5	Block 6
Block 7	Block 8	Block 9

Fig. 4. Separated input image

Furthermore, in the method, the number and norm of optical flows are the key features. First, the average value of the norm of the optical flow is calculated with each frame, and then the maximum value of the norm in the whole video is calculated. This value is used to evaluate the intensity of the motion of the entire video. The average value of the norm of the optical flow is calculated by the following equation. The meanings of the characters in Eq. (2) are shown in Table 2.

$$V_f = \frac{1}{N_f} \sum_{i=1}^{N_f} V_i \qquad (2)$$

Table 2. Meanings of the characters in Eq. (2)

N_f	Total number of optical flows obtained between the fth frame and the f+1th frame
V_i	Norm of each optical flow

2.3.2 *Prior statistics*

Since the estimation of each state is done by comparison with the normal state, it is necessary to know in advance what properties each state has in relation to the normal state. For the three states, feature extraction is carried out using the above procedure with the three videos.

The results of the validation of the properties of the normal state are presented in Table 3. The results of the validation of the properties of the 'looking down' state are presented in Table 4. The results of the validation of the properties of the 'stumble' state are presented in Table 5.

Table 3. Results of the validation of the properties of the normal state

Scene	$\bar{S}$	V_{max}
Video 1	0.53	118
Video 2	0.40	109
Video 3	0.51	124

Table 4. Results of the validation of the properties of the 'looking down' state

Scene	$\bar{S}$	V_{max}
Video 1	1.29	113
Video 2	1.19	110
Video 3	1.06	75.6

Table 5. Results of the validation of the properties of the 'stumble' state

Scene	$\bar{S}$	V_{max}
Video 1	0.61	271
Video 2	0.54	181
Video 3	0.45	194

From these results, it can be read that $\bar{S}$ is greater in the 'looking down' state among the three states. This means that in the 'looking down' state, the similarity of block 2, block 5 and block 8 is greater. It can also be seen that V_{max} in the 'stumble' state is the largest among the three states. This indicates that the motion in the 'stumble' state is the most intense.

2.3.3 *State analysis*

In the proposed method, the motion states are classified into three states. Classification 1: If the value of S satisfies Eq. (3), the image is classified as a 'looking down' image. If the value of S does not satisfy Eq. (3), the image is classified as 'other' image. 'other' can be either 'normal' or 'stumble'. The state of Classification 1 is determined using the following formula.

$$S - \bar{S} \geq TH1 \qquad (3)$$

Here S is the similarity in the current state and $\bar{S}$ is the average of the similarity in the normal state. $TH1$ is the threshold value in Classification 1.

Sr is the proportion of the images in the image group that are classified as 'looking down' in Classification 1, and if the value of Sr satisfies Eq. (4), the image is classified as "looking down". If the value does not satisfy Eq. (4), it is assumed to be in the 'other' state and proceeds to the next state estimation. If the value of $Vmax$ is satisfies Eq. (5), the state is judged as the 'stumble' state. If none of the above conditions is satisfied, the state is judged as normal. Eq. (4) and Eq. (5) are shown below.

$$Sr - \bar{S_r} \geq TH2 \qquad (4)$$

$$Vmax - \overline{V_{max}} \geq TH3 \qquad (5)$$

Here Sr and $Vmax$ are the evaluated values of similarity and intensity of motion in the current condition, respectively. $\bar{S_r}$ and $\overline{V_{max}}$ are the average values of similarity and intensity of motion, respectively, in the normal condition. $TH2$ and $TH3$ are the threshold values for each condition.

3. Experimental Result

3.1. *Experimental methods*

An experiment is conducted with a camera worn on the head. Fig. 5 shows an image of the camera worn by a subject during the experiment.

Iichirou Moribe, Yui Tanjo

(a)   (b)

Fig. 5. Image of a worn camera. (a) side view,(b) rear view.

The details of the experimental method are as follows:

1. Feature extraction was carried out from five videos in the 'normal' state, and a database was created by averaging the feature values in each video.
2. Eight videos containing each of the three states - 'normal', 'looking down' and 'stumbling' - taken at different locations and at different dates and times were used as input videos to estimate each of the states.

The parameters during the experiment are shown in Table 6.

Table 6. Parameters' values

	Value	
Threshold 1	0.35	$TH1$ in Eq. (2-2)
Threshold 2	0.65	$TH2$ in Eq. (2-3)
Threshold 3	55	$TH3$ in Eq. (2-4)

3.2. *Result*

The results of the experiment are shown in Table 7. For each of the eight videos, the number of correct responses for the 'normal', 'looking down' and 'stumble' states were 6, 7 and 7 respectively.

Table 7. Results of the experiment.

state	normal	looking down	stumble
normal	6	0	2
looking down	1	7	0
stumble	1	0	7

4. Conclusion

In this paper, we proposed a method for estimating the activity state of the camera wearer from the self-viewpoint video obtained using a self-viewpoint camera. Three human walk states were detected by comparing H (Hue) histograms and using features calculated from the optical flow. In the experiment, three states were detected: 'normal', 'looking down' and 'stumble' states. The results showed that estimation was successful in more than 75% of all states. However, in the estimation of all states, misrecognition occurred in which a different state was estimated. Future issues include the estimation of states in case of sudden changes in vision and estimating more precise states corresponding to the angle of looking down.

References

1. D. Lee, E. Aronson: "Visual proprioceptive control of standing in human infants", Attention Perception & Psychophysics, vol. 15, No.3, pp. 529-532, 1974.
2. J. K. Tan, T. Kurosaki: "Estimation of self-posture of a pedestrian using MY VISION based on depth and motion network", Journal of Robotics, Networking and Artificial Life, vol. 7, No. 3, pp. 152-155, 2020.
3. K. Onishi, T. Takiguchi, Y. Ariki: "3D human posture estimation using the HOG features from monocular image", 19th International Conference on pattern Recognition, pp. 1-4, 2008.
4. Z. Cao, T. Simon, S. E. Wei, Y. Sheikh: "Realtime multi-Person 2D pose estimation using part affinity fields", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7291-7299, 2017.
5. J. Shi, C. Tomasi: "Good features to track", IEEE Conference on Computer Vision and Pattern Recognition, 1994.
6. B. D. Lucas, T. Kanade: "An iterative image registration technique with an application to stereo vision", Proceeding of Imaging Understanding Workshop, pp.121-130, 1981.

Authors Introduction

Mr. Iichirou Moribe

Mr. Moribe received his Bachelor's degree in Engineering in 2022 from the Control Engineering, Kyushu Institute of technology in Japan. He is currently a master student in Kyushu Institute of Technology, Japan.

Dr. Yui Tanjo

Dr. Tanjo is currently a professor with the Department of Mechanical and Control Engineering, Kyushu Institute of Technology. Her current research interests include ego-motion analysis by MY VISION, three-dimensional shape/motion recovery, human detection, and its motion analysis from video. She was awarded SICE Kyushu Branch Young Author's Award in 1999, the AROB Young Author's Award in 2004, the Young Author's Award from IPSJ of Kyushu Branch in 2004, and the BMFSA Best Paper Award in 2008, 2010, 2013 and 2015. She is a member of IEEE, The Information Processing Society, The Institute of Electronics, Information and Communication Engineers of Japan.

Supporting Safe Walk of a Visually Impaired Person at a Station Platform Based on MY VISION

Shintaro Yamada

Graduate School of Engineering, Kyushu Institute of Technology, 1-1 Sensui, Tobata-ku, Kitakyushu, 804-8550, Japan

Yui Tanjo, Seiji Isihkawa*

Faculty of Engineering, Kyushu Institute of Technology, 1-1 Sensui, Tobata-ku, Kitakyushu, 804-8550, Japan
*Emeritus Professor of Kyushu Institute of Technology
Email: tanjo@cntl.kyutech.ac.jp

Abstract

When individuals with visual impairment go out, public transportation such as trains and buses is commonly used. However, many of them experience accidents, such as falling from train platforms or tripping due to unexpected contact with other passengers. To solve this problem, we propose a method using the MY VISION system which detects obstacles that may pose risks to individuals with visual impairment. The proposed method detects obstacles such as passengers' pillars and platform edges at train stations. We employ an RGB-D camera for capturing frontal views of a user, use depth images to detect the edge of obstacles and level differences, and give warning to the visually impaired user based on the distance between him/her and the detected obstacle. Experimental results show satisfactory performance of the method.

Keywords: Visually impaired, Safety, Train station platform, Obstacles detection, MY VISION

1. Introduction

Public transportation is often used as a means of mobility for individuals with visual impairments, and more than half of them frequently utilize railways [1]. However, approximately 30% of visually impaired individuals have experienced falls from station platforms, resulting in an average of 76 fall incidents annually [2].

There are methods for detecting steps and obstacles to assist visually impaired individuals in walking [3], [4]. However, the methods that use depth images to acquire the walking plane and to detect steps and obstacles are affected by noise, as the distance from the camera increases, it results in a narrow and less accurate detection range. Additionally, some methods use color information, but they are often designed for indoor use and are influenced by the factors such as time of day and weather when used outdoors.

In this paper, we propose a method that detects step edges directly using depth images and determines the distance from the pedestrian to the step. The proposed method, designed to assist visually impaired individuals on station platforms, involves region segmentation using Graph Based Segmentation on depth images and the detection of step lines using Line Segment Detector.

2. Method

2.1 Obstacles detection

Initially, the coordinates conversion is performed from the image coordinate system to the camera image coordinate system. The transformation formula is given as follows.

$$X(x,y) = \frac{(x - c_x) \times Z(x,y)}{f_x} \quad (1)$$

$$Y(x,y) = \frac{(y - c_x) \times Z(x,y)}{f_y} \quad (2)$$

$$Z(x,y) = d \quad (3)$$

Here c_x is the optical center in the *x*-axis direction, c_y is the optical center in the *y*-axis direction, f_x is the focal length in the *x*-axis direction, and f_y is the focal length in the *y*-axis direction.

Changes in $Y(x,y)$ and $Z(x,y)$ with respect to the change in the *y*-coordinate, as well as the changes in $X(x,y)$ and $Z(x,y)$ with respect to the change in the *x*-coordinate, are expressed by the following equations.

Shintoro Yamada, Yui Tanjo, Seiji Ishikawa

$$\frac{dY(x,y)}{dy} = Y(x,y) - Y(x,y+k) \qquad (4)$$

$$\frac{dZ(x,y)}{dy} = Z(x,y) - Z(x,y+k) \qquad (5)$$

$$\frac{dX(x,y)}{dx} = X(x,y) - X(x+k,y) \qquad (6)$$

$$\frac{dZ(x,y)}{dx} = Z(x,y) - Z(x+k,y) \qquad (7)$$

The method computes the change in angle of the gradient of the Z coordinate with respect to the gradient of the Y coordinate at the point (x,y), as well as the change in angle of the gradient of the Z coordinate with respect to the gradient of the X coordinate. If these angles exceed a certain threshold, the method judges it as an obstacle region. The formula and conditional statement are shown as follows.

$$d\_YZ(x,y) = \arctan\left(\frac{\frac{dY(x,y)}{dy}}{\frac{dZ(x,y)}{dy}}\right) \qquad (8)$$

$$d\_XZ(x,y) = \arctan\left(\frac{\frac{dX(x,y)}{dx}}{\frac{dZ(x,y)}{dx}}\right) \qquad (9)$$

$$d\_YX(x,y) = \arctan\left(\frac{\frac{dX(x,y)}{dx}}{\frac{dY(x,y)}{dy}}\right) \qquad (10)$$

$$object = \begin{cases} 1 & if\, d_{YZ} > t\ and\ d_{XZ} > th\ and\ d_{YX} = 0 \\ 0 & otherwise \end{cases} \qquad (11)$$

2.2 *Step detection at a station platform*

The method designates the regions at both ends of the area, excluding the obstacle region as candidate regions where steps are present. Graph Based Segmentation (GBS) is applied to these candidate regions. By applying GBS, it is possible to segment and merge regions with similar features in the image. This allows for the detection of steps by dividing regions according to different planes of varying height.

2.3 *Line detection at a step edge*

Line Segment Detector (LSD) is then applied to the candidate regions which contain the steps detected through region segmentation. LSD detects candidate line segments at the edges of planes with different heights. Linear approximation is performed using RANSAC on the endpoints of the line segments detected by LSD to detect the edges of the train platform. The detected lines are represented by the equation $y = ax + b$ using the coordinates (x, y), and the gradient a helps determine the direction of the platform edge from the perspective of the observer. The conditional statements are as follows.

$$Dir = \begin{cases} Right & if\ a > th_1 \\ Left & if\ a < th_2 \\ Front & else \end{cases} \qquad (12)$$

2.4 *Distance from a step edge*

Let g_c be the point where the user of the proposed method is standing. It is defined as g_c = (width/2, *height*). Let g_p be a point on the detected line segment. Referring to the depth information on the left and the right of the point g_p which is the nearest point from g_c, the distance D_{min}^H is defined as the distance from the platform edge to the user. It is formulated as follows.

$$D = \min_{p=1,2,\dots,p}(|g_c - g_p|) \qquad (13)$$

$$d(x, y, d_e) = D \qquad (14)$$

$$D_{min}^H = \min_{d_{e^*}}(d(x+1, y, d_{e^*}), \\ d(x-1, y, d_{e^*})) \qquad (15)$$

3. Experimental result

In the performed experiment, an RGB-D camera was mounted on the chest of a subject, and video taking was conducted at a train platform. The proposed method was applied to the captured depth images, and comparison with ground truth was performed to assess the accuracy of the method. Experimental videos were captured in four situations involving obstacles such as people and luggage. Each video consists of 95 frames. Accuracy evaluation is done using the following formula.

$$overlap = \frac{GT \cap OA}{OA} > T \qquad (16)$$

$$angular\ error = |a_{GT} - a_d| < T_{q_a} \qquad (17)$$

$$distance\ error = |b_{GT} - b_d| < T_{q_b} \qquad (18)$$

(a)   (b)   (c)   (d)   (e)

Fig.1. Experimental result : (a) Original color image, (b) depth image, (c) ground truth, (d) obstacle detection, (e) result of platform edge detection（in blue color）

Table 1. Experimental result

Location	Obstacle	Percentage of correct line segment location [%]	Percentage of correct angles [%]	Percentage of correct distance [%]
The left platform edge	Person	100	94.5	98.6
The right platform edge	Person	100	91.0	98.0
The left platform edge	Person & luggage	100	88.0	97.0
The right platform edge	Person & luggage	98.1	92.7	89.0

Here GT represents the area of the ground truth line segment, OA represents the area of the line segment detected by the proposed method, and T is the threshold. In this experiment, T is set to 0.5. The success of the detection is determined when the ratio of overlapping line segment areas exceeds the threshold. a_{GT} represents the distance on the image between the midpoint of the ground truth line segment and g_c, whereas a_d represents the distance on the image between the midpoint of the detected line segment and g_c. T_{q_a} is the threshold. In this experiment, T_{q_a} is set to 15 pixels. The success of the detection is declared when the error between a_{GT} and a_d is below the threshold. b_{GT} represents the angle between the ground truth line segment and the *x*-axis, and b_d represents the angle between the detected line segment and the *x*-axis. T_{q_b} is the threshold. In this experiment, T_{q_b} is set to 5 degrees. The success of the detection is declared when the error between b_{GT} and b_d is below the threshold. An example of experimental results is shown in Fig. 1.

The overall experimental results are presented in Table 1. The average accuracy for line segment position, line segment angle, and line segment distance were 99.5%, 91.6%, and 95.6%, respectively.

4. Conclusion

This paper proposed a method of detecting the edges of steps on a train platform automatically using a chest-mounted RGB-D camera. This camera captures frontal scenes of a user with visual impairment, when he/she is walking. By detecting platform edges, the proposed method is able to contribute to the safety of a user with visual impairment who takes a train. Future work includes improving the proposed method for real-time processing and conducting validation experiments.

References

1. Japan Federation of the Blind and Visually Impaired, Mainichi Shimbun/Japan Federation of the Blind Questionnaire Survey on Railway Stations for the Visually Impaired Survey Results, http://nichimou.org/wp-content/uploads/2017/03/mainichiannke-to.pdf, (2022-9-8)
2. Ministry of Land, Infrastructure, Transport and Tourism, Study Group for Safety Improvement at Station Platforms www.milt.go.jp/tetudo/tetudo_fr7_000015.html,(2022-9-8)
3. K. Imai, I. Kitahara, and Y. Kameda, Detecting walkable plane areas by using RGB-D camera and accelerometer for visually impaired people, 2017 3DTV Conference, The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), Copenhagen, Denmark, pp. 1-4, 2017.
4. Yuna Seibu, Haiyuan Wu, Tadayoshi Shioyama, Fundamental Study of Image-Based Solo Walking System for the Visually Impaired - Detection of Platform Edge, Human Interface News and Report, 13, pp.467-474, 1998.
5. Yuki Kawaguchi, Joo Kooi Tan, Seiji Ishikawa, Takashi

Shinomiya, Supporting Safe Walk of a Visually Impaired Person at a Railway Station Platform Based on MY VISION, The 2022 International Conference on Artificial Life and Robotics(ICAROB2022), pp. 447-450, 2022.

Authors Introduction

Mr. Shintaro Yamada

Mr. Yamada received his B.E. degree in Department of Control Engineering in 2022 from the Faculty of Engineering, Kyushu Institute of Technology in Japan. He is acquiring the M.E. in Kyushu Institute of Technology.

Dr. Yui Tanjo

Dr. Tanjo is currently a professor with the Department of Mechanical and Control Engineering, Kyushu Institute of Technology. Her current research interests include ego-motion analysis by MY VISION, three-dimensional shape/motion recovery, human detection, and its motion analysis from video. She was awarded SICE Kyushu Branch Young Author's Award in 1999, the AROB Young Author's Award in 2004, the Young Author's Award from IPSJ of Kyushu Branch in 2004, and the BMFSA Best Paper Award in 2008, 2010, 2013 and 2015. She is a member of IEEE, The Information Processing Society, The Institute of Electronics, Information and Communication Engineers of Japan.

Prof. Seiji Ishikawa

Professor Ishikawa graduated from Tokyo University and was awarded BE, ME and PhD there. He is now Emeritus Professor of Kyushu Institute of Technology. He was Visiting Researcher of the University of Sheffield, UK, and Visiting Professor of Utrecht University, NL. His research interests include visual sensing & 3-D shape/motion recovery. He was awarded The Best Paper Award in 2008, 2010, 2013 and 2015

Development and evaluation of a learning support robot for vector learning

Kosei Machida
Tokyo Gakugei University, 4-1-1 Nukuikita Koganei, Tokyo, 184-8501, Japan

Shinichi Imai
Tokyo Gakugei University, 4-1-1 Nukuikita Koganei, Tokyo, 184-8501, Japan
Email: m224203x@st.u-gakugei.ac.jp, shimai@u-gakugei.ac.jp

Abstract

Vectors studied in high school are a new concept that differs from those studied in previous scalars. Therefore, it has been pointed out that difficulties arise in the conceptual formation of vectors. This study aims to develop a learning support robot that can visualize vector information by robot movements, and to acquire vector concepts for students through classroom practice using the robot. The robot operates on a piece of imitation paper, marking the start and end points and connecting them with arrows to visualize the vectors. In a class using the robot, the student predicts the sum of the vectors, and the robot confirms it. A questionnaire was administered before and after the class. The results revealed significantly higher mean values when comparing the pre- and post-assessments.

Keywords: Educational Support, Robot, Omni-wheel, Omnidirectional movement, High school mathematics

1. Introduction

Recent large-scale surveys such as OECD-PISA and TIMMS have shown that "while there has been some improvement in the percentage of students who responded positively to the enjoyment of learning mathematics and its relationship to the real world, there are still issues in terms of motivation to learn, such as the fact that the percentage is still low compared to other countries" and in upper secondary schools, "the motivation to learn mathematics is not high. [1] Therefore, in the Guidelines for the Course of Study for Senior High Schools announced in 2018, it is stated that in senior high school mathematics "problems that reflect the real world are handled, students who are not highly motivated to study mathematics are made aware of the significance of studying mathematics, and this will lead to increased motivation and the development of mathematical ability. This will lead to the development of mathematical skills.

"Vectors" are "quantities that have magnitude and direction," a new concept that is different from the "quantities that express only magnitude" that students have learned in the past. Therefore, various difficulties have been pointed out in the concept formation of vectors. One reason for this is that vector learning is conducted only at the level of abstraction on the blackboard, making it difficult to formulate concepts. In addition, lack of understanding of the phases of vectors has also been pointed out [2], which may be caused by a lack of understanding of the concept of vectors themselves.

It has been noted that the use of robots in schools can have a positive impact on students. The use of robots, as part of STEM education, has been noted to have a positive impact on students' understanding of science, technology, engineering, and mathematics; students acquire a broader range of knowledge because robots are associated with technological applications that span multiple disciplines; and the use of robots can increase students' motivation and engagement in the classroom. The use of robots can have a positive impact on students by increasing their motivation and engagement in the classroom. [3], [4], [5], [6]

Therefore, the purpose of this study is to develop a learning support robot for learning vectors and to propose a class in which students can acquire the concept of vectors in addition to numerical calculations. By using the robot, students can learn vectors not only on the blackboard but also through concrete objects, which is expected to enhance their conceptual understanding. Through the proposed class, we expect to achieve the learning effects of understanding the concept of vectors and improving the motivation to learn.

2. Robots to be developed.
2.1. Definition of robot requirements

The robot is used to consider the problem of two vectors with respect to the sum of the vectors, and to develop a robot that represents the motion of the vectors in terms of the robot's motion. The angle between the two vectors shall be limited to 0-180°. The robot should be easy to operate for high school students to acquire the concept of

vectors. To visualize the vectors (trajectory of the robot), the robot is placed on a piece of imitation paper.

2.2. Robot Functions

The speed of the robot's motor is defined as the vector magnitude, the direction of the tires is defined as the vector direction, and the trajectory of the robot's progress is defined as the vector motion. When studying vector sums, the robot's trajectory can be defined as the sum of vectors by simultaneously performing the motions of the two vectors before combining the vectors.

2.3. Robot Design

This Robot will be developed using Artec Robo from Artec, Inc. The robot to be developed uses Artec blocks, DC motors, servo motors, and a battery box Studuino: bit. The omni wheel is an omni-directional moving mechanism. [7] The omni wheel is an omni-directional moving mechanism, and the wheels using the omni wheel can move in any direction. [8] Studuino: bit is used as the microcontroller. Artec blocks are used to assemble the robot. The servo motors are used to change the direction of the tires. DC motors are used to rotate the wheels, which allow the robot to move on a flat surface. Servo motors are used to change the direction of the tires. Electric power is supplied from a battery box. The size of the robot is about the same as the size of the classroom. The size of the robot should be large enough for students to hold in their hands and manipulate on a desk.

2.4. Correspondence between robot motion and vectors

Fig. 1 shows the appearance of this teaching material. Hereafter, the two wheels at the bottom of the robot will be referred to as Wheel A and the two wheels at the top of the robot in Fig. 1 as Wheel B. In Fig. 1, the direction of rotation and speed of the tires are specified for Wheel A and Wheel B, respectively, and the trajectory of the robot's motion is referred to as vector motion. The motion of the omni wheels is shown in Fig. 2. As shown in Fig. 2, the velocity in the x-direction is V_x and the velocity in the y-direction is V_y. Let the rotational velocities of the omniwheel be V_1 and V_2, respectively. considering the x-axis as the axis horizontal to V_1, the x-axis as 0 rad, counterclockwise as positive, and the angle between the two wheels as θ, the calculation of the velocity components can be expressed by the Eq. (1). With V_x and V_y known, the direction of movement is determined, and the direction of rotation and rotation speed of V_1 and V_2 are controlled.

$$V = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \cos(\theta) & \sin(\theta) \end{bmatrix} \begin{bmatrix} V_x \\ V_y \end{bmatrix} \quad (1)$$

Fig.1 Robots to be developed.

Fig.2 Omni wheel movement

However, V_1 is assumed to move in the direction V_x. If $V_1 = V_2$, the body moves straight ahead. The use of an omniwheel allows the robot to move without turning, so the sum of vectors can be moved in a direction different from the direction in which the wheels rotate. The wheels can be changed to any position by using servo motors in the upper positions where the wheels are mounted. This allows the direction of the wheel to be changed.

2.5. Image Analysis

Image analysis is performed to check whether the sum of vectors can be represented by the robot's motion. For the image analysis, we use a camera stand and a video shot from the top of the robot with an iPhone13. The frame rate of the video is 30 fps. PV Studio 2D PRO is used as the image analysis software [9]. Place a marker at the center of the robot and record the axis of the robot's movement. The x-axis and y-axis are set so that the center becomes the origin. The frame when the robot starts moving is used as the reference frame. Because of the blurring of the axis, the frame when the axis increases or decrease in a certain direction without any increase or decrease in the axis is the frame of the start of the movement. Similarly, the last frame in which the axis increases or decrease in a constant direction without increasing or decreasing back and forth is the frame at the end of the motion. Image analysis is performed by combining the two motions when the sum of vectors is decomposed. The analyzed data is quantified, theoretical values are derived, and a comparison of the motion and

the theoretical values is made to confirm that the sum of vectors can be expressed.

2.5.1. Image Analysis Theoretical Value

Define the first values of x-axis and y-axis (axis of the initial position) x_1 and y_1. Thus, the x-axis and y-axis values in frame n can be defined as x_n and y_n. The general forms of the x - and y-axis are x_n and y_n, respectively, where x_1 and y_1 are first term values.

The robot is assumed to move in the minus direction of the x -axis. Therefore, if the robot stops moving at frame m, the total number of frames can be represented by m. If the range of frames in the x -axis and y-axis directions is defined as a and b, respectively, then a and b can be expressed by Eq. (2) and Eq. (3), respectively.

$$a = |x_m - x_1| \qquad (2)$$
$$b = |y_m - y_1| \qquad (3)$$

If the values advancing in the x-axis and y-axis directions per frame are defined as g and h, respectively, then g and h can be represented by Eq. (4) and Eq. (5), respectively.

$$g = \frac{a}{m} \qquad (4)$$
$$h = \frac{b}{\frac{m}{2}} \qquad (5)$$

In the case of h, the vector shows two directions, so the number of frames is $m/2$, since the direction changes after half the time. Therefore, the values x_n and y_n of the x and y-axis in the nth frame can be expressed by Eq. (6), Eq. (7) and Eq. (8). However, when $n = 1$, $x_1 = g$ and $y_1 = h$.

$$x_n = x_{n-1} - g \ (n \geqq 2) \qquad (6)$$
$$y_n = y_{n-1} - h \quad (1 < n \leq m/2) \qquad (7)$$
$$y_n = y_{n-1} + h \quad (m/2 < n \leq m) \qquad (8)$$

When the number of frames is half, y_n can be expressed in two equations for $n/2$, since the sum of the vector's changes direction after half the number of frames has elapsed, since the vectors indicate two directions.

2.5.2. Results and Discussion of Image Analysis

In PV Studio 2D PRO, each displacement of one axis results in a movement of one meter. Based on this, comparing the results of image analysis with the theoretical values, the motion when the left and right wheels are moved simultaneously is shown in Fig. 3.

Fig.3 Results of Image Analysis.

When the number of frames is half, that is, when the direction of the vector sum changes, there is a difference of several frames between the measured and theoretical values. This may be due to the motor. We believe that the individual differences in the motors caused the slight discrepancies.

In addition, there was a discrepancy in the y-axis between the initial value and the final value of the measured value. This can be attributed to friction. Friction is considered to have shifted the y-axis by approximately 0.05 meter compared to the initial position.

The results show that the approximate shape of the sum of the vectors can be represented, although there is an error of a few centimeters from the theoretical value. In order to improve these deviations, we believe that the problem can be solved by making the arrows connecting the trajectories thicker. We also believe that the learning about the robot mechanism can be expanded by having the students consider the reasons for the deviations from the theoretical values.

3. Classroom Practice

3.1. Class Structure

In the class, the sum of vectors is represented by moving a robot on paper, attaching points to the starting and ending points of the robot, and connecting these points with arrows. The goal of the class is "to be able to consider the sum of vectors by transforming the decomposed vectors. The structure of the class is shown in Table 1. A questionnaire will be administered before and after the class.

Table 1. Class Flow.

Intro.	Explanation of how to use the robot. About vectors and sums in one direction.
Dev.	Sum of vectors in two directions Same magnitudes on the left and right Different magnitudes of left and right
Con.	Confirmation of the sum of vectors

Kosei Machida, Shinichi Imai

Table 2. Results of pre-post comparison of questionnaires (t-test (with correspondence))

questionnaire	Pre M	Pre SD	Post M	Post SD	t-value	p-value		effect d
1. Do you think you enjoy learning?	3.25	1.24	3.50	1.16	1.17	.261		0.21
2. Do you think you are able to participate willingly in class?	3.50	1.21	3.75	1.00	0.70	.497		0.23
3. Would you like to find out more about what you have learned?	2.31	1.08	3.00	1.10	2.71	.016	*	0.64
4. Can relate what they have learned to what they have studied and to what is happening in their daily lives and soc	2.38	1.20	2.80	1.13	1.00	.333		0.32
5. Can base one's opinions and thoughts on reasons and in a well-sequenced manner.	2.31	1.01	2.81	1.28	1.52	.150		0.44
6. Can imagine and look ahead and think about what could happen in relation to what they are learning.	2.31	0.95	3.63	1.09	3.88	.001	**	1.30
7. Organize materials and information gathered for a purpose into diagrams and tables.	2.56	1.21	3.06	1.34	1.94	.072		0.39
8. Can express facts and results based on the regularity and lawfulness of things.	2.50	1.03	3.25	1.29	2.54	.023	*	0.65

※In the post-survey, add "Through the use of the robot in the class" at the beginning of the survey item. $*p < .05 ; **p < .01$

3.2. Survey Summary

On October 11, 2023, 17 high school sophomores (16 pre-survey, 17 post-survey) will be surveyed. The subjects are students who have not studied vector sums. The questionnaire will be administered using a five-item method with reference to the literature [10]. In order to measure knowledge, survey questions were administered before and after the survey. The pre-survey was conducted before the class, and the post-survey was conducted after the class. A *t*-test (with correspondence) was used to analyze the differences in attitudes toward independent.

3.3. Result

The results of the survey are shown in Table 2; significant differences were found in three items. For the knowledge question, the response rate for Question 1 was 62.4% before and 93.8% after the response. For question 2, the pre-response rate was 18.8% and the post-response rate was 50.0%.

4. Conclusion

In this study, we developed a learning support robot that represents vector sums. The robot was also used in a class. A questionnaire survey was conducted, and significant differences were found. In addition, the percentage of correct answers in the test was higher than before and after the class. In the future, we plan to further analyze the questionnaires and modify the robot and the classes.

Acknowledgements

References

1. Ministry of Education, Culture, Sports, Science and Technology: "The Revisions of the Courses of Study for Elementary and Secondary School (2018) "
2. BRA Latifa, E Purwaningsih, and S Sutopo;"Identification of students' difficulties in understanding of vector concepts using test of understanding of vector", Journal of Physics: Conference Series, vol2098, pp.1-5(2021)
3. Lai Poh Emily Toh, Albert Causo, Pei-Wen Tzuo, I-Ming Chen, and Song Huat Yeo : "A Review on the Use of Robots in Education and Young Children", Educational Technology & Society, 19 (2), pp.148-163(2016) .
4. D. Lancheros-Cuesta and R. Fabregat: "Educational Robotics Intervention in the Motivation of Students," in IEEE Revista Iberoamericana de Tecnologias del Aprendizaje, vol. 17, no. 2, pp. 131-139, (2022).
5. S. Bruder and K. Wedeward: "Robotics in the classroom," in IEEE Robotics & Automation Magazine, vol. 10, no. 3, pp. 25-29, (2003).
6. Kai Wang, Guo-Yuan Sang, Lan-Zi Huang, Shi-Hua Li and Jian-Wen Guo : "The Effectiveness of Educational Robots in Improving Learning Outcomes", A Meta-Analysis, Sustainability 2023, 15, 4637, pp.1-16(2023)
7. Artec Robo official website, (Accessed on November 25, 2023)
8. S. Long, T. Terakawa, M. Komori and T. Ougino: "Analysis of Traveling Strategies for Driving Omni-Wheeled Vehicle Around a Corner," in IEEE Access, vol. 8, pp. 104841-104856, (2020).
9. PV STUDIO website, (Accessed on November 13, 2023)
10. Ministry of Education, Culture, Sports, Science and Technology: Learning Innovation Project Demonstration Study Report (2014)

Authors Introduction

Mr. Kosei Machida

He e received his Bachelor's Degree (Education) in 2022 from the Faculty of Engineering, Toyko Gakugei University in Japan. He is currently a master student in Tokyo Gakugei University, Japan.

Dr. Shinichi Imai

He graduated doctor course at department of engineering in Hiroshima University. He works at department of education in Tokyo Gakugei University. His research area is about control system design, educational engineering.

A Feasibility Study on Methods to Measure the Strain on Young Children's Bodies.

Sachiko Kido
Interdisciplinary graduate school of Agriculture and Engineering, University of Miyazaki,,
1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan

Praveen Nuwantha Gunaratne
Interdisciplinary graduate school of Agriculture and Engineering, University of Miyazaki,,
1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan

Hiroki Tamura
Faculty of Engineering, University of Miyazaki,
1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan

E-mail: t321z01@student.miyazaki-u.ac.jp, ti20060@student.miyazaki-u.ac.jp, htamura@cc.miyazaki-u.ac.jp

Abstract

This paper investigates the effective methods of measuring the burden on the infant's body. Initially, the "the burden on the body and the lumbar spine burden are the same" was defined. The calculations and comparisons were carried out based on the two methods: the AnyBody Modeling System and the smartphone application Yo-bukun (a lumbar spine burden measurement application). The lumbar burden was calculated using the AnyBody Modelling System. In addition, an iPhone running the Yo-bukun lumbar burden measurement application was placed close to the heart position in order to simulate the lumbar burden. The results were compared with the respective average values and validated.

Keywords: Infants, Physical activity, Lunbar burden,

1. Introduction

In contemporary Japan, where the decline in children's physical fitness has emerged as a prominent concern [1], a comprehensive analysis of children's physical activity, particularly among young children, holds significant relevance. Various methodologies are employed to measure physical activity, encompassing continuous heart rate recording, electromyography, pedometers, calorie counters, and accelerometers [2], [3], [4]. Notably, smartwatches like the Apple Watch have gained popularity in recent years. However, the majority of these methods are tailored for adult use. Even among smartwatches, those designed for children are geared towards elementary school students and older, lacking compatibility with younger children. Accelerometers, pedometers, and behavior analysis are frequently employed for measuring physical activity in young children, yet a prevailing challenge in research utilizing these methods is their inadequacy for accommodating the needs of this demographic. Hence, there is an immediate necessity to investigate measurement methods that align with the body size of infants when assessing their physical activity.

The AnyBody Modeling System, developed at Aalborg University in Denmark, is a widely utilized musculoskeletal mechanics analysis software globally. The AnyBody Modeling System [5] employs inverse dynamics analysis to calculate forces acting on each part of the human body, including muscle activity, muscle and antagonist muscle forces, elastic energy of tendons, joint forces, and joint moments, during motion. While extensively used for simulating joint movements and analyzing motion capture data from video, research using this system has predominantly focused on adults, with limited exploration in the context of infants.

The Yo-bukun application [6], [7], a collaborative effort between the Faculty of Engineering at Miyazaki University and Densan Corporation in 2022, is an iPhone application quantifying and displaying lower back burden. This application is calibrated when the user stands upright, and calculates the amount of lumbar burden based on the tilt of the iPhone. This iPhone application facilitates the easy measurement of lower back burden and has not been previously applied to infants.

This paper investigates the effective methods of measuring the burden on the infant's body. The calculations and comparisons were carried out based on the two methods: the AnyBody Modeling System and the smartphone application Yo-bukun (a lumbar spine burden measurement application).

2. Methodology

In this paper, we will compare the results of the anybody modeling system and the "Yo-bukun" analysis application, which are set to the height and weight of infants. As a measurement system, the AnyBody

Sachiko Kido, Praveen Nuwantha Gunaratne, Hiroki Tamura

Modeling System is accurate if the movement is measured correctly. However, the AnyBody Modeling System is not realistic to accurately measure the burden of a moving infant. The "Yo-bukun" analysis application, who wears a smartphone, has fewer restrictions and can take measurements over a long period of time.

The value of the "Yo-bukun" analysis application is unclear how accurate for infants. Therefore, in order to confirm the accuracy of the "Yo-bukun" analysis application for infants, we conduct the comparative verification using the parameter settings as those for infants.

2.1. *The AnyBody Modeling System's analysis procedure*

In this paper, we use the AnyBody Modeling System to analyze the sample basic motion. Set the analysis settings in the AnyBody Modeling System to the height and weight of a typical infant, and calculate the amount of lumbar strain during movement in a pseudo manner.

2.2. *The Yo-bukun's analysis procedure*

Set the settings in the "Yo-bukun" analysis application to the same height and weight as in Chapter 2.1. We move by hand the smartphone according to the sample basic motion. In this way, we can simulate the basic motion on our smartphone.

3. Results and Discussion

3.1. *Analysis results with the AnyBody Modeling system*

Fig. 1 shows a series of balancing movements viewed from behind. The color of the muscles used changes, indicating areas of high strain. The fact that the lower back is colored in addition to the leg on which the body weight is placed indicates that the burden is placed mainly on the lower back. The posture with the maximum burden on the lower back is considered to be posture 5, in which the subject is off-balance. In the "balance" posture, the average value was 231[N] for approximately 8 seconds, and the maximum value was 283[N] (Fig. 1).

Fig. 2 shows a series of squatting movements, viewed from behind. In this movement, too, the color of the lumbar region changes, and it can be said that the lumbar region is burdened. The posture with the maximum lumbar load is posture 5 in the standing up motion. In the crouching movement, the average value for approximately 10 seconds was 322[N], and the maximum value was 484[N].

Fig. 1. Sequence of balance movements from behind with the AnyBody Modeling System analysis

Fig. 2. Sequence of squat movements from behind with the AnyBody Modeling System analysis

According to the AnyBody Modeling System analysis, two of the movements put the burden primarily on the lumbar region, thus meeting the definition of the "the burden on the body and the lumbar spine burden are the same".

3.2. *Analysis results with the Yo-bukun*

In the context of balance movements, the approximate average load over a duration of 8 seconds was 164[N], with a maximum load reaching 204[N].

In the squatting movement, the approximate average load over a duration of 10 seconds was 228[N], with a maximum load of 320[N].

3.3. *Comparison of Analytical Results between Two Software Applications*

In the Squatting movement, the average load was 322[N] in the AnyBody Modeling System and 228[N] in the Yo-bukun, resulting in AnyBody Modeling System exhibiting a slightly higher value of approximately 90[N]. The average value for balance movements was also 231 [N] for the AnyBody Modeling System and 164[N] for Yo-bukun, with the AnyBody Modeling System being about 70[N] higher. Although there was a slight discrepancy in the average values for both "Squatting" and "Balance," the fact that similar numerical values were obtained with both software applications. The burden on the body of infants during these two movements is approximately 200[N] to 350[N]. Although we predict that the difference between the two results is due to differences in the standards of the skeletal models used.

4. Conclusion

In this paper, the physical burden on infants was measured using two different software applications. A comparison of the average values obtained from the two software applications revealed no significant differences, suggesting that the burden on infants during the two movements can be considered to range from approximately 200[N] to 350[N]. The analysis time capability of the AnyBody Modeling System is approximately ten seconds. In contrast, the Yo-bukun enables prolonged analyses. The absence of significant differences between the AnyBody Modeling System and Yo-bukun in this study suggests the potential for measuring the physical load during infant body activities using Yo-bukun. In our future work, the validations are necessary by increasing the number of infant subjects.

References

1. Japan Sports Agency, "Results of the 2022 National Physical Fitness and Exercise Capacity and Exercise Habits Survey", https://www.mext.go.jp/sports/b_menu/toukei/kodomo/zencyo/1411922_00004.html, (Last viewed on December 15, 2023).
2. A. Sugimoto, "The Measurement of Physical Activity : A Review of Recent Progress", The Japanese Journal of Rehabilitation Medicine, vol. 37, 2000, pp. 53-61.
3. S. Tanaka, "A Study of the Relationship between Young Children's Motor Abilities and Physical Activities : Focusing Five Year Old Children's Physical Activity Countsina Day, The Hoikugaku Kenkyu vol. 47, no.2, 2009, pp. 8–16.
4. R. Tsuda, M. Watanabe and K. Suzuki, "Daily physical activities of public nursery school children during winter : a case study of eight public schools in Ishikawa Prefecture", the Japan Journal of Human Growth and Development Research, vol. 90, 2021, pp. 57-64.
5. M. Damsgaard, J. Rasmussen, S. T. Christensen, E. Surma and M. D. Zee, "Analysis of musculoskeletal systems in the AnyBody Modeling System", Simulation Modelling Practice and Theory, vol. 14, 2006, pp. 1100-1111.
6. H. Tamura, K. Sakurai, K. Tanno, and Y. Fuse, "A Study on the Lumbar Burden Evaluation of Work using One Smartphone," Journal of Robotics, Networking and Artificial Life, vol. 5, no. 3, pp. 173–173, Jan. 2018, doi: https://doi.org/10.2991/jrnal.2018.5.3.7.
7. H. Tamura, K. Sakurai, K. Tanno, and Y. Fuse, "A Study on the Lumbar Burden Evaluation of Work using One Smartphone," Proceedings of International Conference on Artificial Life and Robotics, vol. 23, pp. 550–553, Feb. 2018, doi: https://doi.org/10.5954/icarob.2018.os14-1

Authors Introduction

Ms. Sachiko Kido

She received her M.S. from Nakamura Gakuen University, Graduate School of Human Development in 2015. Currently enrolled in the doctoral program at the Graduate School of Agriculture and Engineering, University of Miyazaki. Assistant Professor, Department of Early Childhood Education, Higashikyushu Junior College, 2010. Lecturer, Faculty of Education, Miyazaki International University, since 2019. Her main research interest is the development of motor skills in young children.

Mr. Praveen Nuwantha Gunaratne

He received his Bachelor's degree in Engineering in 2018 from the Faculty of Engineering, University of Moratuwa, Sri Lanka. He is currently a Doctoral student in University of Miyazaki, Japan

Prof. Hiroki Tamura

He received the B.E. and M.E. degree from Miyazaki University in 1998 and 2000, respectively. From 2000 to 2001, he was an Engineer in Asahi Kasei Corporation, Japan. In 2001, he joined Toyama University, Toyama, Japan, where he was a Technical Official in the Department of Intellectual Information Systems. In 2006, he joined Miyazaki University, Miyazaki, Japan, where he was an Assistant Professor in the Department of Electrical and Electronic Engineering. Since 2015, he is currently a Professor in the Department of Environmental Robotics. His main research interests are Neural Networks and Optimization Problems. In recent years, he has had interest in Biomedical Signal Processing using Soft Computing.

An Integration of Contact Force Models with Multibody Dynamics Analyses for Human Joint Mechanisms and Effects of Viscoelastic Ground Contact

Shintaro Kasai
Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku, Kitakyushu, 808-0196, Japan

Dondogjamts Batbaatar
Mongolian University of Science and Technology, 8th khoroo, Baga toiruu 34, Sukhbaatar district Ulaanbaatar, Mongolia 14191

Hiroaki Wagatsuma
Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku, Kitakyushu, 808-0196, Japan
Email: kasai.shintaro660@mail.kyutech.jp, dondogjamts.b@must.edu.mn, waga@brain.kyutech.ac.jp

Abstract

In human movement and rehabilitation analyses, human joint dynamics is a key to consider the incorporation of spring-damper components, flexible bodies and contact forces analytically. In the present study, an analytical method for human gaits were introduced to integrate those essential elements, and viscoelastic properties of musculoskeletal system were modeled with the absolute nodal coordination formula (ANCF) method representing flexible body motions. A contact force model simulates interactions between different body segments and the environment. The proposed system is applied to a slider crank mechanism, demonstrating its capabilities in human joint motion analysis using integrated dynamic model within the framework of multibody dynamics (MBD), which realizes dynamic/inverse dynamics for human biomechanics.

Keywords: Multibody dynamics (MBD), Contact force model, Absolute nodal coordinate formula (ANCF), human biomechanics

1. Introduction

Joint disorders progress irreversibly due to exercise and aging, and the number of patients who have joint disorders is increasing. These joint disorders affect gait and daily life [1] and they are one of the major issues in the current aging society. Physical therapy treatments for joint disorders include step-by-step treatment that restricts extension by restraining joints using joint orthosis, and rehabilitation that applies appropriate joint loads. From the viewpoint of prevention of the disease by joint load reduction, patient-specific order-made assistive devices using 3D printers and flexible materials [2] are also developing. However, the formulation of rehabilitation programs and the design of orthotics and assist devices are performed through subjective evaluations based on the experience of physical therapists and prosthetists. Therefore, the standard numerical evaluation index is needed, and modeling human joint mechanisms and analyzing their movements plays an important role in the formulation of the programs and design for the devices.

The dynamic and inverse dynamic analysis of human motion is possible with multibody dynamics (MBD) [3], [4], [5], [6], [7]. In order to reproduce the movement of a joint and analyze in detail the increase/decrease in the load that occurs on its components, it is necessary to construct and analyze models in which the muscles and ligaments are replaced with spring damper components or flexible bodies. The absolute nodal coordinate formula (ANCF) [8], [9], [10] has been proposed as an analysis method that incorporates finite element methods for systems composed of flexible bodies in MBD. The ANCF allows the analysis of systems that include large deformation of flexible bodies by dividing the body into multiple elements. Furthermore, by considering the effects of viscoelastic ground contact [11], [12], it becomes possible to analyze human motion such as human gait.

In this study, for the implementation of an analytical system for the joint load reduction analysis based on MBD, we demonstrate an integrated dynamic model by applying a flexible body, spring-damper components and ground contact force to the crank model and analyzing it dynamically.

2. Methodology
2.1. *Integrated dynamic analysis based on MBD*

Shintaro Kasai, Dondogjamts Batbaatar, Hiroaki Wagatsuma

The motion of a multibody system (MBS) composed of rigid bodies is described by a differential algebraic equation (DAE) such as Eq. (1) based on MBD theory using Lagrange multiplier λ and acceleration equation γ. Here, M_R, Φ_R and Q_R^A mean the mass matrix of rigid bodies, constraint matrix and external force vector respectively. Subscript q means the Jacobian matrix obtained by partially differentiating the constraint matrix with respect to the generalized coordinate matrix q.

$$\begin{bmatrix} M_R & \Phi_{Rq}^T \\ \Phi_{Rq} & 0 \end{bmatrix} \begin{bmatrix} \ddot{q}_R \\ \lambda \end{bmatrix} = \begin{bmatrix} Q_R^A \\ \gamma \end{bmatrix} \qquad (1)$$

Considering incorporating a flexible body into MBS, the generalized coordinate matrix of the flexible body is defined and its motion is described based on ANCF. According to Fig. 1, the generalized coordinate matrix of the flexible body i is described as Eq. (2) with *x-y* coordinates at each node and their slope [13].

$$q_F^i = \left[e_1^i, e_2^i, e_3^i, e_4^i, \cdots e_{4N_n-3}^i, e_{4N_n-2}^i, e_{4N_n-1}^i, e_{4N_n}^i \right]^T \qquad (2)$$

The DAE of flexible multibody dynamics (fMBD) which incorporates flexible bodies based on ANCF can be expressed as Eq. (3) by expanding Eq. (1).

$$\begin{bmatrix} M_R & 0 & \Phi_{Rq}^T \\ 0 & M_F & \Phi_{Fq}^T \\ \Phi_{Fq} & \Phi_{Fq} & 0 \end{bmatrix} \begin{bmatrix} \ddot{q}_R \\ \ddot{q}_F \\ \lambda \end{bmatrix} = \begin{bmatrix} Q_R^A \\ Q_F^A \\ \gamma \end{bmatrix} \qquad (3)$$

Q_R^A and Q_F^A are external force vectors consist of gravity force and elastic force due to deformation of flexible body. When the analytical model consists of spring-damper components and the ground contact model, the total external force vector Q^A includes force vectors F_{SD} and F_c which are generated by spring-damper and viscoelastic ground contact as Eq. (4).

$$Q^A = \begin{bmatrix} Q_R^A \\ Q_F^A \end{bmatrix} + F_{SD} + F_c \qquad (4)$$

Fig. 1 The definition of the generalized coordinates of the flexible body I based on ANCF.

2.2. *Contact force in external force vector*

In the continuous contact model, the normal force F_N is expressed by Eq. (5) using the generalized stiffness parameter K and the relative penetration or indentation δ. Based on the theory of elasticity by Herts [11], the exponent n is equal to $2/3$.

$$F_N = K\delta^n \qquad (5)$$

Considering the damping term, F_N can be written as follows with hysteresis damping factor χ. Here, $c_r, \dot{\delta}$ and $\dot{\delta}^{(-)}$ mean coefficient of restitution, relative approach/departing velocity and initial approach velocity respectively.

$$F_N = K\delta^n + \chi\delta^n\dot{\delta} \qquad (6)$$

When χ is expressed as Eq. (7), substituting it into Eq. (6), F_N becomes Eq. (8).

$$\chi = \frac{3K(1 - c_r{}^2)}{4\dot{\delta}^{(-)}} \qquad (7)$$

$$F_N = K\delta^n \left[1 + \frac{3(1 - Cr^2)}{4} \cdot \frac{\dot{\delta}}{\dot{\delta}^{(-)}} \right] \qquad (8)$$

When sliding occurs in contact between the body and the ground, the dynamic friction force is described as Eq. (9) using kinetic friction coefficient μ, tangential velocity vector c_f and dynamic correction coefficient c_d based on friction velocity.

$$F_f = -\mu F_n c_f c_d \qquad (9)$$

Consider the case where body i contacts the straight line $ax + by + c = 0$ at contact point *CP* according to Fig. 2, the coordinate and velocity at point cp are as Eq. (10).

$$\begin{cases} \begin{bmatrix} x_{cp} \\ y_{cp} \end{bmatrix} = \begin{bmatrix} x_i \\ y_i \end{bmatrix} + \begin{bmatrix} r_i \\ 0 \end{bmatrix} \begin{bmatrix} \cos\theta_i & -\sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{bmatrix} \\ \begin{bmatrix} \dot{x}_{cp} \\ \dot{y}_{cp} \end{bmatrix} = \begin{bmatrix} \dot{x}_i \\ \dot{y}_i \end{bmatrix} - \begin{bmatrix} r_i \\ 0 \end{bmatrix} \dot{\theta}_i \end{cases} \qquad (10)$$

The relative penetration δ can be expressed as the shortest distance between the contact point and the straight line as Eq. (11).

$$\delta = \frac{|ax_{cp} + by_{cp} + c|}{\sqrt{a^2 + b^2}} \qquad (11)$$

Fig. 2 Viscoelastic contact with the tilted ground
described as $ax + by + c = 0$

Fig. 3 Integrated dynamic model of slider crank

The driving constraint Φ^D is given to L_1 as Eq. (14) and angular velocity ω is equal to $2\pi/5\ [rad/s]$.

$$\Phi^D = \theta_1 - \omega t - \theta_1^{(0)} = 0 \qquad (14)$$

In this slider-crank model, the endpoint at L_2 collides with L_3 and the effect of viscoelastic contact occurs without sliding. The contact force between these links is described as Eq. (15).

$$F_N = K(x_3 - x_2)^{\frac{2}{3}}\left[1 + \frac{3(1 - c_r^2)}{4} \cdot \frac{\dot{x}_3 - \dot{x}_2}{\dot{\delta}^{(-)}}\right]$$
$$F_f = 0 \qquad (15)$$

Since the relative approach velocity $\dot{\delta}$ in Eq. (8) and tangential velocity c_r in Eq. (9) are perpendicular and parallel to the straight line, they can be written as Eq. (12) when the ground line moves as $[\dot{x}_G, \dot{y}_G]^T$. Here the angle between x-axis and the straight line is as $\varphi = tan^{-1}(-a/b)$.

$$\dot{\delta} = (\dot{x}_G - \dot{x}_{cp})\sin\varphi - (\dot{y}_G - \dot{y}_{cp})\cos\varphi$$
$$c_r = (\dot{x}_G - \dot{x}_{cp})\cos\varphi + (\dot{y}_G - \dot{y}_{cp})\sin\varphi \qquad (12)$$

3. Results and Discussion

3.1. *Dynamic analysis of the integrated slider-crank model based on MBD*

Numerical analysis of the dynamic analysis based on MBD was performed by solving the DAE of the analytical system with MATLAB. In numerical computation, the fourth-order Runge-Kutta Gill's method [14] was used as the numerical integration method. The time step width is $h = 1.0 \times 10^{-5}[s]$. The parameters of spring-damper components are $k = 50, c = 0.1$ and the parameters of viscoelastic contact are $K = 4.0 \times 10^5, c_r = 0.1$ and $\mu = 0.45$. The results of the dynamic analysis for each element of the integrated slider-crank model are shown in Fig. 4. The x-coordinates of link L2 with and without flexible material are compared as shown in Fig. 5. The right graph in Fig. 5 shows the difference between each movement at first contact. When the model consists of a flexible body, the viscoelastic contact causes the deformation of it.

2.3. *Slider-crank model for validation of integrated analysis*

In order to validate the MBD integrated analysis, a flexible body, a spring-damper component and the viscoelastic contact model are applied to the slider-crank model as shown in Fig. 3 to compare the changes in the behavior of the mechanism with and without these components. When N_e which is the number of division elements of the elastic beam F_1 based on ANCF is equal to 8, the kinematic constraint equation between F_1 and rigid links L_1 and L_2 of the slider-crank in Fig. 4 is described by Eq. (13).

$$\begin{bmatrix}\Phi_{L_1F_1}^K \\ \Phi_{L_2F_1}^K\end{bmatrix} = \begin{bmatrix}e_1 - x_1 - \frac{l_1}{2}\cos\theta_1 \\ e_2 - y_1 - \frac{l_1}{2}\sin\theta_1 \\ e_{33} - x_2 + \frac{l_2}{2}\cos\theta_2 \\ e_{34} - y_2 + \frac{l_2}{2}\sin\theta_2\end{bmatrix} = 0 \qquad (13)$$

Shintaro Kasai, Dondogjamts Batbaatar, Hiroaki Wagatsuma

Fig. 4 The result of the dynamic analysis of the integrated slider-crank model.

Fig. 5 The transition of x-coordinate at CP. The right graph is looking at the point in time when the effect of viscoelastic contact occurres.

3.2. *Constraint force by adopting viscoelastic contact model*

MBD-based dynamic analysis was performed for three models with different components in Fig. 6 to analyze the changes in contact force on the body of the analytical system. Fig. 7 shows the result of the dynamic analysis, and the deflection of the flexible body and the contraction of the spring-damper component reduces the relative penetration of the contact point, resulting in smaller contact forces compared to the rigid slider-crank model. It is possible to analyze the force by the viscoelastic contact which changes according to the components applied to the MBS.

4. Conclusion

Through this research, the integrated dynamic analysis of models which has contacts between objects or objects and ground was shown to be possible by expanding the

Fig. 6 Integrated dynamic models of slider crank. The model a) consists of rigid linkages, b) is replaced and has a flexible body and c) is applied with a flexible body and a spring-damper component.

Fig. 7 The difference of contact force at contact point between the slider-crank models

analysis system to incorporate flexible bodies and a viscoelastic contact model. The use of flexible materials and spring-damper components in the joint mechanism enables gait analysis of walking robot legs that utilize nonlinear elastic properties and dynamic analysis of the effect of joint assistive devices on human walking motion, and is effective for the design and development of such robot legs and support devices. In further analysis, it is necessary to apply this integrated dynamic analysis to joint drive control for the robot leg models and the combination models of human gaits and assistive devices.

Acknowledgments

acknowledge ARIZONO Orthopedic Supplies Co., Ltd. for assistance with the formulation of exoskeleton-type assistive devices in their products.

References

1. L. Zhang, G. Liu, B. Han, Z. Wang, Y. Yan, J. Ma and P. Wei, "Knee Joint Biomechanics in Physiological Conditions and How Pathologies Can Affect It: A Systematic Review," Applied Bionics and Biomechanics, 2020, Article Number 7451683.
2. Y. J. Choo, M. Boudier-Revéret and M. C. Chang, "3D printing technology applied to orthosis manufacturing: narrative review," An Palliat Med 2020; 9(6): 4262-4270. doi: 10.21037/apm-20-1185
3. J. A. C. Ambrósio, Impact of Rigid and Flexible Multibody Systems: Deformation Description and Contact Models, in Virtual Nonlinear Multibody Systems. NATO ASI Series (Series II: Mathematics, Physics and Chemistry), eds. W. Schiehlen and M. Valášek, Springer, Dordrecht, vol 103, 2003, pp. 57-81.
4. P. E. Nikravesh, Planar Multibody Dynamics: Formulation, Programming with MATLAB, and Applications, 2nd edn., CRC Press, Boca Raton, 2018.
5. K. Komoda and H. Wagatsuma, Energy-efficacy comparisons and multibody dynamics analyses of legged robots with different closed-loop mechanisms, Multibody System Dynamics 40, 2017, pp. 123–153.
6. D. Batbaatar and H. Wagatsuma, A Proposal of the Kinematic Model of the Horse Leg Musculoskeletal System by Using Closed Linkages, Proceedings of the 2019 IEEE International Conference on Robotics and Biomimetics (ROBIO), Dali, China, 2019, pp. 869–874.
7. J. Baumgarte, Stabilization of constraints and integrals of motion in dynamical systems, Computer Methods in Applied Mechanics and Engineering 1(1), 1972, pp. 1–16.
8. A. A. Shabana, "Computer Implemantation of the Absolute Nodal Coordinate Formulation for Flexible Multibody Dynamics," Nonlinear Dynamics, vol. 16, pp. 293-306, 1998.
9. L. G. Maqueda, A. A. Mohamed and A. A. Shabana, "Use of General Nonlinear Material Models in Beam Problems: Application to Belts and Rubber Chains," Journal of Computational and Nonlinear Dynamics, vol. 5, pp. 21003-1-21003-10, 2010.
10. A. A. Shabana, An overview of the ANCF approach, justifications for its use, implementation issues, and future research directions, Multibody System Dynamics, 2023.
11. 11. Flores, P., Machado, M., Silva, M.T. et al. On the continuous contact force models for soft materials in multibody dynamics. Multibody Syst Dyn 25, 357–375 (2011). https://doi.org/10.1007/s11044-010-9237-4
12. D. Batbaatar and H. Wagatsuma, A Viscoelastic Contact Analysis of the Ground Reaction Force Differentiation in Walking and Running Gaits Realized in the Simplified Horse Leg Model Focusing on the Hoof-Ground Interaction, Journal of Robotics, Networking and Artificial Life, Vol. 8(2); September (2021), pp. 78-84.
13. M. A. Omar and A. A. Shabana, "A Two-Dimensional Shear Deformable Beam for Large Rotation and Deformation Problems," Journal of Sound and Vibration, vol. 243, no. 3, pp. 565-576, 2001. doi:10.1006/jsvi.2000.3416
14. Wolfram Research, Inc., Runge-Kutta Gill's method, https://mathworld.wolfram.com/GillsMethod.html

Authors Introduction

Mr. Shintaro Kasai

He received his Master's degree in Engineering in 2023 from the Graduate School of Life Science and System Engineering, Kyushu Institute of technology (Kyutech) in Japan,. He is currently a doctoral course student in Kyutech, Japan and JSPS Research Fellow (DC1).

Dr. Dondogjamts Batbaatar

He received his M.S. in the field of mechatronics from Mongolian University of Science and Technology (MUST), Mongolia and Ph.D. degree from Kyushu Institute of Technology, Japan in 2015 and 2021. He is currently a vice dean of research affair in School of Mechanical Engineering Transportation in MUST. His research interests include computational non-linear dynamics and bio-inspired robotics.

Dr. Hiroaki Wagatsuma

He received his M.S., and Ph.D. degrees from Tokyo Denki University, Japan, in 1997 and 2005, respectively. In 2009, he joined Kyushu Institute of Technology, where he is currently a Professor of the Department of Human Intelli- gence Systems. His research interests include non-linear dynamics and robotics. He is a member of IEEE.

Haptic Sensation Enhancement via the Stochastic Resonance Effect and Its Application to Haptic Feedback for Myoelectric Prosthetic Hands

Yoshitaka Mizumoto
Okayama University, 1-1-1 Tsushima-naka, Kita-ku, Okayama, 700-8530, Japan

Taro Shibanoki
Okayama University, 1-1-1 Tsushima-naka, Kita-ku, Okayama, 700-8530, Japan
Email:mizumoto-y@s.okayama-u.ac.jp, shibanoki@okayama-u.ac.jp

Abstract

In this paper, we propose a highly realistic haptic feedback method for myoelectric prosthetic hands. Although several haptic feedback methods have been studied, this study attempts to create realistic sensory feedback through prosthetic hands by not only using feedback methods but also by improving the haptic sensation based on the stochastic resonance effect. In the experiment, contact information obtained from a microphone attached to the fingertip of a prosthetic hand was transferred through a vibrotactile stimulator near the elbow fossa, and white noise vibration was also applied near the elbow fossa to verify the improvement of tactile sensitivity of the fingertips. The results demonstrated the possibility of transmitting tactile information for myoelectric prostheses through sensory enhancement of the user.

Keywords: Myoelectric Prosthetic Hand, Haptic Sensation, Stochastic Resonance, Haptic Feedback

1. Introduction

To realize a prosthetic hand that resembles a human hand, it is necessary not only to consider the mechanism of the prosthetic hand but also to establish a sensory feedback method, such as tactile sensation. Sensory information has a vital function in daily activities because the sense of touch conveys various properties or information of objects [1]. However, it is said that when prosthetic users touch objects with their prosthetic hands, they feel vibrations through the socket, making it difficult to accurately acquire tactile information through the prosthetic hand. Although myoelectric prosthetic hands have evolved to imitate motor functions of human hands [2], many current devices are deemed insufficient given their inability to deliver relevant sensory information.

The senses are broadly classified as cutaneous and deep senses, and restoration of these senses shows its improvement for the daily life of forearm amputees [3]. Various methods of cutaneous sensation feedback such as touch, pressure, pain, warmth and cold have been investigated. In the past, a simple transducer attached to a small speaker [4] was proposed, and in recent years, Ueda et al. have transmitted the warmth or coldness of an object by transmitting the temperature detected by a temperature sensor at the fingertip to a device on the upper arm [5], Osborn et al. developed electronic skin, which can transmit pain to the user by detecting differences in the shape and texture of objects at the fingertips and sending signals to peripheral nerves [6], and various other skin sensations can now be transmitted. Alessia et al. measured the acceleration of the fingertip and controlled the oscillator of a socket to enable discrimination of roughness [7], Christian et al. detected pressure from a silicon bulb and confirmed that the magnitude of pressure could be discriminated by inflating the silicon pad at the cut using a monofilament [8].

These various sensory feedback methods provide only pseudo feedback, and it is difficult to reproduce actual sensation. This is due to practical issues such as the danger of direct sensory feedback, including pain, and the fact that the cutaneous sensation near the amputation site is different from that at the fingertip, where feedback is provided in the noninvasive approach. Therefore, we consider improving the user's sensation.

Minamizawa et al. have made it possible to reproduce tactile sensations through vibration stimulation by recording and playing back tactile data [9]. Jianyao et al. measure the acceleration of the fingertip and transmit the vibration as an audio signal to enable roughness discrimination [10]. Research has also been conducted to enhance the sensory perception on the human side of receiving feedback, enhancing the ability to detect signals by applying noisy vibrations to the human hand or foot [11], [12]. This phenomenon is called stochastic resonance, and looking at the case of vibration to the upper limb, it has been reported that applying noise to the wrist and fingers improves tactile sensitivity at the fingertips [13],

Yoshitaka Mizumoto, Taro Shibanoki

[14]. However, the application of these sensory enhancements using audio feedback and stochastic resonance to prosthetic hands has not been fully investigated.

This study applies these methods to a prosthetic hand and provides tactile feedback near the elbow fossa using voice. In addition, by focusing on the phenomenon of probability resonance, we attempt to enhance the sensation obtained by the sensory feedback method of the prosthetic hand, thereby increasing the realism of the sensation through the prosthetic hand.

2. Sensory feedback and enhancement

The proposed myoelectric prosthetic hand sensory feedback method is based on the TECHTILE toolkit [9] and consists of a microphone (AT9904, audio-technica), an audio amplifier (LP-2024A+, Lepy), and vibration actuators (AFT14A903A, Alps Alpine). In this paper, a microphone is attached to the fingertip of a prosthetic hand to record friction and vibration sounds when touching an object as an input signal for vibration stimulus generation. The signal is output as vibration through an audio amplifier using an actuator attached near the elbow fossa.

In this study, we further focused on the stochastic resonance effect to improve sensory function. The vibration used is low-pass filtered white noise with a cutoff frequency of 300 Hz and is output from a PC via an audio amplifier with another actuator. These processes generate tactile sensations and simultaneously enhance the sensation near the elbow fossa, resulting reproduction of human fingertip sensation.

3. Verification of differences in fingertip sensation

In this section, we first discuss the difference between the cutaneous sensation at the fingertips of a person and the sensation when using a prosthetic hand. In this section, we first describe the difference between the cutaneous sensation at the fingertips of a person and the sensation when using a prosthetic hand. We compared and verified the sensitivity of the skin sensation at the fingertips when touching an object and the sensation of the prosthetic hand near the elbow through the socket.

The participant was a healthy university student seated on a chair wearing an eye mask and earplugs (Fig. 1) In the experiment, the participant was given a piece of paper with five pieces of sandpaper of different roughness (#400, #600, #800, #1000, and #1200.) on the front side and one randomly selected piece of sandpaper from the five pieces on the front side on the back side, and asked to touch all the sandpapers on the front and back sides. The sandpaper with the same coarseness as the one on the reverse side was selected from the sandpaper on the front side, and the percentage of correct responses was recorded. The test consisted of one trial in which each roughness was presented twice at random (10 in total). One trial was performed with the human hand and one trial with the prosthetic hand.

The discrimination rate of the human hand was 70%, and that of the prosthetic hand was 10%. These results

(a) Human hand   (b) Myoelectric prosthetic hand
Fig. 1. Scenes from the verification experiment.

indicate that fingertip sensation cannot be obtained through the socket of the prosthetic hand. From a questionnaire survey after the experiment, the participant answered that when they touched the sandpaper with their prosthetic hands, they felt almost no vibration through the sockets, suggesting that discrimination was impossible.

4. Sensory enhancement experiment

4.1 Determination of vibration perception threshold

To determine the vibration perception threshold (the maximum vibration that the participant cannot feel) before the experiment, the participant was asked to wear an eye mask, relax with hands on a desk, and present the vibration. The intensity of noise was increased and decreased until the participant was unable to distinguish between it being on or off. The average value was obtained by repeating each trial three times. The threshold is referred to as 1T.

4.2 Verification of the effect of stochastic resonance

We tested whether the tactile sensitivity was affected by applying white noise near the elbow fossa to one healthy university student. The participant was asked to touch a piece of sandpaper with his hand and a prosthetic hand (Fig. 2), and the percentage of correct responses for each was recorded. Five vibration intensity conditions (0.5 T, 0.75 T, 1 T, 1.25 T, and 1.5 T) were applied randomly. One trial was conducted for each vibration condition with a 30-second rest between trials. The hand and prosthetic hand were tested separately.

Fig. 3 shows the experimental results. The result for the no-vibration condition (no-vib) was obtained from the experiment in Section 3. In both cases, the correct response rate tended to be higher than that of the no-vibration. A questionnaire survey of the participant after the experiment showed that there was no significant difference in the perception of the vibration caused by the white noise.

Fig. 4 shows signals recorded from the microphone while touching sandpapers with the prosthetic hand and their FFT results. The threshold value was set to -30 dB, and the range from the first detection of a signal above the threshold value to the last detection of a signal above the threshold value was clipped. The extreme values of the absolute values of the signals were extracted and averaged to obtain the amplitude. The results show that the roughness of #400 tends to be easily discernible, with an

(a) Human hand   (b) Myoelectric prosthetic hand

Fig. 2. Scenes from the sensory enhancement experiment.

(a) Human hand   (b) Myoelectric prosthetic hand

Fig. 3. Results of the sensory enhancement experiment.

(a) Sound waveform

(b) FFT

Fig. 4. Recorded signals and their FFT results.

amplitude of approximately 0.021, which is larger than the others. In addition, #1000 tended to be misidentified as #800. This may be because the amplitude of #800 was about 0.012, close to the amplitude of #1000 of about 0.009, and the frequency band peaks were similar at around 900 Hz. The amplitude of #600 and #1200 were approximately 0.017 and 0.014, respectively. However, #600 had a peak near 2500 Hz and #1200 had a peak near 400 Hz, which may have affected the discriminability of the two samples. The identification rate through the socket of the myoelectric prosthetic hand was 30%.

(a) Experimental scene   (b) Experimental result

Fig. 5. Tactile feedback experiment.

5. Simulation of tactile generation near the elbow fossa

Although we were able to improve the tactile sensitivity of the fingertips by adding white noise near the elbow fossa in Section 4, the result of the prosthetic hand was low (30%). Therefore, we conducted an experiment in which vibration stimuli were applied near the elbow fossa based on the measured audio signals.

The participants were two healthy university students. White noise vibration was added simultaneously with feedback from the vibratory stimulation of the audio signal using two vibrotactile stimulators attached near the elbow fossa. The same procedure as in Section 3 was used to record the percentage of correct responses by touching the sandpaper with the prosthetic hand in five vibration intensity conditions.

The results of the experiment are shown in Fig. 5(b), which compares the cases without voice feedback (No-FB, the result shown as No-vib in Section 4), with tactile feedback only (No-vib), and with tactile feedback and white noise vibrations. The results show the average of two participants. Except for 0.75T, the result was higher than that without vibration, and the highest response rate was obtained when the vibration intensity was 1T. A questionnaire survey of the participants after the experiment suggested that the correct rate increased because the participants were able to feel strong vibrations due to the tactile feedback. These results indicate that some fingertip sensation can be obtained by vibration stimulation based on audio signals near the elbow fossa and tactile sensitivity enhancement by white noise vibration.

6. Verification of tactile sensation generation with a prosthetic hand

Based on the results obtained in Section 5, an experiment was conducted with a healthy university student as a participant to verify whether tactile feedback and white noise vibration could provide a realistic sensation when touching objects with a prosthetic hand. In the experiment, three different textures of cloth (nylon, polyester, and cotton) were traced with the prosthetic hand to test whether it could reproduce the sense of touch.

From the experimental results, when the participant touched the cloth with the prosthetic hand, the participant could not feel much difference in texture and could not reproduce the tactile sensation. Fig. 6 shows the sound waveforms and FFT results of touching the cloth with the

Yoshitaka Mizumoto, Taro Shibanoki

Fig. 6. Recorded signals and their FFT results.

prosthetic hand. The amplitudes of both results were almost the same (0.006) and the peak frequency bandwidths were similar at around 100 Hz, suggesting that the feedback vibration stimuli were similar and therefore could not be determined.

7. Conclusion

This paper showed that vibrotactile feedback and stochastic resonance effects near the elbow fossa can provide some fingertip sensation when touching an object through a prosthetic hand. However, it is difficult to distinguish objects with small amplitude or similar frequency bands. In the future, we plan to improve the feedback method to detect slight differences in vibration.

Acknowledgements

A Robust Approach for Reproducing the Haptic Sensation of Sandpaper With Different Roughness During Bare Fingertip Interaction.

References

1. Kairu Li, Yinfeng Fang, Yu Zhou, Honghai Liu, "Non-Invasive Stimulation-Based Tactile Sensation for Upper-Extremity Prosthesis: A Reviw", *IEEE Sens. J*, Vol. 17, No. 9(2017), pp. 2625-2635.
2. Larisa Dunai, Martin Novak, Carmen Garcia Espert, "Human Hand Anatomy-Based Prosthetic Hand", *Sensors 2021*, 21(1), 137.
3. Luis Vargas, He Huang, Yong Zhu, Xiaogang Hu, "Object Recognition via Evoked Sensory Feedback during Control of a Prosthetic Hand", *IEEE Robotics and Automation Letters*, Vol. 7, No. 1(2022), pp. 207-214.
4. John W. Hill, "A Describing Function Analysis of Tracking Performance Using Two Tactile Displays", *IEEE Transactions on Man-Machine Systems*, Vol. 11, No. 11(1970), pp. 92-101.
5. Yuki Ueda, Chiharu Ishii, "Development of a feedback device of temperature sensation for a myoelectric prosthetic hand by using Peltier element", *International Conference on Advanced Mechatronic Systems*, pp. 488-493, 2016.
6. Luke E.Osborn, Andrei Dragomir , Joseph L.Betthauser, Christopher L.Hunt, Harrison H.Nguyen, Ra-

hul R.Kaliki, Nitish V.Thakor, "Prosthesis with neuro-morphic multilayered edermis perceives touch and pain", *Sci. Robot.*, Vol. 3, No. 19 (2018), pp. 57-61.
7. Alessia Silvia Ivani, Federica Barontini, Manuel G. Catalano, Giorgio Grioli, Matteo Bianchi, Antonio Bicchi, "VIBES: Vibro-Inertial Bionic Enhancement System in a Prosthetic Socket", *2023 International Conference on Rehabilitation Robotics*, 2023.
8. Christian Antfolk, Anders Bjorkman, Sven-Olof Frank, Fredrik Sebelius, Goran Lundborg, Birgitta Rosen, "Sensory feedback from a prosthetic hand based on air-mediated pressure from the hand to the forearm skin", *Journal of rehabilitation medicine*, Vol. 44, No. 8(2012), pp. 702-707.
9. Kouta Minamizawa, Yasuaki Kakehi, Masashi Nakatani, Soichiro Mihara, Susumu Tachi, "TECHTILE toolkit: a prototyping tool for designing haptic media", *ACM SIGGRAPH 2012 Emerging Technologies*, pp. 387-392(2012).
10. Jianyao Zhang, Hiroyuki Kajimoto, "A Robust Approach for Reproducing the Haptic Sensation of Sandpaper With Different Roughness During Bare Fingertip Interaction", *ERONTIERS IN VIRTUAL REALITY*, Vol. 3(2022).
11. Marius Dettmer, Amir Pourmonghaddam, Beom-Chan Lee, Charles S Layne, "Effects of aging and tactile stochastic resonance on postural performance and postural control in a sensory conflict task", *Somatosens Mot Res*, Vol. 32, No. 2(2015), pp. 128-135.
12. Carly C. Sacco, Erin M. Gaffney, Jesse C. Dean, "Effects of White Noise Achilles Tendon Vibration on Quiet Standing and Active Postural Positioning", *J Appl Biomech,* Vol. 34, No. 2(2018), pp. 151-158.
13. Leah R Enders, Pilwon Hur, Michelle J Johnson, Na Jin Seo, "Remote vibrotactile noise improves light touch sensation in stroke survivors' fingertips via stochastic resonance", *Journal of NeuroEngineering and RehabilitationI*, 10:105(2013).
14. Komi Chamnongthai, Takahiro Endo, Shohei Ikemura, Fumitoshi Matsuno, "Stiffness Discrimination by Two Fingers with Stochastic Resonance", *International Conference on Human Haptic Sensing and Touch Enabled*, Springer, Cham, 2020.

Authors Introduction

Mr. Yoshitaka Mizumoto

He received the B.E. degree from Okayama University in 2022. He is currently a master course student in Okayama University His research interests focus on myoelectric prosthetic hand

Dr. Taro Shibanoki

He received the D.Eng. degree from Hiroshima University in 2012. He was a Research Fellow of the JSPS since 2011, an Assistant Professor in Hiroshima University since 2013 and an Assistant Professor, Lecturer, and an Associate Professor in Ibaraki University since 2014, respectively. He is currently an Associate Professor in Okayama University. His current research interests focus on human-machine interfaces.

PID Parameter Tuning of a Low-Cost DC Motor Speed Control for Mobile Robot Application

Munkh-Erdene Ayurzana

Mongolian University of Science and Technology, 8th khoroo, Baga toiruu 34, Sukhbaatar district Ulaanbaatar, Mongolia 14191

Erkhembayar Gankhuyag

Mongolian University of Science and Technology, 8th khoroo, Baga toiruu 34, Sukhbaatar district Ulaanbaatar, Mongolia 14191

Dondogjamts Batbaatar

Mongolian University of Science and Technology, 8th khoroo, Baga toiruu 34, Sukhbaatar district Ulaanbaatar, Mongolia 14191

Naranbaatar Erdenesuren

Mongolian University of Science and Technology, 8th khoroo, Baga toiruu 34, Sukhbaatar district Ulaanbaatar, Mongolia 14191

Email: B232340036@must.edu.mn

Abstract

Precise control of DC motors is vital for robotics, industrial automation, and mechatronics. Traditional PID (Proportional-Integral-Derivative) control methods, while widely used, often require offline parameter tuning, which can be time-consuming and suboptimal for real-time applications. This paper proposes a GUI (Graphical User Interface) based approach for PID tuning of a DC motor model. The proposed method utilizes a MATLAB-based parameter estimation model with GUI to continuously monitor and update PID controller parameters based on real-time data from the Arduino-controlled DC motor setup.

Keywords: PID parameter tuning, Graphical user interface (GUI), DC motor control

1. Introduction

Despite its simplicity, Proportional-Integral-Derivative (PID) control remains remarkably efficient, offering a three-term solution to address both transient and steady-state responses in numerous real-world problems. Since its invention in 1910, largely attributed to Elmer Sperry's ship autopilot, and the development of straightforward tuning methods by Ziegler and Nichols in 1942 [1], [2]. PID control has witnessed widespread adoption [3]. DC motors are widely employed as actuators for wheeled mobile robots, robots controlled by programmed microcontrollers, or computers that navigate autonomously based on sensor information. These robots typically have two, three, or four wheels, each driven by a DC motor for smooth maneuverability [4], [5], [6]. Two primary control applications for DC motors are speed and position control. Speed controllers regulate robot movement speed, while position controllers ensure accurate movement to predetermined positions. [7], [8]. Recent research on DC motor speed control focuses on stabilization, with many works combining PID controllers with other methods like fuzzy logic control, genetic algorithms [1],

and particle swarm optimization. Similarly, PID-based position controllers have been combined with artificial intelligence techniques like fuzzy logic and genetic algorithms to enhance accuracy [9], [10], [11]. While PID control exhibits remarkable effectiveness, optimal tuning remains a challenge, particularly in dynamic environments encountered by wheeled mobile robots. However, these approaches often lack user-friendliness and adaptivity.

This study utilizes one of the DC motors from a two-wheeled mobile robot as the plant. The DC motor's mathematical model is developed by acquiring input and output signals from an open-loop experiment [10]. The applied working voltage serves as the input signal, and the electronically measured speed is the output signal. The resulting model's step response is compared to the original motors to validate its accuracy. A PID-based speed controller is proposed, with initial parameters obtained from the relay feedback experiment and MATLAB based PID

Munkh-Erdene Ayurzana, Erkhembayar Gankhuyag, Dondogjamts Batbaatar, Naranbaatar Erdenersuren

tuning integrated with GUI. These parameters are then manually fine-tuned based on practical PID tuning knowledge to achieve optimal performance without overshoot or steady-state error. The paper is structured as follows: Section 1 presents the background and related work, Section 2 outlines the proposed approach and methodology, Section 3 details the implementation and results, and Section 4 concludes the study and highlights its potential impact.

2. Methodology

2.1. *Mathematical model of DC motor*

Before we create a GUI interface, we need to start off with a Simulink model of the DC motor system with PI controller. The DC motor has to be modeled first, in Simulink and the mathematical equations behind the functioning of the motor has to be understood before advancing   further

Fig.1 DC Motor model

In order to model a DC Motor in Simulink, we have to define the mathematical equations on which it functions, The DC Motor functioning diagram is given in Fig.1.

2.2. *Electrical Characteristics*

By using Kirchoff's voltage law in the DC-motor model according to Fig. 1 DC motor model following equation is derived

$$V_{in} - V_{Ra} - V_{La} - V_{emf} = 0 \qquad (1)$$

Where, V_{in}, V_{Ra}, V_{La} and V_{emf} is the input-voltage, voltage over the armature resistance, voltage drop over the armature inductance and voltage induced by the coil respectively. The equations for the voltage parts in the DC-motor are,

$$V_{Ra} = i_a \cdot R_a \qquad (2)$$

$$V_{La} = L_a \frac{d}{dt} i_a \qquad (3)$$

$$V_{emf} = K_e \cdot \dot{\theta}_a \qquad (4)$$

2.3. *Mechanical Characteristics* [2]

By torque balance (energy balance) in the system the mechanical equations can be stated.

$$T_e - T_{\ddot{\theta}a} - T_b - T_L = 0 \qquad (5)$$

Where T_e is the electromagnetic torque, $T_{\ddot{\theta}_a}$, is torque generated from the rotational acceleration of rotor. T_b, is the torque due to the friction and angular velocity in the motor. T_L is the torque of mechanical load (external load), equations for the first three parts are

$$T_e = K_t \cdot i_a \qquad (6)$$

$$T_{\ddot{\theta}_a} = J_m \cdot \frac{d}{dt} \dot{\theta}_a \qquad (7)$$

$$T_b = b_m \cdot \dot{\theta}_a \qquad (8)$$

Substituting equation (5) with (6), (7) and (8) yields following differential equation,

$$V_{in} - i_a \cdot R_a - L_a \cdot \frac{d}{dt} i_a - K_e \cdot \dot{\theta}_a = 0 \qquad (9)$$

$$\frac{d}{dt} i_a = \frac{1}{L_a} \cdot (i_a \cdot R_a + K_e \cdot \dot{\theta}_a - V_{in})$$

(10)

3. Results and Discussion

The experiment's foundation lay in meticulous data collection, enabling the analysis of the intricate interplay between voltage and RPM. A 140 second data acquisition process captured the voltage profile, characterized by a 20-second ramp-up to 24V followed by a 20-second minimum, and this pattern mirrored the observed RPM variations with relay feedback experiment was shown in Fig. 2.

Fig.2 Relay output

From the relay feedback experiment, the following result was obtained.

Fig.3 System without PID tuning

The open-loop experiment Fig. 3 revealed significant overshoot and sustained oscillations in the DC motor speed due to suboptimal PID settings. Manually configuring the parameters or using the time-consuming Ziegler-Nichols method might not be ideal for real-time applications.

3.1 *Parameter estimation and tuning*

From the mathematical model of DC motor following initial parameters are to be known such as moment of inertia of the rotor J_m, damping ratio of the mechanical system b_m, electric resistance R_m, electric inductance L_m, armature constant K_t, before being tuned with a PID controller. While manufacturers may provide values for some of these DC motor parameters, they are only approximate. Its necessary to estimate these parameters as precisely as possible for our model to ascertain whether it is an accurate representation of the actual DC motor system.

Fig.4 Parameter estimation plot

Again, these values are of important for designing and tuning a PID controller for the DC motor. The PID controller will use these parameters to calculate the motor's input voltage in order to achieve the desired speed. Furthermore, without knowing the specific values of the parameters, it is difficult to say much about the specific performance of the DC motor. However, the fact that these are estimated parameters (Fig. 5) suggests that the motor has not been precisely modeled or

characterized [12], [13]. This could lead to difficulties in tuning the PID controller effectively.

Fig. 5 Iteration for finding motor parameters

As shown in Fig. 4 the measured data overlaid with the simulated data. The simulated data comes from the model with the estimated parameters listed in Table 1. Comparing the response of the system before and after the estimation process clearly shows that the estimation successfully identified the model parameters and the simulated response accurately matches the experimental data.

Table 1. DC Motor parameters

Constants	Value	Unit
K_t	0.0035401	$[N \cdot m/A]$
R_m	0.0065388	$[Ohm]$
L_m	8.758e-05	$[H]$
b_m	0.0001263	$[-]$
J_m	4.3114e-05	$[kg \cdot m^2]$

While from PID tuning result following initial PID parameters are determined in Table 2.

Table 2. Parameters used in the numerical simulation.

Controller	PID Parameters		
	K_p	K_i	K_d
PID	0.01	0.061	0

After the fine-tuning process in MATLAB, it was found that the final $K_p = 0.01$, $K_i = 0.061$ and $K_d = 0$. The initial value of Kp = 0.9 gives big oscillation output. The fine tuning is performed to adjust the K_p, such that the P controller output results in only one overshoot and no oscillation, although steady state error is still not zero. This final value of Kp is 0.01.

Fig.6 System with PID tuning

The Fig. 6 illustrates the response of a DC motor speed control performance after it has been tuned with PID controller. The blue line shows the reference input, the red line shows the actual output speed of the motor, which initially lags behind the reference input and then abruptly rise to meet it. There are some oscillations in the motor speed, which indicates that the system is damped due external disturbances or externally given mechanical load. The system is returned to stable state approximately after 1 second under the varying operating conditions.

3.2. GUI for motor control

In order to simplify the dc motor control process, a GUI is developed that allows a user to quickly identify model, tune controller coefficients and send reference data was created. Two types of GUIs were created.

   a)   GUI for identification

Fig.7 Identification GUI MATLAB

This GUI (presented on Fig. 7) allows user to control motor speed manually, send identification data, identify model, tune PID, set object constants and deploy model with new coefficients and constants to Arduino. After uploading a model to Arduino, user can check the quality of tuned controller by reconnecting to Arduino, setting reference velocity manually and checking the response. The quality of controlling depends on several factors.

First of all, getting a proper data is one of the most important steps. Based on the collected data, picking the correct model is directly affects the PID tuning response and therefore directly affects the possibilities of tuning. At last, figuring the optimal response through MATLAB PID tuner directly affects the quality of controlling.

   b)   GUI for standalone application

Once identification and tests are complete, user can shut down MATLAB and use a standalone application as shown in Fig. 8 to control the DC motor. Standalone application requires less resources to run compared to multistep manual identification.

Fig.8 Speed controller performance at final stage with standalone application

In Fig. 8 graph depicts the reference tracking response of the DC motor after it has been tuned with the PID controller. The blue line shows the reference input, which is direction of motor changes with time. The red line shows the actual output of a motor, which closely tracks the reference input with minimal overshoot, and settling time, and there is no steady-state error. Finally, an analysis was performed to ensure that the controller was well-tuned with GUI of standalone application.

4.   Conclusion

This project developed an interactive GUI to optimize PID controller parameters for a DC motor SIMULINK model in real-time. This approach enables continuous adaptation to varying operating conditions, achieving significant improvements compared to traditional offline or manual tuning methods. Experimental results demonstrate enhanced control accuracy and system stability, confirming the effectiveness of the proposed method. The PID controller, adjusted through the GUI, optimizes the motor's input voltage to minimize the error between the desired and actual output. This is achieved by utilizing a parameter estimation tool based on the DC motor's mathematical model. Careful tuning of the PID gains ensures a system that is both responsive and stable. The developed GUI serves as a valuable tool for robot designers and researchers, enabling real-time control optimization and hands-on learning of control systems concepts. It offers undergraduate students a practical learning experience in the classroom, fostering their understanding of PID control and its applications.

Acknowledgements

References

1. Astrom, K. J. and T., Hagglund, PID Controllers: Theory, Design and Tuning, ISA, Research Triangle, Par, NC, (1995)
2. S. K. Suman and V. K. Giri, "Speed control of DC motor using optimization techniques based PID Controller," *2016 IEEE International Conference on Engineering and Technology (ICETECH)*, Coimbatore, India, 2016, pp. 581-587
3. Visioli, Antonio. (2012). Research Trends for PID Controllers. Acta Polytechnica. 52. 144-150.
4. Supriyo, B., Suharjono, A., & Atmaja, K. W. (2021). Modelling and Control of DC Motor Speed and Position for Wheel Mobile Robot Application. *Proceedings of the International Conference on Innovation in Science and Technology (ICIST 2020)*, 80–84.
5. Fu, Yuheng & Zhou, Qinyou. (2023). Analysis and application research of mobile robot navigation related technologies. Applied and Computational Engineering. 9. 92-96.
6. Khan, H.; Khatoon, S.; Gaur, P.; Abbas, M.; Saleel, C.A.; Khan, S.A. Speed Control of Wheeled Mobile Robot by Nature-Inspired Social Spider Algorithm-Based PID Controller. *Processes* **2023**, *11*, 1202.
7. K. Ogata (2009) "Modern Control Motorering, 5th Edition", Dorling Kindersley,".
8. Kiam Heong Ang, G. Chong and Yun Li, "PID control system analysis, design, and technology," in *IEEE Transactions on Control Systems Technology*, vol. 13, no. 4, pp. 559-576, July 2005
9. Vikhe, Pratap & Punjabi, Neelam & Kadu, C B. (2015). DC Motor Speed Control Using PID Controller In Lab View. IJISME. 3. 38-41.
10. Vikhe, Pratap & Punjabi, Neelam & Kadu, C B. (2014). Real Time DC Motor Speed Control using PID Controller in LabVIEW. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering. 03. 12162-12167.
11. Messaadi, Nassim & Amroun, Abdelkader. (2021). Speed Control of DC Motor Using Fuzzy PID Controller.
12. Nilawar, Vaishnavi & Sonare, Aniket & Samrutwar, Shrirup. (2022). Implement PID using MATLAB for motion control application.
13. Abdul-Kareem Mohammed, J. (2011). Modeling, Analysis and Speed Control Design Methods of a DC Motor. *Engineering and Technology Journal*.

Authors Introduction

Mr. Munkh-Erdene Ayurzana

He is currently studying for his Bachelor's degree in Mechanical Engineering at the Mongolian University of Science and Technology (MUST). His research interests include control system and robotics.

Mr. Erkhembayar Gankhuyag

He is currently studying for his Bachelor's degree in Mechatronics Engineering at the Mongolian University of Science and Technology (MUST). His research interests include control system and mechatronics.

Dr. Dondogjamts Batbaatar

He received his M.S. in the field of mechatronics from Mongolian University of Science and Technology (MUST), Mongolia and Ph.D. degree from Kyushu Institute of Technology, Japan in 2015 and 2021. He is currently a vice dean of research affair in School of Mechanical Engineering Transportation in MUST. His research interests include computational non-linear dynamics and bio-inspired robotics.

Dr. Erdenesuren Naranbaatar

He received his M.S. in the field of mechatronics from Mongolian University of Science and Technology (MUST), Mongolia and Ph.D. degree from University of Ulsan, South Korea in 2002 and 2013. he is currently an Associate Professor of the Graduate department in School of Mechanical Engineering Transportation in MUST. His research interests include computer vision and robotics.

Reinforcement Learning DDPG Algorithm Based Wheeled Mobility Aid Robot Control Methods

Junkai Li, Mohd Rizon Mohamad Juhari, Tiang Sew Sun

Faculty of Engineering, Technology and Built Environment, UCSI University, Kuala Lumpur 56000, Malaysia

E-mail: 1002162416@ucsiuniversity.edu.my, mohdrizon@ucsiuniversity.edu.my, tiangss@ucsiuniversity.edu.my
www. ucsiuniversity.edu.my

Abstract

When using wheeled walker robots to help individuals with limited walking ability improve their mobility, the stability of the robot's motion control and trajectory tracking accuracy are critical. In this paper, a new trajectory tracking method for wheeled walking robots is proposed by combining the Deep Deterministic Policy Gradient (DDPG) algorithm in reinforcement learning with a Proportional Integral Differential (PID) controller. The article first analyzes the kinematic model of the chassis of the wheeled walking robot, and then introduces the design principle and structure of the adaptive PID controller that combines the DDPG algorithm and the PID controller. Finally, the effectiveness of the research scheme and control strategy is verified by joint simulation experiments, and the results show that this DDPG-based PID controller can automatically adjust the parameters when tracking the trajectory to ensure the accuracy of the trajectory, and it has a strong anti-interference capability.

Keywords: Wheeled helper robot, trajectory tracking, DDPG algorithm, PID controller.

1. Introduction

In recent years, with increasing aging, the elderly show significant physiological decline, limb flexibility and other basic abilities, increasing the risk of falls. Meanwhile, as people's lifestyles change, there are more and more lower limb motor dysfunctions caused by sports injuries, traffic accidents, or diseases, making there a huge demand for intelligent assistive devices [1]. Walking robots for the elderly and patients with lower limb motor dysfunction can be categorized into two types: exoskeleton robots [2] and wheeled assisted walking robots [3]. However, the center of gravity control of most exoskeleton rehabilitation robots is not satisfactory, and the safety of the rehabilitation process cannot meet the requirements [4]. Wheeled walking robot itself has a stable chassis, not only can well avoid the patient in the training process due to the center of gravity is not stable and caused by the fall and other secondary injuries, but also in a narrow space to achieve any direction of movement, to give the patient a good sense of spatial movement and real walking experience, wheeled walking robot with its stable chassis and flexible spatial movement ability, in terms of safety and practicality with the Advantages.

When using wheeled walking robots for rehabilitation training, due to the complexity of the external environment, the wheeled walking robots are required to follow the desired trajectory as much as possible, which requires the robots to be able to track an ideal trajectory with a time function according to the actual position and motion state, and complete the tracking of each point in the trajectory according to the time requirements Therefore, this paper applies the reinforcement learning algorithm to the traditional proportional-integral-derivative (PID) controller to realize adaptive PID

parameter tuning in order to adjust the robot position more accurately so that the robot can complete trajectory tracking.

Although the traditional PID control has the advantages of simplicity of use, easy to implement, no static error, etc., its disadvantage is that it cannot realize the online adjustment of parameters, so when it encounters a strong interference, it is bound to have the phenomenon of prolonged recovery time and increased overshooting, which affects the stability of the motion of the chassis of wheeled robots. A large number of scholars have conducted research on adaptive PID, and introduced the idea of online parameter adjustment in the traditional PID, which improves the response speed of the system. At present, adaptive PID control methods mainly include: fuzzy PID controller [5], the method requires a large amount of a priori knowledge, there are parameter optimization problems; neural network-based adaptive PID control [6], the method can be achieved without identifying the complex nonlinear system to achieve effective control, but in the use of supervised learning to optimize the parameters, the acquisition of the teacher's signals is relatively difficult; evolutionary algorithm adaptive PID controller [7], the method, although the acquisition of a priori knowledge of the lower requirements, but the computation time is longer, it is difficult to achieve real-time control in practical applications; reinforcement learning adaptive PID controller [8], proposed the Actor-Critic algorithm to achieve adaptive tuning of PID parameters, which makes use of AC algorithms of model-free online learning The algorithm utilizes the model-free online learning capability of AC algorithm, but the convergence speed of AC algorithm is slow and the training time is long.

In this paper, we adopts the deep deterministic policy gradient (DDPG) algorithm, which is based on the Actor-Critic (AC) framework, to enhance the performance of the PID controller and the deep Q-network (DQN) algorithm is added on the basis of the deterministic policy gradient (DPG) algorithm. This algorithm can not only update in a single step like DQN, but also has the advantages of high data utilization and fast convergence of DPG. In order to realize the adaptive adjustment of PID parameters, a PID controller based on DDPG algorithm is proposed. In the simulation experiments, the kinematic model of the omnidirectional chassis of the wheeled walking robot in our laboratory is used to verify the superiority and generality of the proposed method.

2. Proposed Method

This section will introduce the kinematic model of the wheeled walking robot chassis. Then, it will describe the design principles and structure of the adaptive PID controller that combines the DDPG algorithm and the PID controller.

2.1. *Kinematic modeling of assistive robots*

The schematic and kinematic model of the robot's omnidirectional wheel chassis is shown in Fig. 1. Taking the world coordinate system $x_w o_w y_w$ as a reference, the robot moves in the motion coordinate system xoy with a velocity of magnitude v toward an angle α from the Y_w axis. V_x, V_y denote the horizontal and vertical travel speeds of the robot relative to the motion coordinate system, respectively [9].

Fig. 1 Kinematic modeling of assistive robots

When the robot is moving in any direction, e.g., at an angle α from the y -direction with velocity v , the magnitudes of the horizontal and vertical velocities of the robot respectively:

$$
\begin{aligned}
V_x &= v sin\alpha \\
V_y &= v cos\alpha
\end{aligned}
\tag{1}
$$

Let the angle between the omnidirectional wheels and the x-axis, measured from the center of the robot chassis, be θ ,and the distance from the center of the chassis to the four wheels be a , then the wheel speed of each omnidirectional wheel is degree is:

$$
\begin{aligned}
V_1 &= -sin\theta V_x + cos\theta V_y + \dot{\theta}a \\
V_2 &= \ \ sin\theta V_x + cos\theta V_y - \dot{\theta}a \\
V_3 &= \ \ sin\theta V_x + cos\theta V_y - \dot{\theta}a \\
V_4 &= -sin\theta V_x + cos\theta V_y + \dot{\theta}a
\end{aligned}
\tag{2}
$$

Where V_1, V_2, V_3, V_4 denote the linear velocities of the four omni-directional wheels respectively and $\dot{\theta}$ denotes the angular velocity of the robot writing Eq. (1) in matrix form gives:

$$\begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \end{bmatrix} = \begin{bmatrix} -sin\theta & cos\theta & a \\ cos\theta & sin\theta & a \\ -sin\theta & cos\theta & a \\ cos\theta & sin\theta & a \end{bmatrix} \qquad (3)$$

Let the transformation matrix R be:

$$R = \begin{bmatrix} -sin\theta & cos\theta & a \\ cos\theta & sin\theta & a \\ -sin\theta & cos\theta & a \\ cos\theta & sin\theta & a \end{bmatrix} \qquad (4)$$

Let the positional attitude of the robot at $Y_w O_w X_w$ be $P = [x, y, \theta]^T$, and the differential of this positional attitude be $\dot{P}$, and the velocities of each of the robot's wheels $V_b = [V_1, V_2, V_3, V_4,]$, then the equations of kinematics of the robot's differential omnidirectional wheels are:

$$\dot{P} = R^{-1} \cdot V_b \qquad (5)$$

2.2. Incremental PID Control Principle

Digital PID control can be categorized into two types: positional PID and incremental PID. incremental PID does not require the use of the accumulated value of past deviations, which effectively reduces the system's computational error. Incremental PID is an algorithm that realizes PID control by controlling the increment. The formula is as follows:

$$u(t) = u(t-1) + K_i(t)e(t) + K_p\Delta e(t) + K_d(t)\Delta^2 e(t) \quad (6)$$

$$\begin{aligned} e(t) &= y_d(t) - y(t) \\ \Delta e(t) &= e(t) - e(t-1) \\ \Delta^2 e(t) &= e(t) - 2e(t-1) + e(t-2) \end{aligned} \qquad (7)$$

In these equations, $y_d(t)$ denotes the current actual signal value, $y(t)$ denotes the current system output value of the current system, $e(t)$ denotes the output error of the system, $\Delta e(t)$ denotes the first difference of the error, and $\Delta^2 e(t)$ denotes the second difference of the error.

The use of incremental PID controllers in control system design can optimize the use of computational resources. The method relies only on the last three sampling points to determine the control increments, thus reducing the computational burden and storage requirements. This facilitates fast training of the DDPG algorithm and storage of sample data. In addition, the incremental PID controller ensures stable operation in the event of a system failure, and since its output is the changing value of the control quantity, it makes the rewards obtained in a reinforcement learning

environment more stable, which in turn accelerates the convergence of the algorithm.

In summary, the incremental PID algorithm is used to realize the trajectory tracking control of the wheeled robot, which requires the design of two PID controllers, i.e., the transverse position X controller and the longitudinal position Y controller, and the block diagram of the control system is shown in Fig. 2. The input of the transverse PID controller is x_e, i.e., the transverse deviation of the wheeled robot system; the output is v_x, the transverse speed of the wheeled robot system. the output of the PID controller goes to the speed distribution controller, which calculates the speeds of the four omnidirectional wheels according to Eq. (2) to correct the transverse deviation of the wheeled robot. The principle of longitudinal position Y control is the same as the principle of transverse position X control, i.e., y_e is the longitudinal error, v_y is the longitudinal velocity.

Fig.2 Block diagram of two PID controller systems

2.3. DDPG algorithm

Reinforcement learning is a machine learning method that learns behavioral strategies through trial and error, while deep learning is a machine learning method that performs high-level abstraction and feature extraction through multi-layer neural networks. Deep Reinforcement Learning combines the feature extraction capability of Deep Learning with the decision-making capability of Reinforcement Learning to autonomously learn and improve the performance of decision-making strategies. DDPG (Deep Deterministic Policy Gradient) algorithm is an optimization algorithm based on the Actor-Critic algorithm [5] framework, which is able to better select the optimal policy in continuous actions such as robot movement. The DDPG algorithm is based on deterministic policy gradient and refers to the experience pool replay mechanism and Double-Depth Q-Network (DDQN) objective network method to update the network parameters and realize the self-tuning of PID parameters, and its algorithm structure is shown in Fig. 3.

Fig.3. Block diagram of DDPG algorithm structure

Actor gives a new parameter μ according to the strategy network, denoted as $a_t = \mu(S_t|\theta^\mu) + N_t$, where θ^μ denotes the parameter of the neural network, G_k denotes the noise.After one update iteration, the state changes from S_t to S_{t+1} , and obtains the reward R_t. (S_t, A_t, S_{t+1}, R_t) is stored in the experience pool. Finally, samples are drawn from the experience pool to train the strategy network and evaluation network.

The strategy network and evaluation network are updated as shown in Fig. 4. The actor and the Critic consist of two identical networks, denoted as: actor evaluation network $\mu(s|\theta^\mu)$, actor estimation network $\mu' = (s_t|\theta^{\mu'})$, Critic evaluation network $Q(s, a|\theta^Q)$, Critic estimation network $Q'(s, a|\theta^{Q'})$.

Fig.4. updating network parameters

The actor evaluation network updates the parameters according to the following objective function:

$$\mu = \frac{1}{N} \sum_t \left(\nabla_a Q(s, a|\theta^Q)|_{S=s_t, A=\mu(S_t)} \cdot \nabla_{\theta^\mu} \mu(s|\theta^\mu)|_{S=s_t} \right) \quad (8)$$

where $\nabla_a Q(s, a|\theta^Q)$ is the gradient of the Critic evaluation network, and the gradient is updated by computing the loss function. The loss function is computed as follows:

$$L = \frac{1}{N} \sum_t \left(y_t - Q(s_t, a_t|\theta^Q) \right)^2 \quad (9)$$

where y_t is referred to as the TD target, which consists of an immediate reward and an estimated discounted reward. The formula is as follows:

$$y_t = r_t + \gamma Q'(s_{t+1}, \mu'(s_{t+1}|\theta^{\mu'})|\theta^{Q'}) \quad (10)$$

Then update the Critic evaluation network according to $\nabla_a Q(s, a|\theta^Q)$, and at the same time pass $\nabla_a Q(s, a|\theta^Q)$ to the actor, update the actor evaluation network according to Eq. (9), and finally update the parameters of the target network as follows:

$$\begin{cases} \theta^{Q'} \leftarrow \tau\theta^Q + (1-\tau)\theta^{Q'} \\ \theta^{\mu'} \leftarrow \tau\theta^\mu + (1-\tau)\theta^{\mu'} \end{cases} \quad (11)$$

2.4. *General Structure of Adaptive PID Controller Based on DDPG Algorithm*

The design idea is to combine the DDPG algorithm on the basis of incremental PID controller, so the structure design is shown in Fig. 5.

Fig.5. adaptive PID controller based on DDPG algorithm

The system block diagram is based on the traditional PID with the addition of closed-loop control by the DDPG algorithm. The speed allocation module allocates the speeds of the four wheels according to Eq. (2), and the scores are obtained through a self-designed reward function, and output to the DDPG algorithm. According

to the state obtained by the system output, the DDPG algorithm continuously carries out trial and error training until it selects the action group with the highest score, which is the optimal PID parameter, and outputs it to the two PID controllers, adaptively adjusts the parameter, and the error under the parameter is minimized, and the parameter outputs it to the speed allocation module, and then outputs the speeds of the four wheels to the wheeled walking robots' chassis system, which realizes the trajectory tracking control. The controller uses the parameter vector of the PID controllers, i.e. the 6 parameters to be tuned by the two PID controllers, as the action space.

Critics rate each reinforcement learning cycle using the system's state variables and the defined reward function R_t to generate the TD error $\delta_{TD}(t)$ and evaluate the value function Q_t, where $\delta_{TD}(t)$ is provided directly to the actor and Critic, and the reward R_t is used to evaluate the quality of the current behavior.

$$\begin{cases} R_1 = -0.00001 \\ R_2 = -10, e(t) \geq 1 \times 10^{-5} \\ R_3 = 0, e(t) < 1 \times 10^{-5} \\ R_4 = -200, e(t) > 0.5 \\ R_t = R_1 + R_2 + R_3 + R_4 \end{cases} \quad (12)$$

R_1 is the reward for the speed of the four wheels, which is set to a score that has less impact on the system since the main reward comes from the error rather than the amount of control. R_2 and R_3 denote the scores when the error is lower or higher than the error tolerance interval, respectively, and R_4 denotes the score of -200 when the system error is more than 0.5. Finally, the four scores are summed up to form the final reward function. Based on the continuous trial-and-error training, the maximum reward value is obtained, and thus the optimal PID parameter values are obtained.

3. Simulation experiment

In this paper, trajectory tracking control simulation experiments are carried out in python environment. Based on the chassis modeling of the wheeled walking robot, the system simulation model was established in Simulink. In the simulation process, the initial position of the controlled object is set as [0,1], the desired trajectory is set as a cosine curve along the x-axis, and the sampling time is set as 1ms. The parameters of the adaptive PID control based on the DDPG algorithm are listed in Table 1 and Table 2.

Table 1. DDPG Algorithm Parameters

Parameter	Value
Actor learning rate	0.005
Critic learning rate	0.001
Discount rate	0.99
Soft update parameter	0.001
Total training steps	50000
Experience pool capacity	1000000

Table 2. PID controller Parameters

Parameter	Value
K_{p1}	5.0
K_{p2}	5.0
K_{i1}	2.0
K_{i2}	2.0
K_{d1}	0.2
K_{d2}	0.2

4. Result

Experiment shows the simulation results of trajectory tracking realized by traditional PID controller in the Fig. 6. The solid line shows the desired trajectory and the dashed line shows the tracked trajectory. It can be seen that the method can also realize the trajectory tracking, the overshoot is 4.5%, and the tracking trajectory is far away from the target trajectory, and the error is large. Fig. 7 shows the simulation results of trajectory tracking control using the PID controller based on DDPG algorithm proposed in this paper. The overshoot of this method is 2.7%, and the trajectory is closer to the target trajectory with less error than the traditional PID control.

Fig.6 Traditional PID for trajectory tracking

Junkai Li, Mohd Rizon Mohamad Juhari, Tiang Sew Sun

Fig.7 PID control for trajectory tracking based on DDPG algorithm

Compared with the traditional PID controller, the maximum error of this controller is reduced by 48.12%, as listed in Table 3. In addition, the controller's overshoot is reduced by 40% compared to a traditional PID controller, which improves user safety during assisted walking and training.

Table 3. Comparison of results

Controller	Maximum error	Overshoot
Traditional PID	0.2097	4.5%
DDPG+PID	0.1088	2.7%

Comparative results show that the controller is able to track the desired trajectory more accurately when the user uses the wheeled mobility aid robot for rehabilitation training. It also has a strong anti-interference ability due to the trial-and-error mechanism of reinforcement learning, which increases the safety and comfort of training.

5. Conclusion

In order to ensure that the elderly and patients with lower limb dysfunction can move accurately according to the pre-set desired trajectory when they use wheeled walking robots for rehabilitation training in complex environments, this paper proposes a reinforcement learning method that combines the DDPG algorithm with a PID controller. This method not only solves the problem that the traditional PID cannot adjust the parameters online, but also reduces the storage space requirement of the controlled system, thus reducing the computation time of the system. Simulation results show that the PID control based on the DDPG algorithm has the advantages of high tracking accuracy, small overshooting amount and strong adaptability, which enables the wheeled walking robot to realize accurate

trajectory tracking control. The method has good generality and generalization.

References

1. Hou Zeng-Guang, Zhao Xin-Gang, Cheng Long, Wang Qi-Ning, Wang Wei-Qun. Recent Advances in Rehabilitation Robots and Intelligent Assistance Systems. ACTA AUTOMATICA SINICA, 2016
2. Zhou, Jinman, Shuo Yang, and Qiang Xue. "Lower limb rehabilitation exoskeleton robot: A review." Advances in Mechanical Engineering 13.4 ,2021
3. Neuhaus, Peter, and H. Kazerooni. "Design and control of human assisted walking robot." Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No. 00CH37065). Vol. 1. IEEE, 2000
4. Nokata M, Ikuta K, Ishii H. 11 Safety Evaluation Method of Rehabilitation Robots. In: BIEN Z Z, STEFANOV D. Advances in Rehabilitation Robotics[M]. Lecture Notes in Control and Information Science, 2006, 306: 187-198.
5. Savran A. A Multivariable Predictive Fuzzy PID Control System [J]. Applied Soft Computing,2013, 13(5): 2658-2667.
6. Chen J H, Huang T C. Applying Neural Networks to on-line Updated PID Controllers for Nonlinear Process Control [J]. Journal of Process Control,2004, 14(2): 211-230.
7. Saad M S, Jamaluddin H, Darus I Z M. Implementation of PID controller tuning using differential evolution and genetic algorithms [J]. International Journal of Innovative Computing, Information and Control, 2012, 8(11): 7761-7779.
8. Pomerleau A, Desbiens A, Hodouin D. Development and evaluation of an auto-tuning and adaptive PID controller [J]. Automatica, 1996, 32(1): 71-82.
9. Mollaret C, Mekonnen A A, Lerasle F, et al. A multi-modal perception based assistive robotic system for the elderly [J]. Computer Vision and Image Understanding, 2016, 149: 78-97.

Authors Introduction

Mr. Junkai Li

He is currently pursuing Doctor of Philosophy (Engineering) in Faculty of Engineering, Technology and Built Environment, UCSI University, Malaysia. His research interests are machine learning, optimization algorithm and optical measurement.

Prof. Dr. Mohd Rizon Mohamad Juhari

He is a Professor in Faculty of Engineering at UCSI University in Malaysia. He received his PhD in Engineering from Oita University, Japen in 2002. His research interests are face analysis, pattern recognition and vision for mobile robot.

Assistant Professor Ir Ts Dr Tiang Sew Sun

She is an Assistant Professor in Faculty of Engineering at UCSI University in Malaysia. She received her PhD in Electrical and Electronic Engineering from Universiti Sains Malaysia in 2014. Her research interests are optimization and antenna design.

Unsupervised image registration based on Residual-connected DRMINE for diagnostic metastatic bone tumors

Shogo Baba[1], Tohru Kamiya[1], Takashi Terasawa[2], Takatoshi Aoki[2]

[1] *Kyushu Institute of Technology, 1-1 Sensui, Tobata, Kitakyushu, Fukuoka 804-8550, Japan*
[2] *University of Occupational and Environmental Health, 1-1, Iseigaoka, Yahatanishi, Kitakyushu, Fukuoka, 807-8555, Japan*
Email: baba.shogo274@mail.kyutech.jp, kamiya@cntl.kyutech.ac.jp

Abstract

Metastatic tumors are frequently identified through follow-up surveillance using computed tomography (CT) scans. However, CT scans produce more than 100 images in an examination, which imposes a significant burden on radiologists and entails a potential risk of misdiagnosis. Temporal subtraction is utilized in Computer-Aided Diagnosis (CAD) and proves to be an effective technique in aiding image interpretation process for the radiologists. In this study, we focus on the preliminary stage of CAD development specialized in bone metastasis extraction, with a particular emphasis on rigid registration. We propose a novel rigid registration technique by augmenting DRMINE, which estimates mutual information using neural networks, with skip connections and normalization. From the three datasets, ten images were selected randomly from the cervical, thoracic, and lumbar regions. These images were then augmented with rotation as well as horizontal and vertical translations to create modified versions. The registration accuracy was assessed based on the Full Width at Half Maximum (FWHM) of the difference images. In the proposed method, FWHM values for the thoracic and lumbar regions of the spine exhibited a maximum reduction rate of 2.8% and a minimum reduction rate of 0.533%. However, the cervical spine region exhibited superior FWHM results with DRMINE compared to the proposed methodology. The proposed method was influenced by the capture area, but it indicated the potential to provide stable registration as the standard deviation decreased for all FWHM values.

Keywords: Computed Tomography, Rigid registration, Unsupervised learning, MINE, DRMINE

1. Introduction

As per the estimates of the World Health Organization (WHO) in 2019, cancer is the first or second leading cause of death in 112 of 183 countries [1]. Furthermore, there exists the possibility of cancer metastasizing to distant locations through the vascular and lymphatic systems, with the spine, in particular, being the most susceptible site for skeletal metastasis [2], [3]. Metastatic bone tumors necessitate prompt identification and timely intervention. However, as metastatic bone tumors commonly lack discernible symptoms, they are often detected through Computed Tomography(CT) examinations during follow-up observations. Therefore, the assessment of osseous structures harboring metastases and the early detection thereof predominantly rely upon CT scanning as the most efficacious modality. However, in CT examinations, where 100 to several hundred cross-sectional images can be obtained in an examination, the burden on the diagnosing radiologist increases significantly, and the results vary due to differences in interpretation experience, which poses a significant challenge.

The Temporal Subtraction technique (TS) [4] is a method that involves performing calculations between two images taken at different times to extract lesions that appeared during a specific period. This technique is applied in computer-aided diagnosis (CAD) systems for supporting radiographic interpretation. However, the mere subtraction of a simple current image and a past image results in the occurrence of differential artifacts due to positional discrepancies between the images. Therefore, it becomes crucial to perform registration processing as a preprocessing step. In typical image registration frameworks, prior to performing non-rigid registration, the target image pairs are often pre-aligned based on rigid registration or affine transformations. However, unless there is a significant improvement in the accuracy of affine deformations, the performance of alignment will be greatly diminished [5], [6].

In recent years, numerous studies have been conducted on registration techniques, particularly in the domain of deep learning. However, deep learning requires a substantial amount of annotated data for training, which hampers the practical application of deep learning methods in medical research. Consequently, unsupervised learning, which does not rely on annotated data, has garnered significant attention as a potential solution to this limitation. Differentiable registration with mutual information and matrix exponential (DRMINE) [7] is an unsupervised learning-based rigid registration model that utilizes Mutual Information Neural Estimation (MINE) [8] to estimate mutual information using a neural network, instead of relying on conventional approximations of mutual information.

Shogo Baba, Tohru Kamiya, Takashi Terasawa, Takatoshi Aoki

In this study, we focused on rigid registration and proposed a model incorporated with skip connections and normalization in the MINEnet framework within the DRMINE model.

2. Materials and methods
2.1. Dataset

We conducted experiments employing TOSHIBA Aquilion PRIME and TOSHIBA Aquilion ONE, utilizing data from three datasets encompassing imaging ranges spanning from the cervical to lumbar spine levels. The image size was set at 512×512, with a slice thickness of 1.0 mm and pixel size ranging from 0.625 to 0.820 mm. We conducted experiments by randomly selecting ten images each from the cervical, thoracic, and lumbar regions based on the data of these three datasets cases.

Furthermore, in order to replicate the displacement in positioning, an image was created with horizontal and vertical movements of 15 pixels each along the x-axis and y-axis, along with a 10° rotation (Bicubic interpolation). Fig. 1 illustrates the image transformation, depicting the image before and after the conversion.

Our computing system consisted of an Intel Core i7-7500U CPU and an NVIDIA Tesla V100 GPU with 16GB of memory.

Fig. 1 The image transformation
Left: Original image, Right: transformed image

2.2. DRMINE registration algorithm

MINE utilizes the principle of Donsker-Varadhan (DV) duality to compute mutual information (MI),

$$MI = \sup_{f} J(f) \tag{1}$$

where $J(f)$ is the DV lower bound.

$$J(f) = \int f(x,z)P_{XZ}(x,z)dxdz$$
$$- \log(\int exp(f(x,z))P_X(x)P_Z(z)dxdz) \tag{2}$$

MINE uses a neural network to compute $f(x,z)$ and uses Monte Carlo technique to approximate the right hand side of (2). MINE claims that computation of (1) scales much better than histogram-based computation of MI.

The optimization of DRMINE is as follows.

$$\max_{\substack{v_1,\cdots,v_6 \\ v_1^1,\cdots v_6^1 \\ \theta}} \{ \sum_{l=2}^{L} MINE(T_l, Warp(M_l, M_{exp}(\sum_{i=1}^{6} v_i B_i)))$$
$$+ MINE(T_1, Warp(M_1, M_{exp}(\sum_{i=1}^{6}(v_i + v_i^1)B_i)))\} \tag{3}$$

B_i is an affine transformation on a two-dimensional plane. This group comprises six generators, where $Mexp$ denotes the matrix exponential operation.

$$B_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} B_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} B_3 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$B_4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} B_5 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} B_6 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Using the multi-resolution pyramid technique [9], [10], [11], the fix image pyramid $:T_l$ and the moving image pyramid $:M_l$ $(l = 1, \dots, L)$. L is the maximum level in the image pyramid.

MINE is a method for estimating mutual information utilizing neural network architectures. MINE learns using the error backpropagation method, similar to conventional neural networks, in order to maximize the mutual information between the fixed image and the distorted moving image. θ denotes the parameters of the neural network utilized to accomplish MINE. MINEnet is constructed with two hidden layers composed of 100 neurons each, and utilizes the non-linear activation function ReLu between the hidden layers.

However, note also that image structures are slightly shifted through multi-resolution image pyramids. So, a transformation matrix suitable for a coarse resolution may need a slight correction when used for a finer resolution. To alleviate this issue, in the second item, we employ the parameterization of matrix exponential functions and introduce a dedicated additional parameter vector $v^1 = [v_1^1, \cdots, v_6^1]$ for the finest level of resolution, optimizing the multiresolution approach. Fig. 2 illustrates DRMINE registration algorithm.

Fig. 2 DRMINE algorithm

2.3. Our method

In order to enhance the registration technique utilizing DRMINE, it is of paramount importance to precisely estimate mutual information. MINEnet is a method that utilizes neural networks for the estimate mutual information, and thus, we postulated that the effective propagation of differentials plays a pivotal role in the estimate mutual information. The conventional approach comprises a simplistic two-layer structure, whereas the proposed methodology constructs a three-layer architecture by incorporating additional layers into MINEnet. Furthermore, we introduced skip connections

[12] between each layer. Additionally, we believed that the normalization of the hidden layers also influences the precision of the neural network. Hence, we incorporated layer normalization [13] into the model, as depicted in Fig. 3. Moreover, we set the number of neurons in each layer to 100, similar to DRMINE.

To demonstrate the enhancement in accuracy attributed to the incorporation of normalization, this study conducted a comparative analysis of precision among three methodologies: the conventional approach, a model enriched solely with skip connections in the DRMINE framework (Method [a]), and the proposed technique (Method [b]).

2.4. Evaluation methods

The evaluation of image registration accuracy involved the computation of the Full Width at Half Maximum (FWHM) from the histogram of the difference image. This difference image was aligned using a reference image and registered with the moving image (Fig. 4). FWHM represents the width at which the distribution encompasses half of its maximum value. A smaller FWHM indicates a reduction in the presence of artifacts attributed to residuals.

3. Experiments and results
3.1. Results

From Table 1, the method[b] yielded the lowest average FWHM for both thoracic and lumbar regions of all data of datasets. The mean reduction rate was found to be 1.389%, with a maximum reduction rate of 2.8% and a minimum reduction rate of 0.294%. However, in the cervical spine region, DRMINE yielded superior precision results across all data of datasets. The model constructed solely with skip connections without normalization did not show a clear improvement in accuracy compared to DRMINE. Therefore, the enhancement of accuracy cannot be achieved solely through the addition of skip connections. By combining Layer Normalization, we have achieved an enhancement in the accuracy of estimations.

Fig. 5, Fig. 6, Fig. 7 depict the results obtained from algorithms applied to the cervical, thoracic, and lumbar regions. The proposed methodology led to the observable reduction of residuals in the contour region.

Fig. 3 Proposed network

Fig. 4 FWHM (Subtract image histogram)

Table1. Average FWHM and Rate of increase or decrease
Rate[a]: The comparison between DRMINE and Method[a], Rate[b]: The comparison between DRMINE and Method[b]
▲: Rate of increase , ▼: Rate of decrease

Datasets	Level	DRMINE	Method[a]	Method[b]	Rate[a] (%)	Rate[b] (%)
Data[A]	C-spine	1.2686±0.0182	1.2805±0.0177	1.2712±0.0151	0.928▲	0.200▲
	T-spine	1.2251±0.0249	1.2065±0.0196	1.1918±0.0140	1.543▼	2.800▼
	L-spine	1.1812±0.0237	1.1794±0.0194	1.1633±0.0136	0.158▼	1.546▼
Data[B]	C-spine	1.3628±0.0224	1.3724±0.0193	1.3667±0.0116	0.699▲	0.286▲
	T-spine	1.3955±0.0197	1.3984±0.0111	1.3819±0.0127	0.254▼	0.981▼
	L-spine	1.3187±0.0252	1.3153±0.0279	1.2958±0.0115	0.207▲	1.764▼
Data[C]	C-spine	1.2645±0.0180	1.2688±0.0200	1.2693±0.0166	0.344▲	0.381▲
	T-spine	1.3194±0.0147	1.3140±0.0121	1.3101±0.0044	0.026▲	0.712▼
	L-spine	1.2680±0.0106	1.2684±0.0154	1.2613±0.0067	0.417▼	0.533▼

Shogo Baba, Tohru Kamiya, Takashi Terasawa, Takatoshi Aoki

Fig.5 Cervical spine region

Fig. 6 Thoracic spine region

Fig. 7 Lumbar spine region

4. Discussion and conclusion

From Table 1, it can be inferred that the contour's shape exerts an influence on the FWHM. However, it is posited that the stratification of the proposed methodology was rather shallow, resulting in an insufficient learning efficacy. Therefore, we contemplate that further deepening the layers and establishing interconnections between them through skip connections may lead to an improvement in precision. Furthermore, the mere addition of skip connections did not exert any influence on the enhancement of accuracy and stabilization of estimation precision. This demonstrates the contribution of adding layer normalization not only to the improvement of accuracy but also to the stabilization of estimate mutual information.

The proposed methodology has demonstrated further enhancement and stabilization of precision in rigid registration. Additionally, it has suggested its utility in rigid registration of medical images from an unsupervised learning perspective.

References

1. OECD, Health at a Glance 2021: OECD indicators, pp.143, OECD publishing, 2022.
2. David Stegner et al., "Mechanistic Explanation for Platelet Contribution to Cancer Metastasis", Thrombosis Research, Vol.133, No.52, pp.149-157, 2014.
3. R. Guillevin et al., "Spine Metastasis Imaging: Review of the Literature", Journal of Neuroradiology, Vol. 34, No.5, pp.311-321, 2007.
4. Kano A. et al., "Digital image subtraction of temporally sequential chest images for detection of interval change", Medical Physics, Vol.21, No.3, pp.453-461, 1994.
5. Maintz et al., "A survey of medical image registration. Medical image analysis", Vol.2, No.1, pp.1-36, 1998.
6. Pluim et al., "Mutual-information-based registration of medical images: a survey", IEEE Transactions on Medical Imaging, Vol.22, No.8, pp.986-1004. 2003.
7. Nan A. et al., "Drmime: Differentiable mutual information and matrix exponential for multi-resolution image registration", In: Medical Imaging with Deep Learning. PMLR, pp.527-543, 2020.
8. Belghazi et al., "Mutual information neural estimation., In: International conference on machine learning", PMLR, pp.531-540. 2018.
9. P. Thevenaz, et al., "A pyramid approach to subpixel registration based on intensity", IEEE Transactions on Image Processing, Vol.7, No.1, pp.27-41, 1998.
10. S. Kruger et al., "Image registration using multiresolution frequency domain correlation", in BMVC, pp.1-10, 1998.
11. H. S. Alhichri et al., "Multi-resolution image registration using multi-class Hausdorff fraction", Pattern Recognition Letters, Vol.23, No.1-3, pp.279-286, 2002.
12. He K. et al., "Deep residual learning for image recognition", In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770-778, 2016.
13. Ba Jimmy Lei et al., "Layer normalization", arXiv:1607.06450, 2016.

Authors Introduction

Mr. Shogo Baba

He received his Master degree in Medical Science from Kagoshima University Japan in 2021. He is a Ph.D. student at Kyushu Institute of Technology Japan.

Tohru Kamiya, Prof. Ph.D

He received his B.A. degree in Electrical Engineering from Kyushu Institute of Technology in 1994, the Masters and Ph.D. degree from Kyushu Institute of Technology in 1996 and 2001, respectively. He is a professor in the Department of Mechanical and Control Engineering at Kyushu Institute of Technology. His research interests are focused on image processing, medical application of image analysis.

MD. Takashi Terasawa

He received his MD degree from university of occupational and environmental health japan in 2011 and PhD in medicine in 2020. He is a radiologist at the university.

MD. Takatoshi Aoki

He is a professor of Radiology at University of Occupational & Environmental Health in Japan, and the vice president of Japanese Musculoskeletal Society. His clinical and research interests include the imaging modalities of musculoskeletal pathology, especially of tumor, arthritis, and bone metabolism.

Developing a Smart Belt for Monitoring Elderly Activities Based on Multi-Modal Sensors Integration and Internet of Things

Abdul Jalil*, Pujianti Wahyuningsih, Najirah Umar, Muhammad Risal, Suwatri Jura, A. Edeth Fuari Anatasya

Universitas Handayani Makassar,
Adhyaksa Baru Street, No. 1, Makassar, South Sulawesi, 90231, Indonesia.
Email: abdul.jalil@handayani.ac.id, ujiwahyuningsih@handayani.ac.id, najirah@handayani.ac.id,
risal@handayani.ac.id, suwatrijura@handayani.ac.id, edetanatasya@handayani.ac.id.

Abstract

This study aims to develop a smart belt that can be used to monitor elderly activities at home by integrating Multi-Modal Sensors (MMS) and the Internet of Things (IoT). The MMS used in this study to collect elderly information is the IMU sensor, vibration sensor, push button, and ESP32 to process IoT data. Furthermore, the IoT platform used to transmit elderly information data from the smart belt to the family smartphone is Blynk. In this study, the smart belt can monitor the activities of the elderly when walking, sitting, lying down, and sleeping. After that, the smart belt can provide warning alerts when the elderly carry out abnormal activities at home. The results of this study show that the smart belt can monitor the activities of the elderly when walking, sitting, lying down, sleeping, and warning alerts on the family smartphone based on the integration of MMS and IoT.

Keywords: Smart Belt, Elderly, Multi-Modal Sensor, Internet of Things.

1. Introduction

The elderly is a person who has reached the age of 60 years or older [1]. During this phase, the elderly will have a reduced ability to perform activities, which means that their family will need to provide them with special assistance. In addition to monitoring the condition of the elderly at home, the busy activities of family members outside the home can pose an additional problem when caring for an elderly relative who is alone at home. Therefore, the urgency of this research is to build a smart belt device that can help family members monitor their elderly activities based on the integration of Multi-Modal Sensors (MMS) and the Internet of Things (IoT).

Some researchers have conducted studies on the development of devices to monitor elderly activities based on IoT technology. Hua et al. have developed a device that can be used to monitor elderly activities using ICE (IoT Cares for the Elderly) [2]. In that study, the researchers used a heartbeat sensor, a body temperature sensor, and an Intel Edison platform to monitor elderly activities based on the Internet of Things. Narasinghe et al. have conducted a study to monitor and detect activities of elderly care by integrating various technologies of wearable and non-wearable devices connected to the wireless network [3]. The researchers in that study used the heart rate sensor and the PIR sensor to detect elderly activities and the cloud system to store elderly information using an internet connection. Cheng et al. have presented an application to monitor elderly care with IoT [4]. In that study, the researchers used the ThingSpeak IoT platform and the ADXL345 accelerometer to monitor the elderly while doing activities at home. Rupasinghe et al. conducted the design and development of an IoT-based device to track the physical activities of the elderly using an accelerometer sensor [5]. Firebase IoT platform was used to transmit elderly information from the device to the smartphone of the elderly family. Furthermore, Naeim et al. have developed a device for monitoring elderly activities for safety based on mobile IoT [6]. In that study, the researchers used a low-cost prototype device to measure heart pulse, detect falls, and determine the location of the elderly, then used Blynk, Firebase, and Google Assistant for the IoT platform.

According to the results of the study shown previously, it can be seen that some researchers have developed devices to monitor elderly activities using IoT. Therefore, the contribution of this study is to develop a smart belt that can be used to monitor elderly activities based on the integration of MMS and IoT. In this study, we used the IMU sensor, vibration sensor, and push-button for MMS, and Blynk as an IoT platform to transmit elderly information from the smart belt to elderly activities when the elderly is walking, sitting, lying down, and sleeping.

This paper consists of a fourth chapter to elaborate on our study results. In the second chapter, we explain the method used in our study to develop a smart belt to monitor elderly activities. The third chapter presents the results of our study. Then in the last chapter, we show the conclusion and future work of our study.

2. Methodology

Figure 1 shows the architecture design developed in our study to build the smart belt based on the integration of MMS and IoT to monitor elderly activities.

Abdul Jalil, Pujianti Wahyuningsih, Najirah Umar, Muhammad Risal, Suwatri Jura, A. Edeh Fuari Anatasya

Fig.1. Architecture design

According to the information in Fig. 1, it can be seen that the sensor devices used to develop MMS in our study consist of the IMU sensor, vibration sensor, and push button. The function of the IMU sensor in this study is to detect the activities of the elderly when they walk, sit, lie down, and sleep. We detected it when the degree of motion of the IMU sensor was observed following the activities of the elderly. IMU sensor is a sensor device that can be used to detect the degree of motion of an object according to acceleration and gyroscope [7]. The vibration sensor function in the smart belt is to confirm that the elderly is moving when the IMU sensor detects that the elderly walk, sit, and lie down. Then after that, when the IMU sensor detected the elderly in a sleep position and the vibration sensor did not detect the elderly moving, the system confirmed that the elderly was in a sleep state. The vibration sensor is a sensor device that can detect the vibration of an object based on the mechanical quantity received by the sensor and then convert it into electrical current [8]. Furthermore, the push button function is a security button that sends elderly information when the elderly need helps or are in an abnormal condition.

In this study, the microcontroller used to read and process MMS and IoT data is ESP32. ESP32 is a microcontroller that can be used to process sensor data for the controller, is low-cost and low-power, and is integrated with the Wi-Fi and Bluetooth module on a board [9]. The IoT platform used to transmit IoT data between the smart belt and the smartphone in this study is Blynk. Blynk is an IoT platform that can be used to communicate IoT data between devices to an iOS or Android smartphone over an Internet connection [10]. Furthermore, Fig. 2 shows the flowchart system developed in our study to operate the smart belt based on IoT.

Fig.2. Flowchart systems

Based on the information in Fig. 2, it can be seen that, in the first step, the smart belt reads the data input received from the IMU sensor, vibration sensor, and push button, then processes the IMU and vibration sensor data to detect elderly activities. When the system detects that the elderly is walking, the system then sends the information from the smart belt to the family's smartphone to inform the elderly that they are walking using Blynk IoT, connected over the Internet connection.

Furthermore, when the elderly is sitting, lying down, and sleeping, the system sends the information data to the family's smartphone based on the information from the IMU and vibration sensor, respectively. When the system detects that the push button is pressed, the Blynk IoT then sends the warning information to the family that the elderly need helps or are in an abnormal condition. The systems then run repeatedly until the power on the smart belt is turned off.

3. Results and Discussion

We have developed a smart belt that can be used to monitor the activities of the elderly based on the integration of MMS and IoT. Figure 3 shows the smart belt developed in our study.

Fig.3 The smart belt developed in this study.

Based on the information in Fig. 3, to implement the experiment, we used an IMU sensor (KKHMF, MPU-6050) to detect elderly activities when walking, sitting, lying down, and sleeping according to the degree of movement of the IMU sensor (acceleration and gyroscope), a vibration sensor (DKARDU, Piezoelectric) to confirm elderly activities when walking, sitting, lying down, and sleeping according to the vibration of elderly moving, then push button as trigger for the smart belt to send a warning to the family when the elderly is an abnormal condition. The microcontroller used in our study to process MMS data and IoT is ESP32 (ESP32 DEVKIT V1) and then used the Blynk platform to transmit IoT data between the smart belt and the family smartphone. Furthermore, Table 1 shows the DATASTREAM ID and data type used to transmit the elderly information from the smart belt to the family smartphone using Blynk.

Table 1. The DATASTREAM ID and data type are used to transmit elderly information.

Elderly Information	DATASTREAM	Data Type
Walking	V1	INT, 0/1
Sitting	V2	INT, 0/1
Lying Down	V3	INT, 0/1
Sleeping	V4	INT, 0/1
Warning	V5	INT, 0/1

According to the information in Table 1, we use INT data type 1 or 0 to confirm the activities of the elderly when they walk, sit, lie down, sleep, and have abnormal conditions. When the system detected that the elderly was doing activities (walking, sitting, lying down, sleeping, and abnormal conditions), the system then sent the INT data of 1 to confirm that the elderly was doing the activities. If not, the system sends 0 data information to the family smartphone. Furthermore, Fig. 4 shows the

activities of the elderly when walking, standing up, sitting, and lying down, respectively. Fig. 5 shows the elderly information on the family smartphone using the Blynk application based on the Android OS when the elderly walk.

Fig.4. Activities of the elderly when (a) walking, (b) standing, (c) sitting, and (d) lying down.

Fig.5. Information of elderly when walking.

4. Conclusion

In this study, the development of a smart belt based on the integration of MMS and IoT that can be used to monitor elderly activities when walking, sitting, lying down, sleeping, and under abnormal conditions was carried out. We used an IMU sensor, vibration sensor, and push button for MMS and ESP32 as a microcontroller to process MMS and IoT data, then Blynk for the IoT platform. In our study result, the smart belt can detect the elderly walking, sitting, lying down, and sleeping on the family smartphone.

References

1. J. P. Putri and J. Natalia, The Dynamics of Aging Process Adaption From the Late Adulthood to the Elderly in Panti Werdha X. Psikostudia: Jurnal Psikologi, Vol. 11, No. 3, 2022, pp. 467-478.
2. B.D.C. Hua, H. Fahmi, L. Yuhao, C.C. Kiong and A.Harun, Internet of Things (IoT) Monitoring System for Elderly, International Conference on Intelligent and Advance System (ICIAS), 2018.
3. N. Narasinghe, RPS. Kathiarachchi and M.W.P. Maduranga, IoT based Health Monitoring and Activity Detection for Elderly Care, 13th International Research Conference General Sir John Kotelawala Defense University, 2020, pp. 396–401.
4. B.J. Cheng, M.M.A. Jamil and R. Ambar, Elderly Care Monitoring System with IoT Application, ICITAM SCI 863, 2020, pp. 525–537.
5. I.D.M.S. Rupasinghe and M.W.P. Maduranga, Towards Ambient Assisted Living (AAL): Design of an IoT-based Elderly Activity Monitoring System, I.J. Engineering and Manufacturing, 12(2), 2022, pp. 1–10.
6. M.K.M. Naeim, G.C. Chung, I.E. Lee, J.J. Tiang and S.F. Tan, A Mobile IoT-based Elderly Monitoring System for Senior Safety, International Journal of Technology, Vol. 14(6), 2023, pp. 1185–1195.
7. W. Suryanto, Supriyanto, E. Hartantyo, T. Juhana, Budiarjo, A.S. Handaru and J. Satriani, IMU (Inertial Measurement Unit) Device for Internet of Things Based Disaster Early Warning System: Application and Innovation, IOP Conf. Series: Earth and Enviromental Science, 2020.
8. D.N. Ilham, R.A. Candra, A. Budiansyah, E. Sipahutar, M.K. Harahap and F. Anugreni, Implementation of Vibration Sensor and Pin Lock using Keypad for Charity Box Security, International Journal of Multidisciplinary Sciences and Arts, 1(2), 2022, pp. 125-133.
9. D. Hercog, T. Lerher, M. Truntič and O. Težak, Design and Implementation of ESP32-Based IoT Devices, Sensors, 23(15), 2023, pp. 1-20.
10. D D. Babić, I. Jovović, T. Popović, N. Kovač and S. Čakić, An Internet of Things System for Environmental Monitoring Based on ESP32 and Blynk, 26th International Conference on Information Technology (IT), Zabljak, Montenegro, 2022, pp. 1-5.

Authors Introduction

Abdul Jalil

He received his Bachelor's and Master's degrees in Computer Systems, 2012 and 2015, respectively, from STMIK Handayani Makassar, Indonesia. He then received his Doctor of Engineering in 2023 in the Department of Computer Science and Systems Engineering, Kyushu Institute of Technology, Japan. His research interest includes robotics, embedded systems, network, and Internet of Things.

Ms. Pujianti Wahyuningsih

She received her Bachelor's degree in Information Systems in 2015 from STMIK AKBA Makassar, Indonesia. Then received her Master's degree in Computer Systems in 2017, from STMIK Handayani Makassar, Indonesia. Her research interests include information systems engineering and management information systems.

Ms. Najirah Umar

She received her Bachelor's degree in Informatics Engineering in 2000 from STMIK Handayani Makassar, Indonesia. She then received her Master's degree in Electrical Engineering in 2007, from Hasanuddin University, Indonesia. His research interests include software engineering and database management systems.

Mr. Muhammad Risal

He received his Bachelor's degree in Computer Systems in 2006 from STMIK Handayani Makassar, Indonesia. Then received her Master's degree in Electrical Engineering in 2011, from Hasanuddin University, Indonesia. His research interests include embedded systems, microcontroller, networking, and Internet of Things.

Ms. Suwatri Jura

She received her Bachelor's and Master's degrees in Computer Systems, 2010 and 2015, respectively, from STMIK Handayani Makassar, Indonesia. His research interests include information systems and software engineering.

Ms. A. Edeth Fuari Anatasya

She received her Bachelor's degree in Information Systems in 2012 from STMIK AKBA Makassar, Indonesia. She then received her Master's degree in Computer Systems in 2015, from STMIK Handayani Makassar, Indonesia. Her research interests include informatics engineering and management information systems.

A penalized motion detection model for extracting ionospheric echoes from low signal-to-noise ratio Ionogram video images

Yuu Hiroshige
*Department of Creative Informatics, Kyushu Institute of Technology, 680-4 Kawazu,
Iizuka, 820-8502, Japan*

Akiko Fujimoto
*Department of Artificial Intelligence, Kyushu Institute of Technology, 680-4 Kawazu,
Iizuka, 820-8502, Japan*

Akihiro Ikeda
*National Institute of Technology, Kagoshima College, 1460-1 Shinko, Hayato,
Kirishima, Kagoshima, 899-5193, Japan*

Shuji Abe
*International Research Center for Space and Planetary Environmental Science, Kyushu University,
744, Motooka, Nishi-ku, Fukuoka, 819-0395, Japan*

Akimasa Yoshikawa
*International Research Center for Space and Planetary Environmental Science, Kyushu University,
744, Motooka, Nishi-ku, Fukuoka, 819-0395, Japan*

*Email: hiroshige.yuu147@mail.kyutech.jp, fujimoto@ai.kyutech.ac.jp, a-ikeda@kagoshima-ct.ac.jp,
abeshu@i-spes.kyushu-u.ac.jp, yoshikawa.akimasa.254@m.kyushu-u.ac.jp
www.kyutech.ac.jp*

Abstract

Measuring the altitude distribution of electron density in the upper atmosphere, known as the ionosphere, using High-Frequency radio wave reflections often causes the low signal-to-noise ratio of ionospheric echoes due to radio frequency interference. We propose a model for converting low-signal-to-noise-ratio ionospheric echo video images (Ionogram) into noise-reduced images using image processing techniques, for tracing the ionospheric echoes from Ionogram. The proposed method consists of three processing parts: noise removal optimized for individual Ionogram images, extraction of ionospheric echoes by penalized background subtraction technique, and fine-tuning of ionospheric echo signals using a minimum spanning tree algorithm. The proposed model successfully reproduces fine Ionograms with 98% recall and 99% precision.

Keywords: Ionospheric Observation, Computer Vision, Motion Detection, Time-series data

1. Introduction

The ionosphere is a part of Earth's upper atmosphere from about 100km to high altitudes and is very closely related to our modern life. The ionosphere has a significant impact on the radio propagation for communications, satellite positioning system, etc. The structure of ionosphere is not stable and variable in the time of day, season, latitude, and solar activity, because of the differences of the energetic photo-ionization process. The irregularities and disturbances in the ionosphere cause various severe problems, for example, lower accuracy of the radio navigation due to the global positioning system/Global Navigation Satellite System scintillation, disruptions in High Frequency (HF) radio wave bands or interferences due to anomalous propagation, etc. In order to utilize radio propagation safely and securely in modern society, it is essential to measure and analyze the fluctuating ionospheric environment at high speed and with high accuracy.

The fluctuating ionospheric environment is characterized by electron density variations in this region. The fluctuating ionospheric environment is characterized by electron density variations in this region. Ionosonde, a vertical ionospheric sounder, has been widely used to detect the ionospheric electron density as a function of height, using HF radio wave reflections. Ionosonde emits radio impulses with increasing frequency, measuring the time delay of radio signals received back from the different ionospheric layers. Fig. 1a shows an ionogram, which is a visualization of the output of an ionosonde measurement. The ionogram represents the received echo intensities in a virtual height (time-of-flight) against transmitted frequency. To analyze the fluctuating ionospheric environment, it is important of the interpretation of the ionogram, the so-called "scaling" of

Yuu Hiroshige, Akiko Fujimoto, Akihiro Ikeda, Shuji Abe, Akimasa Yoshikawa

ionograms. In other words, the scaling is the detection of the ionospheric echo on the ionograms and the extraction of characteristics of E and F layers (virtual heights: h'E, h'F, h'F2, etc. or critical frequencies: foE, foF1, foF2, etc.) from ionograms [1].

Classical manual scaling by professional researchers has been used, but several methods of automatic scaling have been proposed in recent decades. There are two main approaches to automatic scaling: image processing (computer vision) approaches [2], [3] and deep-learning image generation approaches [4], [5]. The both methods perform the scaling of ionograms with higher accuracy, but the expected ionogram for scaling is high signal-to-noise ratio (SNR) image. Here, the signal is the ionospheric echo, and the noise is the background noise, such as radio frequency interference (RFI) from other emissions. In Fig. 1a, the noise appears as the vertical strip. The ionogram visualization process also generates noise [6]. Ionogram images are visualized using spectral analysis techniques such as the Fast Fourier Transform (FFT) algorithms on original data obtained by the ionosonde. Artificial noises are generated by this process. The noise due to RFI and FFT complicates automatic information extraction from ionogram through scaling [7].

Ionosonde transmitting power by a typical HF radar is high power, several tens of kW. In ionogram images acquired by such high-power ionosondes, the signal strength of the ionospheric echo and the background noise are distinctly different, and the ionospheric echo is sufficiently strong. This means that scaling can be performed from ionogram images with high SNR. On the other hand, our ionosonde observations with Frequency Modulated Continuous Wave (FM-CW) type radar, operated by the joint research with Kyushu Institute of Technology and Kyushu University, have a maximum peak power of 20 watts. Although the advantage of FM-CW radar is its high gain even at low power, the reflected echo strength from the ionosphere is relatively weak compared to a typical HF radar. Therefore, the signal strength of the ionospheric echoes is often comparable to the background noise, and it is difficult to separate from each other. Since the scaling process is performed from ionogram images with low SNR, a simple noise elimination processing is not sufficient.

There are two challenges in automatic ionogram scaling for FM-CW ionogram. First, as described above, scaling accuracy is poor when the input image is an ionogram with low SNR. Second, in terms of the quasi-real-time monitoring of the ionospheric condition, the processing time on scaling is required within few seconds. We sequentially operate the FM-CW ionosonde with every 3 minutes. A low-cost computer, which is usually expected to perform with no high processing speed, is required on the observation site. The deep learning approaches have a processing time of 2-3 minutes, and real-time scaling requires a computer with high processing speed. We try to propose a new algorithm without deep-learning

approach for scaling which includes the elimination process of background noises.

a. Original Ionogram   b. Full denoised Ionogram

c. Histogram of signal intensity for Fig .1a

d. Histogram for signal intensity of Fig. 1b

Fig. 1. Comparison of the signal intensities of original and full denoised ionograms recorded at 00:06:00 UT, January 1, 2019.

In this paper, we provide a new method to scaling of FM-CW ionogram with low SNR. The new scaling method can be applied to ionograms containing any ionospheric echo class (E and F layers) for quasi-real-time detection. The shape and position of the ionogram echoes vary with time, due to daily and seasonal changes in the ionospheric electron density structure and its solar activity dependence. Therefore, our study focuses on time-variant scaling based on time-series analysis. The proposed method consists of three parts. First, we remove the noise from the ionogram with low SNR, optimizing the noise elimination threshold value for individual ionogram image. Second, we extract the ionospheric echoes based-on penalized background subtraction technique. Finally, the ionogram with high SNR are regenerated by fine-tuning of ionospheric echo signals using a minimum spanning tree algorithm.

2. Dataset

The dataset used in this study consists of 960 ionogram data observed every 90 seconds on January 1, 2019, at the FM-CW radar ionosonde located at Sasaguri (130.34° E, 33.37° N). Fig. 1a shows a sample of the ionogram image data obtained on January 1, 2019, at 00:06:00 UT (Universal Time). Ionogram image shows the signals with frequency values from 2 MHz to 9 MHz and altitude values from 0 km to 700 km. In order to treat the ionograms as a two-dimensional array, we consider the class width as an array with 278 x 238 bins with a frequency of 0.25 MHz and an altitude of 2.9296875 km.

As mentioned above, the ionogram is usually contaminated with RFI noise appearing as vertical strips,

shown in Fig. 1a. As shown in Fig. 1, the signal intensity distribution in histograms is significantly different between the original ionogram (Fig. 1c) and the full denoised ionogram (Fig. 1d). Fig. 1b was produced by manual removal of the background noise from Fig. 1a. The signal distribution in Fig. 1d only corresponds to the ionospheric echoes. Comparing between Fig. 1c and Fig. 1d, it is obvious that ionospheric echoes are highly intense relative to noises. In other words, the almost background noises are in lower value of signal intensity.

3. Methodology

The proposed method consists of three processing parts: (1) Weak Signal Reduction, (2) Penalized Background Subtraction, and (3) Fine-tuning Ionogram. In the processing part on Weak Signal Reduction, the divided value between the ionospheric echoes and the background noises is automatically estimated based-on Ridge Regression [8]. After this process, almost background noises are eliminated, and we get the denoised ionogram remaining enough ionospheric echoes. The processing part of Penalized Background Subtraction enables the extraction of ionospheric echoes with motion detection algorithms. Finally, the processing part of Fine-tuning Ionogram provides the regenerated ionogram, removing the remained noises and interpolating the ionospheric echoes from the detection region using the minimum spanning tree in the graph structure. The following subsections describe the detail algorithm of each part.

3.1. *Weak Signal Reduction*

As pre-processing, the weaker RFI noises are eliminated by aligning the average signal strength in the altitude direction to a constant value, which is the mode of histogram of signal intensity. We define the elimination function of RFI noises as below:

$$\boldsymbol{R}_{f,h} = \boldsymbol{I}_{f,h} + MO - \mu_f \tag{1}$$

where $\boldsymbol{R}_{f,h}$ and $\boldsymbol{I}_{f,h}$ represent the power of the original ionogram as input and the eliminated ionogram as output, respectively. The subscript f and h indicate the positions of the frequency and altitude values on the ionogram. MO represents the mode value of signal intensity histogram. μ_f is the average of signal intensity at each frequency bin.

As shown in Fig. 1, it is possible to separate the ionospheric echoes and the background noise by estimating the boundary value between them. However, the distribution of signal intensity is different each ionogram. This suggests that it is necessary to estimate the threshold value according to the signal distribution of each ionogram. We estimate this threshold of boundary between the ionospheric echoes and the background noise by ridge regression [8].

Ridge regression is a method which further improves generalization performance by adding an L2 regularization term to multiple regression analysis. For each raw ionogram, Ridge regression estimates the coefficients, w, of multiple regression models for the boundary threshold value as the objective variable, y, and when you give the frequency, mean and variance of the mode of the signal intensity as the explanatory variable, X. The loss function of the ridge regression, L, is shown in Eq. (2), where the weight of the L2 regularization term is $\alpha = 0.7$.

$$L = (y - Xw)^T(y - Xw) + \alpha\|w\|_2^2 \tag{2}$$

$$\hat{y} = X^T w \tag{3}$$

The equation for the objective variable is given by Eq. (3). The above equation yields a model for estimating the threshold for removing an appropriate low-intensity signal. The ionogram after applying this algorithm is shown in Fig. 2.

Fig. 2. Ionogram with Weak Signal Reduction applied to Fig. 1a

3.2. *Penalized Background Subtraction*

As shown in Fig. 2, it is still impossible to completely remove the noise from the ionogram, even after removing the low intensity signal. This is because the noise signal is more intense than the ionospheric echoes. We use Background Subtraction method [9] to extract of the ionospheric echoes from the ionogram and to ignore the remained intense background noise. The Background Subtraction method is one of the motion detection techniques to detect moving objects in a video frame. A moving object is detected by comparing a frame of video with its corresponding background image. It is important to generate a background image without moving objects. In this study, we consider an ionospheric echo as the moving object and an ionogram image with only noise present as the background image. The processing part of Penalized Background Subtraction is divided into four steps and the following subsections describe the detail algorithms.

Yuu Hiroshige, Akiko Fujimoto, Akihiro Ikeda, Shuji Abe, Akimasa Yoshikawa

3.2.1. *Penalty based on ionospheric echo shape*

As shown in Fig. 1b, ionospheric echoes in the ionogram are observed mainly as line shapes extending in the frequency direction. This suggests that the signal intensity of ionospheric echoes is higher at specific altitudes than at other altitudes in each frequency band bin. In fact, as shown in Fig. 3b, the distribution of signal intensity against the altitude in interest frequency bin, 4.0 MHz, shows two significant peaks corresponding to the ionospheric echoes.

a.   Original Ionogram data and the interest frequency, 4.0 MHz, in red vertical line

b.   Signal intensity as a function of the altitude for the interest frequency, 4.0 MHz

Fig. 3. Signal distribution in altitude direction, recorded at 06:01:30 UT, January 1, 2019

For generating the background image from the input ionogram image including the ionospheric echoes, we only penalize the corresponded frequency into the significant ionospheric echoes in each ionogram, to reduce the amplitude of signal power. The detail operating flow is as shown in Fig. 4. First, a vector of the signal intensity is extracted for each frequency bin from the two-dimensional matrix of the ionogram. Next, the components that penalize the extracted vectors are selected according to the flowchart shown in the Fig. 4.

Fig. 4. Determination flow of penalization for the generation of background image

The penalty, $f_{f,h}^p$, is given as Eq. (4), where $I[t]_{f,h}$ is the signal intensity of the unprocessed ionogram array at time t. We use anomaly detection in deciding whether to penalize. This is designed based on the Hotelling's theory

[10]. In the present work, a normal distribution is assumed for each signal with the same frequency value after removing weak intensity signals, and signals that deviate from the 95% confidence interval are penalized by detecting them as anomalies. By penalizing with these procedures, we ideally aim to get the background image contains only noise.

$$f_{f,h}^p(I[t]_{f,h}) = \begin{cases} \dfrac{1}{I[t]_{f,h}} & (penalized) \\ I[t]_{f,h} & (\text{else}) \end{cases} \quad (4)$$

3.2.2. *Penalty based on one previous detection frame image*

The background subtraction method can efficiently detect moving objects under the constraint that the background scene is stationary. However, the accuracy decreases under dynamically changing background image. Ionograms are no exception to this problem, as the position and shape of background noise as well as ionospheric echoes change from moment to moment. Therefore, the ionospheric echoes detected from the previous ionogram are used to estimate the background image.

This method assumes a two-dimensional Gaussian distribution centered on the detected component. Here, the frequency value is f and the height value is h, from which the mean μ_f, μ_h, and standard deviations σ_f and σ_h are calculated, respectively. The covariance ρ is given by Eq. (5), considering the characteristic that the ionospheric echoes vary in the frequency direction. Eq. (6) and (7) show the weight Gaussian. The generated background image is penalized by considering the previous ionospheric echo detection region.

$$\rho = \begin{pmatrix} 8 & 0 \\ 0 & 4 \end{pmatrix} \quad (5)$$

$$f_{f,h}^g = \frac{1}{2\pi\sigma_f\sigma_h\sqrt{1-\rho^2}}e^\alpha \quad (6)$$

$$\alpha = -\frac{1}{2(1-\rho^2)}\left\{\frac{(f-\mu_f)^2}{\sigma_f{}^2} - \frac{2\rho(f-\mu_f)(h-\mu_h)}{\sigma_f\sigma_h} + \frac{(h-\mu_h)^2}{\sigma_h{}^2}\right\} \quad (7)$$

3.2.3. *Penalized background image generation algorithm*

In addition to the background image update algorithm for the Background Subtraction method [9], the background image, $B[t]$, is updated by adding the weight terms of Eq. (4) and Eq. (6) as follows.

$$\boldsymbol{B}[t] = \alpha \cdot \boldsymbol{B}[t-1] + (1-\alpha) \cdot f^p_{f,h}(\boldsymbol{I}[t])$$
$$-\beta \cdot f^g_{f,h}(\boldsymbol{F}[t-1]) \quad (8)$$

where α and β are weights for $f^p_{f,h}$ and $f^g_{f,h}$. In this study, we set $\alpha = 0.5$ and $\beta = 8.0$, which are assumed the same level of influence of each term.

3.2.4. *Detection region determination*

The positions where the result of subtracting the background image obtained by Eq. (8) from the input image is a positive value is detected as ionospheric echoes, $\boldsymbol{D}[t]_{f,h}$, as shown in Eq. (9). It is assumed that background regions are distributed where the difference values are small. Thus, *thresh* in Eq. (9) is set to the class value of the mode in the histogram of difference values between the input and background images.

$$\boldsymbol{D}[t]_{f,h} = \begin{cases} 1 & (\boldsymbol{I}[t]_{f,h} - \boldsymbol{B}[t]_{f,h} \geq thresh) \\ 0 & (else) \end{cases} \quad (9)$$

Fig. 5 shows that our method can detect ionospheric echoes. This indicates that our method is also effective for ionograms with high signal-to-noise ratios.

a.   Original Ionogram   b.   Ionogram detection data after processing for Fig. 2

Fig. 5. Ionospheric echo detection by Penalized Background Subtraction, original ionogram data recorded at 00:06:00UT, January 1, 2019

3.3. *Fine-tuning Ionogram*

As shown in Fig. 6a, the background subtraction method can be used to detect ionospheric echoes, but it can also leave isolated noises. In addition, the signal should be detected might have missing or incorrectly detected as the ionospheric echo. For solving these problems, we propose two post-processing methods: the convolution operation to remove noise that appears as isolated echoes, and the optimization for interpolating and denoising of the detection region using an algorithm based on a minimum spanning tree.

3.3.1. *Denoise impulse noise*

The convolution is, an algorithm of image processing, used to reduce of impulse noises in the present work. The multiplication a two-dimensional square matrix composed of odd-numbered widths several times is applied to the detected ionogram with two-dimensional array. If there are many signals around the element of interest that have been labeled as ionospheric echoes by the background subtraction method, then they should be ionospheric echoes, otherwise they should be considered as noise.

The filter $H_{i,j}$ consists of an $n \times n$ square matrix, with all elements 1. It is also scanned with stride 1 to maintain the size of the two-dimensional array of the ionogram. The convolution equation using this filter is shown in Eq. (9). Let D^{bs} be the array output by the algorithm in the processing part of Penalized Background Subtraction and D be the array after impulse noise removal.

$$D_{f,h} = \begin{cases} 1 & \left(\sum_{i=-n}^{n} \sum_{j=-n}^{n} D^{bs}_{f+i,h+j} \cdot H_{i,j} \geq 2+n\right) \\ 0 & (else) \end{cases} \quad (10)$$

In this study, impulse noises are removed by performing the convolution operation, changing the filter width n in the order of 5, 7, 9, and 11. Ionograms before and after the removal of impulse noises are shown in Fig. 6. It is obvious that this process reduces extracted excess noises and does not disappear the detected signal of ionospheric echoes (Fig. 6b).

a.   Ionospheric echo detection results before processing   b.   Ionospheric echo detection results after processing

Fig. 6. Comparison of ionograms with and without Impulse noise reduction processing

3.3.2. *Optimization by a minimum spanning tree*

First, for all points detected as ionospheric echoes, these detection points are connected for each E and F layer. The points are connected so that they form a minimum spanning tree, considering the linear shape of the ionospheric echoes. Next, edges with large distances between vertices are removed. The graph to which the removed vertices belong is judged to be noise and removed. For the remained edges after the above processes, the detection area is interpolated between the two vertices that compose the edge. The result is shown in Fig. 7.

Yuu Hiroshige, Akiko Fujimoto, Akihiro Ikeda, Shuji Abe, Akimasa Yoshikawa

a. Ionospheric echo detection results before processing

b. Ionospheric echo detection results after processing

Fig. 7. Comparison of ionograms with and without optimization process by minimum spanning tree

4. Result and discussion

To evaluate the accuracy of our automatic scaling method, we compared the accuracy of ionograms obtained by our algorithm with that of manually scaled ionograms. The evaluation is based on the absolute error of the critical frequencies' foE and foF of each layer of the ionogram and the percentage of absolute error within 0.5 MHz indicating that sufficient accuracy was achieved. The accuracy of the proposed method is compared with that of the Weighted Background Subtraction without each weight term (Table 1). The average processing time of the proposed algorithm was 7.1 seconds per ionogram. The proposed model successfully reproduces fine ionograms with 98% recall and 99% precision.

Table 1 shows that the detection results were satisfactory despite the low SNR of the ionograms. From Table 1 and the detection result shown in Fig. 7b, it is clearly found that the addition of a weight term to the background subtraction algorithm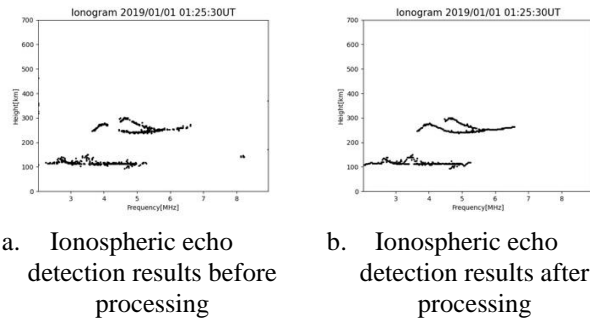 was effective. The processing time with few of seconds indicates that the proposed method is useful to generate quasi-real-time ionograms with high SNR.

Table 1. Performance Comparison with different weight terms in this study

Model		Main result	Base+ $f_{penalty}$	Base+ f_{gauss}	Base
Absolute error (MHz)	foE	0.47	0.93	0.92	0.99
	foF	0.26	0.44	0.42	2.6
Rate of accuracy (%)	foE	67%	50%	50%	47%
	foF	83%	49%	61%	15%

On the other hand, the continuity of ionospheric echoes remains an issue. Some ionospheric echoes are missing from the detection results of our method. This is due to the altitudinally banded intermodulation noise in the ionograms. This is a factor that reduces the signal strength, and thus causes the background subtraction method to fail to detect the ionospheric echoes. Also, although not present in this study's data set, scattered

signals such as those seen in equatorial spread F are not expected to give meaningful results.

5. Conclusion

Automatic scaling of ionograms presents various challenges for the ionogram image contaminated with noises. In this study, we proposed an algorithm to automatically extract ionospheric echoes from ionograms with low signal-to-noise ratio.

Our algorithm is based on statistical thresholding according to the signal intensity distribution of each ionogram, image processing techniques focusing on time variability and ionospheric echo characteristics, noise reduction filtering, and graph theory. Although it is necessary to define some parameters, it is possible to extract ionospheric echoes from ionograms in quasi-real-time without the huge amount of teacher data as in deep learning.

On the other hand, for frequent noise and special ionospheric echo classes, the algorithm may not guarantee accuracy. The algorithm should be improved to be robust in these cases as well for the future work.

References

1. B. Zolesi, L. R. Cander, "Ionospheric Prediction and Forecasting", Springer, 2014
2. Z. Chen, S. Wang, S. Zhang, G. Fang, J. Wang, "Automatic scaling of F layer from ionograms", Radio Science, Vol. 48, No. 3, pp. 334–343, 2013.
3. Z. Chen, Z. Gong, F. Zhang, G. Fang, "A new ionogram automatic scaling method", Radio Science, Vol. 53, No. 9, pp. 1149–1164, 2018.
4. C. De La Jara, C. Olivares, "Ionospheric echo detection in digital ionograms using convolutional neural networks", Radio Science, Vol. 56, No. 8, 2021.
5. Z. Xiao, J. Wang, J. Li, B. Zhao, L. Hu, L. Liu, "Deep-learning for ionogram automatic scaling", Advances in Space Research, Vol. 66, No. 4, pp. 942–950, 2020.
6. F. Arıkan, O. Arıkan, S. Salous, "A new algorithm for high-quality ionogram generation and analysis", Radio Science, Vol. 37, No. 1, pp. 1-11, 2002.
7. M. D. E. Turley, A. J. Heitmann, R. S. Gardiner-Garden, "Ionogram RFI Rejection Using an Autoregressive Interpolation Process", Radio Science, Vol. 54, No. 1, pp. 135–150, 2019.
8. D. E. Hilt, D. W. Seegrist, "Ridge: a computer program for calculating ridge regression estimates", Research Note NE-236. Upper Darby, PA: U.S. Department of Agriculture, Forest Service, Northeastern Forest Experiment Station, pp. 1-7, 1977.

9. R. Cutler and L. Davis, "View-based detection and analysis of periodic motion", Proceedings. Fourteenth International Conference on Pattern Recognition (Cat. No.98EX170), Brisbane, QLD, Australia, Vol. 1, pp. 495-500, 1998.

10. H. Hotelling, "Multivariate Quality Control Illustrated by Air Testing of Sample Bombsights", In: Eisenhart, C., Hastay, M.W. and Wallis, W.A., Eds., Techniques of Statistical Analysis, McGraw Hill, New York, pp. 111-184, 1947.

Authors Introduction

Mr. Yuu Hiroshige

He received his Bachelor's degree in Computer Science and Systems Engineering in 2022 from the Faculty of Computer Science and Systems Engineering, Kyushu Institute of technology in Japan. He is currently a master student in Kyushu Institute of Technology, Japan.

Dr. Akiko Fujimoto

She is an Associate Professor at the Faculty of Computer Science and Systems Engineering at Kyushu Institute of Technology in Japan. She graduated from the Department of Earth and Planetary Sciences, Kyushu University, in 2005. She received her Ph.D. degree in Science from Kyushu University in 2010. Her research interest is Space Weather and Upper Atmospheric Sciences.

Dr. Akihiro Ikeda

He is an Associate Professor of the Department of Liberal Arts and Sciences at National Institute of Technology, Kagoshima College in Japan. He graduated from the Department of Earth System Science, Fukuoka University, in 2005. He received his Ph.D. degree in Science from Kyushu University in 2010. His research interest is Upper Atmospheric Sciences.

Dr. Shuji Abe

He is a Research Fellow at the International Research Center for Space and Planetary Environmental Science, Kyushu University in Japan. He graduated from the Department of Earth and Planetary Sciences, Kyushu University, in 1999. He received his Ph.D. degree in Science from Kyushu University in 2005. His research interest is space weather.

Dr. Akimasa Yoshikawa

He is a Professor at the Faculty of Sciences, Kyushu University in Japan. He graduated from the Department of Physics, Yamaguchi University, in 1989. He received his Ph.D. degree in Science from Kyushu University in 1998. His research interest is space plasma physics.

Verification Experiments on the Lower Back Burden caused by Posture and Environment during Lifting Operations

Tomoka Kimura

Graduate School of Engineering, University of Miyazaki, 1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan

Yutaro Fujino

Faculty of Engineering, University of Miyazaki, 1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan

Sachiko Kido

Interdisciplinary Graduate School of Agriculture and Engineering, University of Miyazaki, 1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan

Praveen Nuwantha Gunaratne

Interdisciplinary Graduate School of Agriculture and Engineering, University of Miyazaki, 1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan

Hiroki Tamura

Faculty of Engineering, University of Miyazaki, 1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan

E-mail: hi19014@student.miyazaki-u.ac.jp, hi20035@student.miyazaki-u.ac.jp, z321t01@student.miyazaki-u.ac.jp, ti20060@student.miyazaki-u.ac.jp, htamura@cc.miyazaki-u.ac.jp

Abstract

One of the measures to prevent back pain is the use of appropriate posture corrections. In generally, the Squat method (method of keeping the knees bent and the waist as straight as possible) is recommended over the Stoop method (method of keeping the knees straight and the waist bent). The results of previous studies using acceleration and muscle potentials have shown that the lifting with knees in a kneeling position reduces the amount of load born by the lower back. However, few studies discuss the case where there is an obstacle between the subject and the object to be lifted and further, the scenarios where the object is away from the body's center of gravity. Therefore, this research focus on analyzing three types of movements using the AnyBody musculoskeletal mechanics analysis software, the Yo-bukun back pain prevention application, and the Delsys surface EMG and verifying the amount of lower back burden when there is an obstacle between the lifting object and the subject. This paper presents the verification results.

Keywords: Lower back burden, Posture correction, EMG, The body's center of gravity

1. Introduction

Frequent lifting activities in the construction and nursing care industries place a heavy burden on the lower back and can easily lead to lower back pain. One of the measures that can be taken by individuals to prevent back pain is to improve posture. There are two main types of lifting movements: the Stoop method (keeping the knees straight and the hips flexed) and the Squat method (keeping the knees flexed and the hips as straight as possible). The Japanese Ministry of Health, Labour and Welfare's "Guidelines for Prevention of Back Pain in the Workplace" recommends the knee-bend lifting method. The results of the previous study and the analysis using the "Yo-bukun" application for measuring back strain showed that the lifting method with the knees bent reduced the amount of back strain compared to the lifting method without knees bent [1], [2], [3]. In addition, the previous study showed that the lifting method with the knees bent was associated with a lower EMG burden

Tomoka Kimura, Yutaro Fujino, Sachiko Kido, Hiroki Tamura

when the lifting object was located close to the body's center of gravity [4], [5].

However, few previous studies have discussed the case where there is an obstacle between the lifting object and the subject and also when the object is far from the body's center of gravity. Therefore, the purpose of this paper is to analyze three types of movements using the AnyBody musculoskeletal mechanics analysis software, Yo-bukun back pain prevention application, and the Delsys surface electromyograph, in order to verify the amount of back burden when there is an obstacle between the lifting object and the subject.

2. Methodology

In this paper, lumbar burden was measured using AnyBody, Yo-bukun applications, and Delsys EMG.

2.1. Measuring equipments

AnyBody is a musculoskeletal mechanics analysis software from Terrabyte Co. Basically, it calculates the forces acting on each part of the human body, such as muscle activity, muscle force, and joint moments, when a movement is applied to a musculoskeletal model using inverse dynamics analysis.

Yo-bukun is an application developed by the University of Miyazaki in collaboration with DENSAN Corporation for the prevention of back pain for the iPhone. It is calibrated when the user stands upright, and calculates the amount of lumbar burden based on the tilt of the iPhone.

Delsys is the generic name for Trigno EMG system incorporated with some additional sensors developed by Delsys. In addition to EMG, IMU sensors (accelerometer, gyroscope, and magnetometer) are built in to collect kinematic data (angles) at the same time.

2.2. Measurement method

A video camera footage was used as the method of measurement for the analysis by AnyBody application. Under Yo-bukun approach, subjects were asked to keep their iPhones installed with Yo-bukun application, in their chest pockets for the measurement. EMG based approach employs two EMG sensors placed on the subject's erector spinae. The positioning of the lumbar EMG sensors is shown in the Fig. 1. The left side sensor is considered as the first channel and the right side sensor is the second channel.

The subjects were all healthy adult males, 174±3 cm tall and 22±1 years old. Subject identifiers are A, B, and C respectively. The movement to be measured was the lifting and lowering of an object while kneeling on one knee. The dimensions of the object (a box), are 35(W)× 23.5(D)×31(H)cm and the weight is 1 kg. Three types of movements were measured to examine the effects of

the presence of an obstacle and the distance between the centers of gravity. The size of the obstacle was 51(W)×10(D)×51(H)cm. The subject's center of gravity was set at the pelvis. The experimental apparatus is shown in the Fig. 2.

The initial measurement was a subject lifting the object as close as possible to its body's center of gravity without placing any obstacles. The step wise process involves a subject lifting the object and remaining for 5 seconds, before lowering it. The distance between the center of gravity of the object and the subject's center of gravity was 10 cm. In the second measurement, the obstacle is in place, and the subject lowers the object to behind the obstacle starting from an apparatus configuration as in Fig. 2, subsequently returns it to the top of the obstacle. Upon lowering the object behind the obstacle, the subject stood upright and maintained the position for a duration of 5 seconds.

Fig.1 . EMG sensor placements

Fig.2 . Experimental apparatus (includes the lifting object and the obstacle)

Fig.3 Posture in which the center of gravity of both the subject and the object 10 cm

Fig.4 . Posture in which the center of gravity of both the subject and the object 30 cm

Then, the subject knelt down once more and raised the object back over the obstacle. The distance between the object's center of gravity and the subject's center of gravity was 30 cm. The third measurement closely resembled the second, with the distance between the center of gravity of the object and that of the subject being 45 cm. The postures observed during such measurements at the distance between the center of gravity of the object and that of the subject being 10 cm and 30 cm are depicted in Fig 3 and Fig. 4.

After that the transition of the amount of burden applied on the lower back during such movements were calculated using AnyBody and Yo-bukun. Later, the analysis incorporates the maximum and average values of the amount of burden observed here. The Delsys measured the EMG potentials of the lower back during movements. To reduce the influence of noise, the rate of increase was calculated by dividing the difference between the mean value of the EMG during movement and the mean value of the EMG at rest by the mean value of the EMG at rest. The resting EMG potentials refer to the potentials recorded when the subject is at rest for a 5-second interval between lifting and lowering movements.

3. Results

This chapter outlines the measurement outcomes for the three equipment types and provides a comparison between them. The movements are individually analyzed for lifting and lowering at center-of-gravity distances of 10 cm, 30 cm, and 45 cm, respectively.

3.1. Similarity between AnyBody and Yo-bukun

To validate the reliability of Yo-bukun, the similarity between AnyBody and Yo-bukun was assessed. Verification utilized the measurement data of subject B, analyzable with A. The coefficients of determination for each center-of-gravity distance are presented in Table 1. A somewhat stronger correlation was evident in the data at a center-of-gravity distance of 10 cm, while robust correlations were observed at center-of-gravity distances of 30 cm and 45 cm.

Table 1. Correlation between AnyBody and Yo-bukun (Subject B)

Center of gravity distance (cm)	R^2
10	0.5582
30	0.9289
45	0.9812

Table 2. The rates of increase in Yo-bukun measurements

Subject Identity	Center of gravity distance (cm)	Raising movement (Up) %	Lowering movement (Down) %
A	10	97.39	107.02
	30	110.71	104.21
	45	120.70	105.34
B	10	90.68	108.78
	30	140.64	132.38
	45	172.14	163.56
C	10	137.98	157.31
	30	197.84	193.90
	45	196.63	162.03

3.2. Analysis results using Yo-bukun

The rate of increase was also determined for the Yo-bukun in comparison with myoelectric potentials. The rate of increase in the Yo-bukun measurements for the three movements is shown in Table 2. The comparisons

Tomoka Kimura, Yutaro Fujino, Sachiko Kido, Hiroki Tamura

for the raising and lowering movements are shown in Fig. 5 and 6, respectively. The average rate of increase was obtained by excluding the rate of increase at 10 cm. The average rate of increase was 156.44% for the lifting motion and 143.57% for the lowering motion.

Fig.5. Yo-bukun Increase rate (up)

Fig.6. Yo-bukun Increase rate (down)

3.3. Analysis results using Delsys device

The rates of increase in lumbar EMG potentials during the three movements are shown in Table 3, and comparisons for raising (up) and lowering (down) movements are shown in Fig. 7 and Fig. 8, respectively. In all cases, the rate of increase was lowest when the center of gravity distance was 10 cm for each subject. For Subjects A and C, the rate of increase was higher when the center of gravity was 45 cm than when the center of gravity was 30 cm, for both lifting and lowering. For subject B, the rate of increase was higher when the center of gravity was 30 cm than when the center of gravity was 45 cm. The average rate of increase was obtained by excluding the rate of increase at 10 cm. The average rate of increase was 404.43% for the lifting motion and 394.71% for the lowering motion. The average increase was 2.59 and 2.75 times that of Yo-bukun for each operation.

Table 2. The rates of increase in lumbar EMG potentials

Subject Identity	Center of gravity distance (cm)	Raising movement (Up) %	Lowering movement (Down) %
A	10	30.43	5.73
	30	230.24	239.78
	45	339.78	296.69
B	10	59.08	68.08
	30	352.31	357.99
	45	247.76	284.30
C	10	16.20	16.48
	30	518.76	534.77
	45	737.72	654.72

Fig.7. EMG Increase rate (up)

Fig.8. EMG Increase rate (down)

Fig.9 . Posture (Left: Subject B, Right: Subject C)

4. Conclusion

The rate of increase of EMG potentials during the movement was different in Subject B than in the other subjects. The possible reason for this is the difference in posture. The most notable difference is the curvature of the back. A comparison of the posture of Subject B and Subject C at a center of gravity distance of 45 cm is shown in the Fig. 9. The yellow circle at the pelvis position is the center of gravity. The white line is a simplified representation of the back. It can be seen that subject B has a straighter back than subject C. This indicates that subject B has a straighter back than subject C. This suggests that Subject B had less muscle stretch than Subject C, and that the rate of increase in myopotential was less pronounced in Subject B than in the other subjects.

In this paper, we analyzed movements using three types of measurement devices and verified the amount of lumbar burden in a situation where there is an obstacle between the subject and the lifting object. The analysis results shows that the change in the amount of lumbar burden due to the difference in the center of gravity distances was more pronounced in the electromyograms than in the measurements from the posture-based devices, such as the Anybody and the Yo-bukun. Therefore, it is possible that even in the one-kneeling lifting method, a significant distance between the object and the operator's center of gravity might not result in a reduction of electromyographic back burden.

References

1. H. Tamura, K. Sakurai, K. Tanno, and Y. Fuse, "A Study on the Lumbar Burden Evaluation of Work using One Smartphone," *Journal of Robotics, Networking and Artificial Life*, vol. 5, no. 3, pp. 173–173, Jan. 2018, doi: https://doi.org/10.2991/jrnal.2018.5.3.7.
2. H. Tamura, K. Sakurai, K. Tanno, and Y. Fuse, "A Study on the Lumbar Burden Evaluation of Work using One Smartphone," *Proceedings of International Conference on Artificial Life and Robotics,* vol. 23, pp. 550–553, Feb. 2018, doi: https://doi.org/10.5954/icarob.2018.os14-1.
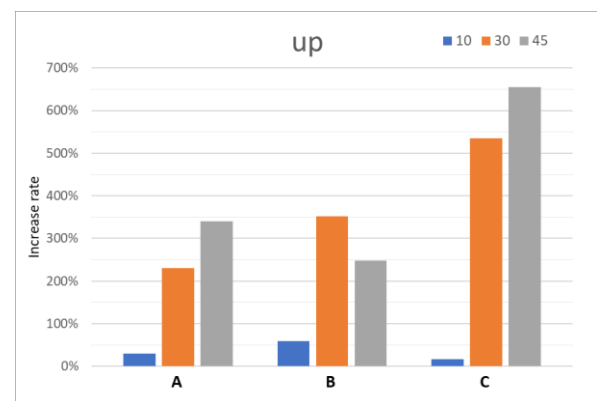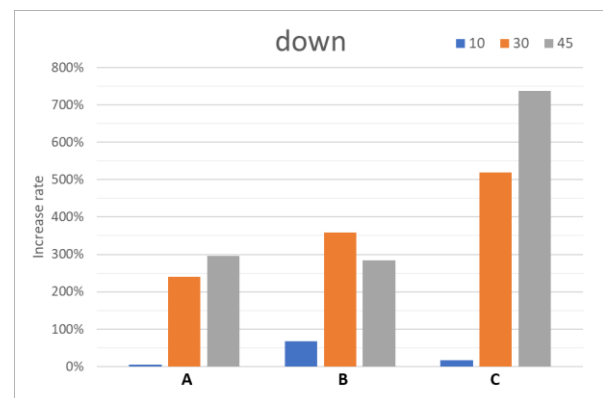3. M. F. Antwi-Afari, H. Li, D. J. Edwards, E. A. Pärn, J. Seo, and A. Y. L. Wong, "Biomechanical analysis of risk factors for work-related musculoskeletal disorders during repetitive lifting task in construction workers," *Automation in Construction*, vol. 83, pp. 41–47, Nov. 2017, doi: https://doi.org/10.1016/j.autcon.2017.07.007.
4. I. Kingma et al., "Lumbar loading during lifting: a comparative study of three measurement techniques," *Journal of Electromyography and Kinesiology*, vol. 11, no. 5, pp. 337–345, Oct. 2001, doi: https://doi.org/10.1016/s1050-6411(01)00011-6.
5. C. Larivière, A. Bertrand. Arsenault, D. Gravel, D. Gagnon, and P. Loisel, "Surface electromyography assessment of back muscle intrinsic properties," *Journal of Electromyography and Kinesiology*, vol. 13, no. 4, pp. 305–318, Aug. 2003, doi: https://doi.org/10.1016/s1050-6411(03)00039-7.

Authors Introduction

Ms. Tomoka Kimura

She received her Bachelor's degree in Engineering in 2023 from the Faculty of Engineering, Miyazaki University in Japan. She is currently a master student in Miyazaki University, Japan.

Mr. Yutaro Fujino

Yutaro Fujino Born in 2001. He is currently studying in the department of environment robotics, and will receive the B,Eng from University of Miyazaki in 2024. His current reseasrch is Analysis of Low Back Burden Utilizing a Smartphone Application Called Yo-bukun.

Ms. Sachiko Kido

She received her M.S. from Nakamura Gakuen University, Graduate School of Human Development in 2015. Currently enrolled in the doctoral program at the Graduate School of Agriculture and Engineering, University of Miyazaki. Assistant Professor, Department of Early Childhood Education, Higashikyushu Junior College, 2010. Lecturer, Faculty of Education, Miyazaki International University, since 2019. Her main research interest is the development of motor skills in young children.

Mr. Praveen Nuwantha Gunaratne

He received his Bachelor's degree in Engineering in 2018 from the Faculty of Engineering, University of Moratuwa, Sri Lanka. He is currently a Doctoral student in University of Miyazaki, Japan.

Tomoka Kimura, Yutaro Fujino, Sachiko Kido, Hiroki Tamura

Prof. Hiroki Tamura

He received the B.E. and M.E. degree from Miyazaki University in 1998 and 2000, respectively. From 2000 to 2001, he was an Engineer in Asahi Kasei Corporation, Japan. In 2001, he joined Toyama University, Toyama, Japan, where he was a Technical Official in the Department of Intellectual Information Systems. In 2006, he joined Miyazaki University, Miyazaki, Japan, where he was an Assistant Professor in the Department of Electrical and Electronic Engineering. Since 2015, he is currently a Professor in the Department of Environmental Robotics. His main research interests are Neural Networks and Optimization Problems. In recent years, he has had interest in Biomedical Signal Processing using Soft Computing.

Development of Smartphone Application for Calculating the Low Back Pain Risk

Seigo Imura
Faculty of Engineering, University of Miyazaki, 1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan

Praveen Nuwantha Gunaratne
Interdisciplinary graduate school of Agriculture and Engineering, University of Miyazaki, 1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan

Hiroki Tamura
Faculty of Engineering, University of Miyazaki, 1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan
Email: hi20006@student.miyazaki-u.ac.jp, ti20060@student.miyazaki-u.ac.jp, htamura@cc.miyazaki-u.ac.jp

Abstract

This study primarily relies on a smartphone application, developed within our research institute known as Yo-bukun, for the real-time estimation of lumbar burden. The Yo-bukun is capable of estimating the lumbar burden of a subject (the user) by placing an app. installed smartphone in the subject's chest pocket, while the subject is lifting/ relocating an object. The subject's movements are assessed through sensors embedded in the smartphone and certain aspects of their physical information initially fed into the app. are used in the estimation formula for determining lumbar burden. In the current scenario, Yo-bukun lacks the capability to ascertain whether the user is holding an object or not; consequently, it can only estimate the lumbar burden for limited cases of a subject holding an object. To address such limitations, the proposed system integrates voice recognition to facilitate lumbar burden estimation, considering the presence or absence of an object. Further, it was incorporated with the capability to recalculate lumbar burden after measurements, enabling prospective studies.

Keywords: Lumbar burden estimation, smartphone application, voice recognition.

1. Introduction

The smartphone application called Yo-bukun allows you to estimate the lumbar burden when holding an object in real time by placing a smartphone in the chest pocket [1], [2]. The estimation method of the Yo-bukun involves utilizing smartphone sensors such as the acceleration sensor and gyro sensor to assess the user's state (standing, sitting, turning, crouching, walking). Subsequently, the user's physical information is incorporated, and the estimation formula used in MIYADAI Pro developed by Professor Mitarai [3] is applied to perform the estimation. Here, the user needs to set the user's physical information and the weight of the object being held by the user before taking measurements. During measurement, the lumbar burden is estimated at every second in real-time.

However, few constraints have been identified in its functionality. The initial challenge is the inability to determine whether the user is actively holding an object during measurement or not. Therefore, throughout the measurement the lumbar burden is approximated under the assumption that the user is consistently holding an object. As a result, the Yo-bukun can't estimate the lumbar burden regarding movements occurring before the user holds an object or after putting an object down.

The second issue arises upon the completion of measurement, in which the user has not been given the capability to modify the set information and reevaluate the lumbar burden. In that case, if a measurement has already been conducted with incorrect information, it demands repeating the same measurement. Therefore, this paper outlines the functionalities introduced to address and resolve these identified challenges.

2. Methodology

This paper describes the proposed method that can be resolved only using a smartphone, emphasizing user-friendly ease of operation.

2.1. First problem-solving method

To address the problem of measurement being possible only in the limited situation where the user holds an object, a voice recognition function has been introduced to the Yo-bukun. The user articulates speech, while initially holding an object and when releasing it, and this voice information is utilized to determine the object possession status. The judgment method can be described as follows: upon detecting the initiation of object holding, the system assumes continuous possession until the

Seigo Imura, Praveen Nuwantha Gunaratne, Hiroki Tamura

moment of object release is identified. Conversely, if the system detects the timing when the user puts down an object, it assumes that the user does not hold an object until the point of object holding initiation is recognized. In this way, the status of object possession can be determined using voice recognition.

2.2. Second problem-solving method

The limitation was explained wherein, the post-measurement, user is unable to modify pre-measurement information and recalculate lumbar burden again. Utilizing the estimation formula utilized in Yo-bukun, the key parameters for lumbar burden estimation, extending beyond the pre-measurement information provided by the user were identified. These additional parameters are recorded for each lumbar burden estimation. Consequently, even if the user modifies the pre-measurement information after the completion of measurement, the proposed system facilitates the use of recorded parameter data for any recalculations of lumbar burden.

3. Implementation Procedure
3.1. First problem-solving Procedure

With the use of Apple's speech recognition framework [4], a voice recognition feature was incorporated with the Yo-bukun and developed a system that assesses the lumbar burden based on the user's object possession status. This advancement enables the estimation of lumbar burden solely through a smartphone, considering the user's object possession status. The proposed system estimates lumbar burden based on a user-selected preset weight if an object is held. In instances where the user is not holding an object, the system assumes an object weight of 0 kg. Furthermore, this approach replaces traditional button operations as the commencement (Start) and cessation (End) of measurement have been automated through voice recognition.

3.2. Second problem-solving Procedure

A review was conducted on the fundamental parameters crucial for lumbar burden estimation and the system development was carried forward for their systematic recording during measurements. Moreover, the system was facilitated with the use of recorded parameters to reassess lumbar burden subsequent to the measurement. As a result, users possess the capability to modify pre-measurement information, enabling recalculation of lumbar burden based on these alterations.

4. Experiment

Two experiments were conducted to validate the functionality of the introduced system.

4.1. First experiment

In the first experiment, Yo-bukun was employed for lumbar burden assessment, using specific verbal cues during both the Start and End of the object handling process. The weight of the object was set at 2 kg. To ensure the optimal accuracy in estimation result, the smartphone was securely affixed to a vest, positioned in close proximity to chest area. Initially, the configuration was set for the subject to commence the measurement without holding an object as the Yo-bukun also offers configurability to specify the status of holding an object or not, at the initialization. The Fig. 1 depicts the following operational sequence proceeded during the experiment:

a) Start the measurement without any object.
b) Uttered a specific word and lifted an object after 11 seconds.
c) Raised the upper half of the body.
d) Safely place the object on the floor by uttering a word after 21 seconds.
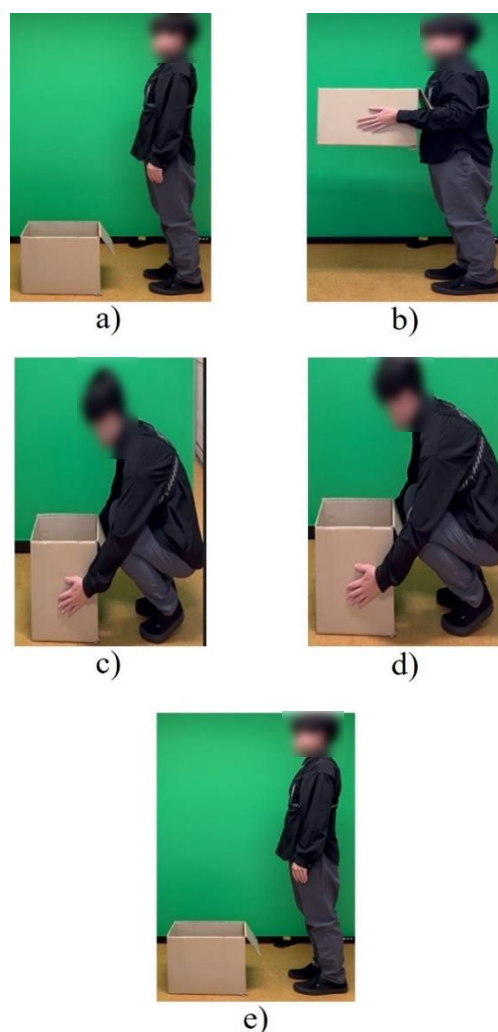e) End the measurement by raising the upper half of the body.

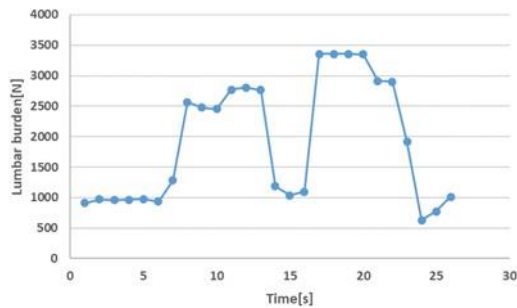Fig. 1. Operation sequence of the first experiment

Fig. 2. Graph illustrating lumbar burden estimation during the first experiment.
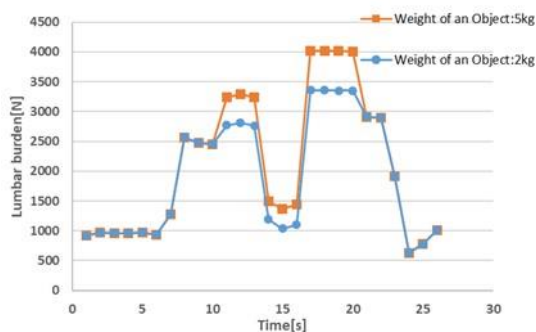

Fig. 3. Graph showing the change in lumbar burden estimation during the second experiment; graph for the object weight 2kg (blue), 5kg object (orange)

4.2. Second experiment

The second experiment was conducted to reassess the estimation results obtained from the initial trial by modifying user-input information. In this subsequent trial, an examination was made to determine the impact on estimation results by altering the weight of an object to 5 kg.

5. Experiment result
5.1. First experiment result

The experimental results are illustrated in Fig. 2. The first experiment result outlines the procedure for determining whether the presence or absence of an object can be considered. During the measurement, the Yo-bukun is modified to enable the recording of user-predefined information and the other parameters essential for estimating lumbar burden. Consequently, the recorded parameters included the weight of an object, with the weight being logged every second. From this, it is evident that the Yo-bukun assesses lumbar burden based on the criteria that a weight of 0 kg indicates the person is not holding an object, whereas a weight of 2 kg implies the person is holding an object. The first experiment result revealed that the weight was 0 kg for the initial 10 seconds, 2 kg from seconds 11 to 20 Sec., and returned to 0 kg thereafter until the end of the

measurement. Therefore, it was deduced that lumbar burden could be estimated through voice recognition, considering the presence or absence of an object.

5.2. Second experiment result

The reevaluated results, depicting changes on object weight from 2 kg to 5 kg, are illustrated in Fig. 3. During different measurements when no object is held, the recorded weight is maintained at 0 kg, yielding consistent measurements. Subsequently, the object's weight is adjusted to 5 kg, and recalculation is executed during object holding. As a result, no variations were observed during the intervals from 0 to 11 seconds and from 21 Seconds until the conclusion of the measurement when the person was not holding an object. The fluctuations in results were only evident from the 12 to 20 seconds time interval when the person was holding an object. Therefore, the results of the second experiment illustrate the possibility of parameter adjustments and post-measurement recalculations.

6. Conclusion

Through the integration of voice recognition into Yo-bukun, the system is now capable of estimating lumbar burden, considering whether the user is holding an object or not. The start and end of measurements which were previously reliant on manual button press within the application, can now be automated via voice recognition. By recording the parameters essential for lumbar burden estimation during the measurement, users have the flexibility to modify pre-set information even after completing the measurement, allowing the re-calculation of lumbar burden. The implementation of this system has also simplified the prospective studies of how changes in object weight or physical information during the same movement influence lumbar burden.

References

1. Hiroki Tamura, Keike Sakurai, Koichi Tanno, Yasufumi Fuse, "A Study on the Lumbar Burden Evaluation of Work using One Smartphone," Journal of Robotics, Networking and Artificial Life 5(3) (2018) 173-179
2. Mizuki Maiguma, Hiroki Tamura, Koichi Tanno, "A Study on the Lumbar Burden Evaluation of Work using One Smartphone," The 2018 International Conference on Artificial A Life and Robotics 1-4
3. K. Mitarai, "Software Resources at Miyazaki University," Miyazaki University. [Online]. Available: https://www.cc.miyazaki-u.ac.jp/mitarai/room707/soft%20MIYADAI.html.
4. Apple Inc., "Speech Framework," Apple Developer. [Online]. Available: https://developer.apple.com/documentation/speech?language=objc.

Authors Introduction

Mr. Seigo Imura

Seigo Imura born in 2001. He is currently studying in department of environmental robotics, and will receive the B.Eng from University of Miyazaki in 2024.

His current research is Development of Smartphone Application for Calculating the Low Back Pain Risk.

Mr. Praveen Nuwantha Gunaratne

He received his bachelor's degree in engineering in 2018 from the Faculty of Engineering, University of Moratuwa, Sri Lanka. He is currently a doctoral student in University of Miyazaki, Japan.

Prof. Hiroki Tamura

He received the B.E. and M.E. degree from Miyazaki University in 1998 and 2000, respectively. From 2000 to 2001, he was an Engineer in Asahi Kasei Corporation, Japan. In 2001, he joined Toyama University, Toyama, Japan, where he was a Technical Official in the Department of Intellectual Information Systems. In 2006, he joined Miyazaki University, Miyazaki, Japan, where he was an Assistant Professor in the Department of Electrical and Electronic Engineering. Since 2015, he is currently a Professor in the Department of Environmental Robotics. His main research interests are Neural Networks and Optimization Problems. In recent years, he has had interest in Biomedical Signal Processing using Soft Computing.

A Computational Approach for Global Trade Analysis Sensitive to Free Trade Agreement Circumstances: A Case Study Focusing on the Great Mekong Subregion

Ahmad Altaweel
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu-ku, Kitakyushu, 808-0196, Japan
Email altaweel.ahmad770@mail.kyutech.jp

Bo-Young Lee
Logistics Revolution Korea Co., Ltd.
5th FI. 47, Gangnam-daero 101-gil, Seocho-gu, Seoul 06034, Republic of Korea
Email: bylee@logisroad.com

Masayuki Fujiwara
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu-ku, Kitakyushu, 808-0196, Japan
Email: m-fujiw@brain.kyutech.ac.jp

Jang-Sok Yoon
Logistics Revolution Korea Co., Ltd.
5th FI. 47, Gangnam-daero 101-gil, Seocho-gu, Seoul 06034, Republic of Korea
Email: kjsyoon@logisroad.com

Hiroaki Wagatsuma
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu-ku, Kitakyushu, 808-0196, Japan
Email: waga@brain.kyutech.ac.jp

Abstract

In the Global Trade Analysis Project (GTAP), GDP and economic statistical indices can be analyzed for forecasting future trends in multiple countries by using the GTAP database and GEMPACK utilities, which realize the numerical calculation based on the concept of Computable General Equilibrium (CGE) models. Even with such well-prepared tools with the official database, appropriate forecasting is still difficult due to the sensitivity of the Free Trade Agreement (FTA) circumstances. FTA scenarios with a uniform tariff reduction can be assumed in general, while an abrupt and unreasonable change may occur in the simulation depending on the network structure of trading countries and the upper and lower bounds of tariffs in the time course. In the present study, we focused on the Great Mekong subregion (GMS) and explored possible methods to calculate substantially.

Keywords: computable general equilibrium (CGE), General Equilibrium Modelling PACKage (GEMPACK), Global Trade Analysis Project (GTAP), Free trade agreements (FTA).

1. Introduction

Computable General Equilibrium (CGE) analysis has emerged as a competitive tool for policy evaluation. Its widespread application includes research on trade policy, regional infrastructure development, and environmental protection. However, challenges in the technical implementations explained in [1], [2], [3], data preparation, and interpreting the simulation results pose significant burdens in the current state of CGE modeling.

To lower this technical barrier involved, this research investigates the roadmap to handle the implementation of the GTAP model. It discusses preparing data and measures the effect of altering the growth rates and reduction in the tariff value. Moreover, it forecasts the changes in the real GDP among the countries in the Great Mekong subregion (China, Lao-PDR, Myanmar, Vietnam, Thailand, and Cambodia).

2. Methodology

Ahmad Altaweel, Bo-Young Lee, Masayuki Fujiwara, Jang-Sok Yoon, Hiroaki Wagatsuma

2.1. *Simulation environment*

2.1.1. *General Equilibrium Modelling PACKage (GEMPACK)*

One of the fundamental characteristics of the CGE models is their computational capability, specifically, they can generate numerical solutions to shocks (changes) using the provided database and parameters. GEMPACK short for General Equilibrium Modelling Package is a set of software created for solving applied general equilibrium models. This package comprises various tools for managing economic models, databases, and simulations. In GEMPACK, data is typically presented in the format of HAR files, which are header-array files capable of storing multiple numeric and string arrays with dimension labels, accommodating up to seven dimensions. The outcome of the simulation is also multiple HAR files or solution files (e.g., .slc, .sl4) can be saved in header files.

2.1.2. *Modifying header files*

It is often required to adjust the input parameters, shock files, and process the model's outputs, and that involves modifying the data within header files. The conventional approach for handling HAR files has been to employ executable programs provided by GEMPACK (e.g., seehar.exe or modhar.exe). However, utilizing these programs may pose a challenge. Manual transfer of data between HAR files and data editing software (e.g., Excel) using a basic clipboard method is also a common practice. However, this process is highly manual, time-consuming, and prone to errors. Another approach for modifying HAR files is to create a data manipulation TABLO Input file using ViewHAR. That can be done by exporting the header file as a TABLO Code from the Export menu within the ViewHAR program. This process automatically generates the foundational code, which can be further edited using the Tablo language to include specific filenames, formulas, updates, assertions, and other adjustments tailored to the model's requirements. In addition, this approach requires creating a Command file (CMF) to define the input, output, and auxiliary file. This can also be generated automatically through the menu (view/Create CMF) within the Tablo file (.tab file).

2.1.3. *Executing a model in GEMPACK*

GEMPACK compiles a model by interpreting a model written in the TABLO language into a system of equations and generating an executable file (e.g., 'gtapv7.exe'). Users must specify the endogenous and exogenous variables, along with the chosen solution method (e.g. Gragg, Johansen, Euler, 1,3,5 steps) in the CMF file. Subsequently, the model is executed, and the coefficients of the model are filled with the actual data. The simulation generates a set of files (.tab, .cmf, .exe, .sl4, .slc, .har) associated with the model. To retrieve the solution, users can utilize ViewSol, AnalyzeGE, or ViewHAR to access data in '.slc' and '.sl4' files.

2.2. *Data and Model Preparation*

2.2.1. *GTAP Database*

The GTAP-RD model is calibrated to the most recent database produced by the Global Trade Analysis Project GTAP database version 11 [2]. The database GTAP 11 reconciles different data sources at a global scale and offers a time series of 5 reference years (2004, 2007, 2011, 2014, and 2017), distinguishing 65 sectors in each of 141 countries and 19 aggregate regions [4]. The GTAP database is typically supplied in the default format of the HAR file (flows file, parameters, sets,..etc). This database has been precisely adjusted to achieve equilibrium in the world's initial state [5].

2.2.2. *The GTAP-RD model*

This study uses GTAP-RD, a dynamic multi-region model that is described in detail in [6]. GTAP-RD model was built on the latest GTAP 7 framework [7] and offers features including, capital accumulation, and international capital and investment mobility. In addition, GTAP-RD offers flexibility in terms of closures as well as a more convenient way to perform a modification to the model due to consistent variables, coefficients, and equation-naming conventions.

2.3. *Forecasting Process*

Quantitative policy analysis is typically framed within the confines of a specific scenario (numerical projection). This scenario assumes that the economy will be affected by future adjustments in the macroeconomic forecasts of each country and the expected policy changes. The baseline scenario provides a reference point for evaluating the impact of policy changes [3] and should reflect as closely as possible the changes expected to occur in the world economy. For example, FTAs that are already in effect [8]. In most baselines, external information about certain future trends of the macroeconomic drivers, for example, population, labor force growth, and other drivers are taken from external projections, such as the United Nations Population Division, World Bank, IMF, OECD, CEPII…, etc.). The forecasts gathered from the mentioned sources may lack

5	vnm	Vietnam	18	sas	Oceania
6	mmr	Myanmar, East Timor	19	eas	East Asia
7	phl	Philippines	20	nwa	North America
8	idn	Indonesia	21	lam	Latin America
9	mys	Malaysia	22	wer	Western Europ
10	brn	Brunei	23	men	Middle East
11	sgp	Singapore	24	ssa	Sub-Saharan and North Africa
12	jpn	Japan	25	row	Rest of World
13	kor	South Korea			

The analysis begins with the 2017 GTAP database serving as the base year (Fig. 1). It shifts forward by three years to establish 2020 as the benchmark year. Subsequently, for each period, there is a five-year increment until reaching the year 2050 in all scenarios.

Fig. 1. Forecasting Process and sequence of years

We utilized the GTAP-RD aggregation utility database [10] by making the necessary adjustments to align it with GTAP database version 11. This involved modifying the region mapping from the original 230 countries to fit the 160 countries of the GTAP database and aggregate years to match the relevant years in the study.

Scenario 1 (High Case): The SSP5 scenario percent change for each of real GDP (qgdp), Population (pop), skilled and unskilled labor qe(ENDWL,REG) is applied as target shocks to the years from 2025~2050. In addition, import tariff within the six regions in the GMS area was reduced gradually in the S1 Scenario but maintained between the rest of the world and GMS regions. The reduction periods are as follows (Table 2):

Table 2. Tariff Reduction Schedule between the GMS countries.

Period	Percent change
(2020~2025)→	-20%
(2025~2030)→	-20%

information and may not align. This necessitates some processing to guarantee the availability of data for all countries and relevant years. Shaping a baseline is important for long-term economic issues, such as climate change, the Intergovernmental Panel on Climate Change (IPCC) has created standard scenarios for future greenhouse gas (GHG) emissions so-called Shared Socioeconomic Pathways (SSPs). The SSPs are five scenarios that form a consistent set of socioeconomic drivers widely used in the climate modeling community, and also in the broader global modeling community [3].

2.3.1.   *Simulation design and policy scenarios*

The baseline projection runs from 2017 to 2050 and includes past data and projections for growth in real GDP, population, and the skilled and unskilled labor supply for all regions. The baseline projection data comes from the Shared Social-Economic Pathways (SSP) database macro forecasts for 230 countries for the period of 2007-2100. This paper uses the SSP2 projections for the baseline (business-as-usual scenario). In the baseline scenario, we did not incorporate policy changes caused by existing and ongoing FTAs. Absence of the policy changes in the baseline may affect the simulation results of the policy scenarios [9]. The baseline projection is run under the default GTAP-RD closure for all regions and to make the GTAP model follow a chosen growth path, real GDP (qgdp) is swapped with the region-wide technological change (afereg) to be calculated endogenously. In subsequent model runs, technological change are then maintained at those estimated levels, while GDP is endogenously computed in the base rerun.

The study focuses on global logistics in Southeast Asia and the Great Mekong subregion, and therefore, countries in this area are the targets of the analysis and are treated as individual countries. Countries in other regions are consolidated into regional units, with each treated as a hypothetical country in the GTAP Model.
We aggregated the GTAP database to 25 countries/regions and 10 sectors (Table 1) The region of the rest of Southeast Asia contains Myanmar and East Timor after making all the other countries in this region as individual countries.

Table 1. Regional aggregation of the GTAP Data Base.

NO	ID	COUNTRY/REGION	NO.	ID	COUNTRY/REGION
1	lao	Lao-PDR	14	ind	INDIA
2	chn	China	15	aus	India
3	tha	Thailand	16	nzl	Australia
4	khm	Cambodia	17	xoc	New Zealand

Ahmad Altaweel, Bo-Young Lee, Masayuki Fujiwara, Jang-Sok Yoon, Hiroaki Wagatsuma

(2030~2035)→	-20%
(2035~2040)→	-40%

It is improbable that an agreement would lead to the complete removal of all import barriers, and therefore, this experiment provides an upper bound for the benefits of a free trade agreement that can be captured by the model.

To control the change in the tariff, we used the "rate%" function in the CMF file of the shocks as a policy file after associating the new levels variable TMS_L (power of import tariff) Eq. (1) with the previously declared linear variable tms in the GTAP model [2].

$$TMS\_L(c,s,d) = VMSB(c,s,d)/VCIF(c,s,d) \quad (1)$$

VMSB: the value of imports of commodity c from s to d at domestic (basic) prices.
VCIF: the value of imports of c from s to d at CIF prices (tradeable only).

After adding the level variable TMS_L, it is possible to model the removal of all import taxes by utilizing the "final_level" function Eq. (2) from GEMPACK as follows:

$$Final\_level \ TMS\_L(COMM,GREG,GREG) = uniform \ 1 ; \quad (2)$$

This function will remove the import taxes; thereby ad valorem (tariff rate) is equal to zero and the power level of the tariff is equal to one from the year 2040 in the S1 scenario. The uniform reduction in the tariff rate is applied to all commodities within the six countries and regions in the Great Mekong subregion by making a new region set GREG in the model to contain the GMS countries.

Scenario 2 (Low Case): The SSP4 scenario for the percent changes in the real GDP, population, and labor force are implemented for the years 2025~2050. This scenario takes a pessimistic stance on the advancement of a trade agreement, envisioning no reduction, and predicts a modest growth rate.

3. Results and Discussion

3.1. *Estimated Gross domestic product growth*

The estimated values of real GDP are displayed in Fig 2. As described in Section 2.3.1, It is exogenously given in the baseline scenario. Looking at the high case the increments in the value of GDP come from the benefit of changes in the tariff reduction and the increase in the total

factor of production growth rate in the remarkable countries.

Fig.2. GDP value change in GMS countries

4. Conclusion

In this study, the authors employed the GTAP model to estimate the future GDP value in the Great Mekong subregion area. The study delves into the technical intricacy of implementing modifications in the growth rate and reduction of the tariff rate. It simulates a free trade agreement among the involved countries.

Subsequently, the authors intend to expand the range of Free trade agreements, encompassing the agreements currently in force in the baseline scenario. Additionally, they plan to refine the calibration to the growth rate in the GMS area.

Acknowledgment

References

1. K. Pearson, M. Horridge, and E. Corong, "Hands-on Computing with RunGTAP and WinGEM To Introduce GTAP and GEMPACK.
2. M. Horridge, "GEMPACK manual." Dec. 17, 2019.
3. R. Dellink, D. Van Der Mensbrugghe, and B. Saveyn, "Shaping baseline scenarios of economic activity with CGE models: introduction to the special issue," *J. Glob. Econ. Anal.*, vol. 5, no. 1, pp. 1–27, Jun. 2020, doi: 10.21642/JGEA.050101AF.
4. A. Aguiar, M. Chepeliev, E. Corong, and D. Van Der Mensbrugghe, "The Global Trade Analysis Project (GTAP) Data Base: Version 11," *J. Glob. Econ. Anal.*, vol. 7, no. 2, pp. 1–37, Dec. 2022, doi: 10.21642/JGEA.070201AF.

5. M. Ivanic, "GEMPACK simulations in R: Ademonstration of running the GTAP modeland processing its results entirely in R using packages HARr and tabloToR," *J. Glob. Econ. Anal.*, vol. 8, no. 1, pp. 1–20, Jun. 2023, doi: 10.21642/JGEA.080101AF.

6. A. Aguiar, Corong, Erwin, and van der Mensbrugghe, Dominique, "The GTAP Recursive Dynamic (GTAP-RD) Model: Version 1.0," *Present. GTAP-RD Model GTAP Conf. Organ. Sess. OnImproving Baseline Proj. GTAP-Recursive Dyn. RD Model Framew.*, Feb. 2019.

7. E. Corong, H. Thomas, M. Robert, M. Tsigas, and D. Van Der Mensbrugghe, "The Standard GTAP Model, version 7," *J. Glob. Econ. Anal.*, vol. 2, no. 1, pp. 1–119, Jun. 2017, doi: 10.21642/JGEA.020101AF.

8. K. Itakura, "Impact of the Regional Comprehensive Economic Partnership (RCEP): A Global Computable General Equilibrium (CGE) Simulation," *Econ. Res. Inst. ASEAN East Asia*, Oct. 2022.

9. K. Itakura, "Impact of liberalization and improved connectivity and facilitation in ASEAN," *J. Asian Econ.*, vol. 35, pp. 2–11, Dec. 2014, doi: 10.1016/j.asieco.2014.09.002.

10. GTAP-RD Data Aggregation Utility.

Authors Introduction

Mr. Ahmad Altaweel

He received his Bachelor's degree in Electronics and Communication Engineering, from the Faculty of Electrical and Mechanical Engineering, ALBAATH University, Syria in 2013. He is currently a Ph.D. student at Kyushu Institute of Technology, Japan.

Ms. BoYoung Lee

She received her Bachelor's degree in Human Relations from Keio University in Japan. She then received her M.S. in logistics from Inha University in Korea and is currently a Ph.D. student. She works as a logistics consultant in the Logistics Revolution Korea Co., Ltd.

Dr. Masayuki Fujiwara

He graduated from the Fukuoka Institute of Technology(FIT) in 2015. He then went on to earn his M.Sc. and Ph.D. (Knowledge Science) from Japan Advanced Institute of Science and Technology(JAIST) in 2017 and 2023. Until 2020, he was a Research Fellow of the Japan Society for the Promotion of Science (JSPS) (DC1) and now is a researcher at Kyushu Institute of Technology(Kyutech). His research focuses on the EEG studies and neural modeling of (non-)symbolic communication from the perspective of cognitive neurodynamics.

Dr. JangSok Yoon

He received his M.S. and Ph.D. degrees from Kyunghee University, in South Korea, in 2004 and 2008. He is the founder, CEO, and Consultant over 23 years of Logistics Revolution Korea Co., Ltd He has served as a logistics and supply chain consultant to more than 300 companies and governments.

Dr. Hiroaki Wagatsuma

He received his M.S., and Ph.D. degrees from Tokyo Denki University, Japan, in 1997 and 2005, respectively. In 2009, he joined the Kyushu Institute of Technology, where he is currently an Associate Professor of the Department of Human Intelligence Systems. His research interests include non-linear dynamics and robotics. He is a member of IEEE.

Trigger circuit design and system integration
for simultaneous measurement of human EEG, motion, and gaze

Masayuki Fujiwara, Phan Hoang Huu Duc
Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku, Kitakyushu, 808-0196, Japan
Email: m-fujiw@brain.kyutech.ac.jp, phan.duc-hoang516@mail.kyutech.jp

Laurent Bougrain, Patrick Hénaff
Université de Lorraine, CNRS, LORIA, F-54000 Nancy, France
Email: {laurent.bougrain, patrick.henaff}@loria.fr

Hiroaki Wagatsuma
Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku, Kitakyushu, 808-0196, Japan
Email: waga@brain.kyutech.ac.jp

Abstract

Simultaneous measurement of EEG, motion, and gaze in humans has the potential to lead to the discovery of new scientific insights. In order to achieve these simultaneous measurements, it is necessary to manage triggers and time information between measurement devices, as well as to correct time offsets. However, the management of accurate triggers and time information requires the design of a dedicated circuit board and the integration of Transistor-Transistor Logic signal voltage information. In this study, we report on the building of a trigger circuit and an experimental system using it to solve these problems. We created a home-made trigger circuit board for voltage integration and combined it with a commercially available microcomputer to realize an integrated trigger circuit and measurement system.

Keywords: Trigger circuit design, Electroencephalography (EEG), Motion, Gaze

1. Introduction

Simultaneous recording of physiological and behavioral data in standard data formats and analysis of these data in combination may lead to new scientific discoveries. So, in recent years, multimodal measurement methods and data formats have been attracting attention. In response to this situation, data acquisition through the Lab Streaming Layer (LSL) [1], new standard data formats, such as Brain Imaging Data Structure (BIDS) [2], [3], Neuroscience Information eXchange [4] and Open metaData Markup Language [5] (NIX-odML) [6], have been proposed to integrate and handle multimodal data and meta information.

For these simultaneous recordings, an experimenter is faced with the necessity to manage the trigger and time information between measuring instruments and to compensate for time offsets. In fact, while there are paid software and hardware types of equipment that manage the trigger and time information, it is so challenging to handle between measuring instruments that do not support such information.

The simultaneous measurement of various types of information, such as movement and gaze, in conjunction with human electroencephalography (EEG), holds particular significance due to their intimate involvement with neural activities. As introduced by Iturrate et al. [7], in the case of humans, the information derived from the gaze is directed to the V1 area within a span of 60-90 ms, reaches the motor cortex in approximately 220 ms, and the expression of fingertip movements is processed with a latency roughly between 280-400 ms (it is noteworthy that the process is expedited in monkeys [8]). Measuring and analyzing brain activities during this process is beneficial to understanding human cognitive functions.

In the previous work, we have proposed a methodological design [9] to reveal human-robot interactions in the real world by managing Transistor-Transistor Logic (TTL) signals. This proposal provided a specific event analysis based on data from multiple instruments and contributed to a reproducible experimental scheme in which each element could be interchangeable in the future.

However, to achieve accurate trigger and time information to be managed, a dedicated and versatile circuit board and its system must be designed to integrate

Masayuki Fujiwara, Phan Hoang Huu Duc, Laurent Bougrain, Patrick Hénaff, Hiroaki Wagatsuma

Fig. 1 Recording System Integration for human EEG, motion, and gaze data synchronization.

the voltage information of the TTL signal.

Therefore, this study reports on the fabrication of a trigger circuit and an experimental system using it to solve the problems mentioned above. We built a trigger circuit board for voltage integration and combined it with a commercially available microcontroller to integrate the trigger circuit and the measurement system.

2. Proposed Trigger circuit design and system integration s

2.1. *Recording System Integration*

The integrated measurement system of our recording instruments is shown in Fig. 1. The system uses a home-made trigger circuit and a microcomputer for integrated management of the voltage and time information output from each measurement instrument with and without trigger output function. The following is a description of the actual measurement instruments that were used:

- g.tec USBamp and g.tec GAMMAbox for EEG recording;
- Perception Neuron 2.0 for Motion Capturing;
- Tobii Pro Glasses 2 for Eye Tracking.

Perception Neuron 2.0 does not have a trigger output function. Nevertheless, the trigger and time information can be obtained by combining a trigger circuit and a microcomputer, as described in the next section. Note that the timing of the first output of the trigger signal

depends on detecting the pressing of the Axis Neuron's recording button using software for the perception neuron with a global hook function on Windows.

2.2. *Home-made Trigger Circuit and Microcomputer*

Fig. 2 shows the fabrication and design of the trigger circuit board. We used Autodesk Fusion 360 [10] to unify circuit design (Fig. 2a) and board design (Fig. 2b). The designed circuit was fabricated to be integrated with the microcomputer (Fig. 3). Moreover, this circuit has two input ports (blue and yellow cables), but multiple input ports can be managed by increasing the number of level-shift circuits. For measurement instruments that do not have a trigger output, a timer function in the microcomputer is used to output a square wave (e.g., 5.00 [V]) with accuracy in the order of microseconds. The trigger and synchronization are managed by adding/subtracting the above-mentioned multiple trigger inputs to/from this voltage. If other measurement equipment has a trigger output, it can be integrated after checking the output voltage.

2.3. *Testing Equipment information*

We observed the trigger signals using a digital oscilloscope to test the integrated trigger circuit board and its recording system. Actually, we connected multiple measurement instruments to the input ports of the trigger circuit to capture the changes in the trigger

a.   Schematic circuit diagram

b.   Circuit board diagram

Fig. 2 Fabrication and Design of Trigger Circuit Board

Fig.3 Integrated Trigger Circuit Board

a.   Only Integrated Trigger Circuit Board

b.   Combined with Other Measuring Instruments

Fig.4 Recorded Trigger Signals

signal. The settings used for observation were an amplitude of 2.00 [V]/scale and a period of 100 ms/scale.

3. Results and Discussion

We measured the trigger signal from the integrated trigger circuit board (Fig. 4). Fig. 4a shows the case in which only the trigger circuit was used, and a square wave (5.00 [V]) was measured. Fig. 4b shows a simultaneous measurement, including trigger signals from other measurement devices and the timing of multiple trigger inputs. These recorded trigger signals showed the management we expected.

As a limitation, if the measurement start time of the measurement device depends on the recording software, the accuracy of the trigger may be affected by the system clock of the Operating System (OS), leaving room for further study. Currently, in this case, a dedicated application is developed and used to output the trigger from the Universal Serial Bus port to detect the measurement start time, which may cause an accuracy problem.

4. Conclusion

We designed a custom trigger circuit board to measure EEG, motion, and eye movement simultaneously. This trigger circuit board and its system integrate the voltage information of multiple TTL signals, allowing for accurate triggering and time synchronization between different measurement instruments. It also may suggest a potential way to provide data from measurement instruments that do not have trigger output functions with respect to existing measurement methods for integrating multimodal data.

Future work is to investigate the accuracy of triggering at system time, etc., when relying on measurement software., i.e., the management of time information between the OS and the trigger circuit.

Masayuki Fujiwara, Phan Hoang Huu Duc, Laurent Bougrain, Patrick Hénaff, Hiroaki Wagatsuma

Acknowledgements

References

1. The lab streaming layer, https://github.com/sccn/labstreaminglayer
2. K. Gorgolewski, T. Auer, V. Calhoun, et al. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. Scientific Data, 3, 2016, 160044.
3. C. R. Pernet, S. Appelhoff, K. J. Gorgolewski, et al. EEG-BIDS, an extension to the brain imaging data structure for electroencephalography. Scientific Data, 6, 2019, 103.
4. A. Stoewer, C. J. Kellner, J. Benda, T. Wachtler and J. Grewe, File format and library for neuroscience data and metadata, Frontiers in Neuroinformatics. Conference Abstract: Neuroinformatics 2014.
5. J. Grewe, T. Wachtler, and J. Benda, A bottom-up approach to data annotation in neurophysiology, Front. Neuroinformatics, 5(16), 2011.
6. nix-odML-converter, https://github.com/G-Node/nix-odML-converter
7. A. Grimaldi, A. Gruel, C. Besnainou, J. N. Jérémie, J. Martinet, L. U. Perrinet, Precise spiking motifs in neurobiological and neuromorphic data, Brain sciences, vol.13 no.1, 68, 2022.
8. S. J. Thorpe and M. Fabre-Thorpe, Neuroscience. Seeking categories in the brain. Science, vol.291 no.5502, pp.260–263, 2001.
9. M. R. V. Sanchez, S. Mishima, M. Fujiwara, G. Ai, M. Jouaiti, Y. Kobryn, S. Rimbert, L. Bougrain, P. Hénaff, H. Wagatsuma, Methodological Design for Integration of Human EEG Data with Behavioral Analyses into Human-Human/Robot Interactions in a Real-World Context, ICIC Express Letters, 14(7), pp.693-701, 2020.
10. Autodesk Fusion 360, https://www.autodesk.co.jp/products/fusion-360/overview

Authors Introduction

Dr. Masayuki Fujiwara

He graduated from Fukuoka Institute of Technology (FIT) in 2015. He then went on to earned his Master and Ph.D (Knowledge Science) degrees from Japan Advanced Institute of Science and Technology (JAIST) in 2017 and 2023. Until 2020, he was a Research Fellow of the Japan Society for the Promotion of Science (JSPS) (DC1) and is now a researcher at Kyushu Institute of Technology (Kyutech). His research focuses on the EEG studies and neural modeling of (non-)symbolic communication from the perspective of cognitive neurodynamics.

Mr. Phan Hoang Huu Duc

He received Bachelor degree in Biomedical Engineering, from Department of Biomedical Engineering, Faculty of Applied Science, Ho Chi Minh City University of Technology (HCMUT) in 2021. He is currently studying Neuroscience in Kyushu Institute of Technology from 2022.

Dr. Laurent Bougrain

He earned his MSc in artificial intelligence from université Paris 6, his BS in psychology from université Nancy 2 and his Ph.D in Computer Science from Université de Nancy 1 in 2000. He is leading the Neurosys/Neurorhythms team since 2015 at LORIA. His research focuses on brain-computer interfaces, motor systems, machine learning, feature extraction and selection, post-stroke rehabilitation, artificial intelligence, and monitoring of general anesthesia. He is currently an Associate Professor of Computer Science at the University of Lorraine.

Dr. Patrick Hénaff

He earned his M.Sc and Ph.D in Robotics from Université Pierre et Marie Curie in 1989 and 1994. He worked as a lecturer at the University of Cergy-Pontoise and a researcher at LISV (Systems Engineering Laboratory at the University of Versailles) and the University of Cergy-Pontoise. Since 2013, he has been a Professor at the Ecole des Mines de Nancy, University of Lorraine, and worked in the LORIA laboratory. He led the Complex Systems, AI, and Robotics department at LORIA from 2015 to 2022, and the Computer Science department at Mines Nancy from 2018 to 2023.

Dr. Hiroaki Wagatsuma

He received his M.S., and Ph.D. degrees from Tokyo Denki University, Japan, in 1997 and 2005, respectively. In 2009, he joined Kyushu Institute of Technology, where he is currently an Associate Professor of the Department of Human Intelligence Systems. His research interests include non-linear dynamics and robotics. He is a member of the IEEE society.

Terminal Synergetic Controller for Car's Active Suspension System Using Dragonfly Algorithm

Tinnakorn Kumsaen
Department of Chemical Engineering, Faculty of Engineering, Khon Kaen University, Khon Kaen, 40002, Thailand

Sorn Simatrang
Nacres Co., Ltd, Bangkok 10150, Thailand

Arsit Boonyaprapasorn
Department of Mechanical Engineering, Chulachomklao Royal Military Academy, Nakhon-Nayok, 26000, Thailand

Thunyaseth Sethaput
Sirindhorn International Institute of Technology, Thammasat University, Pathum Thani, 12120, Thailand
Email: tinnakor@kku.ac.th, urarl.urarl9@gmail.com, robosorn@gmail.com, thunyaseth@siit.tu.ac.th

Abstract

This research introduces a terminal synergetic controller (TSC) designed for the active suspension system of automobiles through the implementation of the dragonfly algorithm (DA). The proposed controller aims to enhance the dynamic performance of a car's suspension using the DA in tuning the system parameters. The stability of the designed controller is proved through the application of Lyapunov stability theory. Through iterative optimization processes, the TSC approach seeks to achieve an optimal balance between ride comfort and vehicle handling. The simulation results demonstrate that the proposed controller enhances convergence properties and alleviates the presence of chattering. The results indicate that the proposed approach with the optimal parameters provided insights into its potential application in improving the overall suspension system.

Keywords: Active suspension, Synergetic control, Feedback control, Metaheuristic; Swarm-based optimization

1. Introduction

Since the primary objective of controlling active suspension systems is disturbance rejection, in this control application, it is imperative to select a potential feedback control method that can effectively counteract disturbances. Even though the robustness and enhancement of the convergence of the control system can be achieved by using the concept of sliding mode control (SMC), the method has a key drawback from the chattering phenomenon. This drawback can be avoided by using the synergetic control (SC) which was proposed by Kolesikov et al. [1], [2]. The SC method allows the designer to achieve the desired characteristics of the control system including parameter insensitivity and robustness if the macro variables are selected properly. Sliding surfaces in sliding mode control can be utilized as macro variables in the SC method. Consequently, the convergence time of the control system under the SC method can be improved by choosing macro variables and/or dynamic evolutions with terminal attractors [3]. Swarm-based optimization algorithms, including the dragonfly algorithm (DA) [4], enable the optimal selection of controller parameters. The benefit of achieving an improved convergence rate is obtained when a feedback controller is combined with the DA [5].

In this study, we present the TSC method specifically designed for the active suspension system of automobiles. Based on the quarter-car model, the primary objective of

the proposed controller is to improve the stability of the car by utilizing the DA for the precision tuning of system parameters.

2. Methodology

Lemma 1: Consider a nonlinear system in Eq. (1):

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t)), \qquad (1)$$

where $\mathbf{x}(t)$ denotes a state vector and $\mathbf{x}(\cdot): \square \to \square^n$,. $\mathbf{f}(\cdot): \square^n \to \square^n$ is sufficiently smooth nonlinear mapping. If there exists a positive-definite and continuous Lyapunov function $V(\cdot,\cdot): \square^n \times \square \to [0,\infty)$ satisfied with Eq. (2)

$$\dot{V}(\mathbf{x}(t),t) \le -\lambda V^{\eta}(\mathbf{x}(t),t), \ \forall t \ge t_0, \ V(t_0) \ge 0, \quad (2)$$

where $\lambda > 0$, η is a constant exponent such that $0 < \eta < 1$, and $t_0 > 0$ is the initial time. Then,

$$V(\mathbf{x}(t),t) = 0, \ \forall t \ge T_s, \qquad (3)$$

where T_s is so call the settling time and determined as

$$T_s = t_0 + \frac{V^{1-\eta}(\mathbf{x}(t_0),t_0)}{\lambda(1-\eta)} \qquad (4)$$

2.1. Mathematical model

In this study, the model of the suspension system derived from Newton's law presented in [6], [7] was used. Fig. 1 shows a quarter-car model of an active suspension which includes the hydraulic actuator in the suspension system. The model can be extended by considering the hydraulic dynamics can be found in [6]. However, in this study, the model views the hydraulic actuator in the suspension system as the control input u_a and ignores the hydraulic dynamics as shown in Eq. (5):

$$M_b \ddot{x}_s + K_a(x_s - x_w) + C_a(\dot{x}_s - \dot{x}_w) - u_a = 0$$

$$M_{us}\ddot{x}_w + K_a(x_w - x_s) + C_a(\dot{x}_w - \dot{x}_s) + K_t(x_w - r) + u_a = 0 \tag{5}$$

where M_b represents the car body mass and M_{us} the wheel. The displacement of car body and wheel are denoted by x_s and x_w. The stiffness of the active suspension system are defined by K_a and K_t. The damping coefficient is represented by C_a. The force u_a is the actuator force from the hydraulic system. The road disturbance is defined by r and it is assumed to be bounded.

According to [6], [7] by letting $x_1 = x_s$ $x_2 = \dot{x}_s$ $x_3 = x_w$ and $x_4 = \dot{x}_w$, the corresponding state space representation is shown in Eq. (6):

$$\dot{x}_1 = x_2$$
$$\dot{x}_2 = -\frac{1}{M_b}\left[K_a(x_1 - x_3) + C_a(x_2 - x_4) - u_a\right]$$
$$\dot{x}_3 = x_4 \tag{6}$$
$$\dot{x}_4 = \frac{1}{M_{us}}\left[K_a(x_1 - x_3) + C_a(x_2 - x_4) - K_t(x_3 - r) + u_a\right]$$

The control objective is to control the error in Eq. (7) to converge to zero:

$$e = x_1 - \bar{x}_3, \tag{7}$$

where $\bar{x}_3$ is the filtered signal of the wheel displacement, x_3. The filter is defined as:

$$\dot{\bar{x}}_3 = \varepsilon(x_3 - \bar{x}_3). \tag{8}$$

where $\varepsilon > 0$ is the parameter to be determined later. The output is defined as:

$$y \,\square\, e = x_1 - \bar{x}_3. \tag{9}$$

It is worth noting that the reasons for introducing ε and additional dynamics of Eq. (8) were elaborated in detail in [6]. To summarize, introducing ε can provide us the degree of freedom to guarantee the stability of the zero dynamics corresponding to the chosen output in Eq. (9) as we will further discuss at the end of section 2.3. In addition, it will play an important role in the tradeoff between ride quality and rattle space usage.

Fig. 1. A quarter-car model with active suspension system

2.2. Controller design

According to [1], [2], [3], the terminal synergetic controller design procedure can be presented as follows: Firstly, define the macro variable of the controller.

$$\psi = e + \frac{(\dot{e})^{\frac{p}{q}}}{\alpha}. \tag{10}$$

where p and q are positive odd numbers such that $p > q$ Second, define the dynamic evolution of the macro variable as Eq. (11):

$$(\beta\dot{\psi})^{\frac{m}{n}} + \psi = 0 \tag{11}$$

where $1 < \frac{m}{n} < 2$, m and n are positive odd numbers, $\beta > 0$ is parameters to be determined by the designer and affect the convergence time of the macro variable to converge to zero. Finally, determine the control input as follows. By Eq. (10), the derivative of the macro variable is,

$$\dot{\psi} = \dot{e} + \frac{p(\dot{e})^{\frac{p}{q}-1}\ddot{e}}{q\alpha} \tag{12}$$

Recall that $e = x_1 - \bar{x}_3$, then.

$$\dot{e} = \dot{x}_1 - \dot{\bar{x}}_3 = x_2 + \varepsilon(\bar{x}_3 - x_3) \qquad (13)$$

And,

$$\ddot{e} = -\frac{\left[K_a(x_1 - x_3) + C_a(x_2 - x_4)\right]}{M_b} - \varepsilon^2(\bar{x}_3 - x_3) - \varepsilon x_4 + \frac{u_a}{M_b} \qquad (14)$$

By using the Eq. (11) – Eq. (12) and Eq. (14), we can determine the controller u_a as

$$u_a = M_b \left(\left(\frac{(-\psi)^{\frac{n}{m}}}{\beta} - \dot{e} \right) \frac{q\alpha}{p(\dot{e})^{\frac{p}{q}-1}} + \varepsilon^2(\bar{x}_3 - x_3) + \varepsilon x_4 \right) \\ + \left[K_a(x_1 - x_3) + C_a(x_2 - x_4) \right] \qquad (15)$$

Note that to guarantee that the term $(\dot{e})^{\frac{p}{q}-1}$ will not cause the singularity, the additional condition must be imposed on the parameters p, q, m, n, which the following

i) $\dfrac{pn + m(q-p)}{qn} > 0$ and ii) $\dfrac{m(p-q)}{pn} < 1$.

2.3. Proof of stability

The Lyapunov function is defined in terms of the macro variable as

$$V(\psi) = \frac{1}{2}\psi^2 \qquad (16)$$

By Eq. (12), the derivative of Eq. (16) is determined as

$$\dot{V}(\psi) = \psi\dot{\psi} = \psi \left(\dot{e} + \frac{p(\dot{e})^{\frac{p}{q}-1}\ddot{e}}{q\alpha} \right) \qquad (17)$$

By substituting the controller in Eq. (15) into Eq. (17), we yield,

$$\dot{V}(\psi) = \psi \frac{(-\psi)^{\frac{n}{m}}}{\beta}$$

Recall that m and n are positive odd number such that $0.5 < {n}/{m} < 1$ and let define $\eta = \dfrac{n+m}{2m}$ and $\lambda = \dfrac{2^\eta}{\beta}$.

Note that $0.75 < \eta < 1$ and $\lambda > 0$ by construction, then,

$$\dot{V}(\psi) = -\psi\frac{(\psi)^{\frac{n}{m}}}{\beta} = -\frac{1}{\beta}\left(2\frac{\psi^2}{2}\right)^\eta = -\lambda V^\eta \qquad (18)$$

Based on Lemma 1, the settling time T_1 is determined as

$$T_1 = t_0 + \frac{(\psi(t_0))^{2-2\eta}}{2^{1-\eta}\lambda(1-\eta)}. \qquad (19)$$

Then, this implies that the manifold $\psi = 0$, $\forall t > T_1$. From Eq. (10),

$$\dot{e} = -\bar{\alpha}e^{\bar{\eta}} \qquad (20)$$

whenever the manifold $\psi = 0$, where $\bar{\alpha} \triangleq \alpha^{\frac{q}{p}}$ and $\bar{\eta} \triangleq {q}/{p}$. Furthermore, notice that $\alpha > 0$ and both of p, q are positive odd numbers such that $p > q$, then $\bar{\alpha} > 0$ and $0 < \bar{\eta} < 1$ by construction. Again, by Lemma 1, the error $e(t)$ converges to zero within the time:

$$T_2 = T_1 + \frac{(e(T_1))^{2-2\bar{\eta}}}{2^{1-\bar{\eta}}\bar{\alpha}(1-\bar{\eta})} \qquad (21)$$

Hence, $x_1(t)$ converges to $\bar{x}_3(t)$ within $T = t_0 + T_1 + T_2$ which is a finite real number. Note that, since we choose the output $y \triangleq x_1 - \bar{x}_3$, we can obtain the zero dynamics in the matrix form as Eq. (22) [6] by setting the output identically equal to zero.

$$\dot{z} = A_z z + B_z r \qquad (22)$$

where $z \triangleq \begin{bmatrix} \bar{x}_3 & x_3 & x_4 \end{bmatrix}^T$

$$A_z \triangleq \begin{bmatrix} -\varepsilon & \varepsilon & 0 \\ 0 & 0 & 1 \\ -\varepsilon^2\dfrac{M_b}{M_{us}} & \left(\dfrac{\varepsilon^2 M_b - K_t}{M_{us}}\right) & -\varepsilon\dfrac{M_b}{M_{us}} \end{bmatrix} \text{ and}$$

$$B_z \triangleq \begin{bmatrix} 0 & 0 & {K_t}/{M_{us}} \end{bmatrix}^T.$$

and r is assumed to be a bounded road disturbance i.e., $|r| \le M_r$ where M_r is a positive constant.

One can easily realize that the matrix A_z is Hurwitz if and only if $\varepsilon > 0$ by using the Ruth-Hurwitz criterion. It is also worth noting that there is no finite escape time phenomenon that occurs no matter how large the magnitude of road disturbance $|r|$ is since the zero dynamics Eq. (22) is a linear system with z is the state vector of the system and r is the disturbance. From the design perspective, one can choose a parameter ε such that the matrix A_z is sufficiently robust enough to dominate the bounded road disturbance r to guarantee the convergence of the state z to zero. Then, the state x_1 and x_2 converge to zero for the following reasons. Since x_1 converges to $\bar{x}_3$ in finite time by Eq. (18) and the argument from Eq. (20) i.e., $\dot{e} = -\bar{\alpha}e^{\bar{\eta}}$ when the manifold $\psi = 0$. Then, whenever $z \square \begin{bmatrix} \bar{x}_3 & x_3 & x_4 \end{bmatrix}^T$ is identically equal to zero, it implies that x_1 is identically equal to zero and then it also implies that $x_2 = 0$ i.e., $x_2 \square \dot{x}_1$. Hence, the proof of locally terminal time stability is completed $\square$

2.4. *Controller parameter tuning*

By using the algorithm of DA, the controller parameters of the design control input as shown in Eq. (15) are selected optimally to minimize the cost function or performance index which is defined as

$$J = \sqrt{u_a^{\ T} u_a}$$

3. Simulation Results and Discussions

To show the capability of the designed controller, the model in Eq. (6) with the system parameters from [8], [9] are used as the simulation example. The system parameters are defined as follows: $M_b = 290$ kg, $M_{us} = 59$ kg, $K_a = 16{,}812$ N/m, $K_t = 190{,}000$ N/m, $C_a = 1000$ Ns/m. The controller parameters are given by DA with 40 search agents and 20 iterations. The controller parameters before tuning are arbitrarily defined as $\alpha_1 = 0.228$, $\beta_1 = 0.9431$, $\varepsilon_1 = 1.3356$, and the controller parameters after tuning are $\alpha_2 = 3.9648$ $\beta_2 = 0.10942$, $\varepsilon_2 = 3.7505$. Both cases are using $p = 5$, $q = 3$, $m = 7$, $n = 5$. We formulate the road disturbance based on the principles outlined in [6]. This disturbance is mathematically described as a single bump with a height of 5 cm ($a = 0.025$), defined as:

$$r = \begin{cases} a(1 - \cos 8\pi t), & 3.5 \leq t \leq 3.75 \\ 0, & otherwise \end{cases}$$

Fig. 2 and Fig. 3 illustrate that the proposed TSC method can effectively guide all state responses to the desired levels of displacement and velocity. Nevertheless, upon comparing the controller with and without optimal parameters obtained from DA, it is evident that the TSC with optimal parameters exhibits a superior convergence rate performance. Fig. 4 demonstrates that the TSC with optimal parameters can alleviate the presence of chattering.
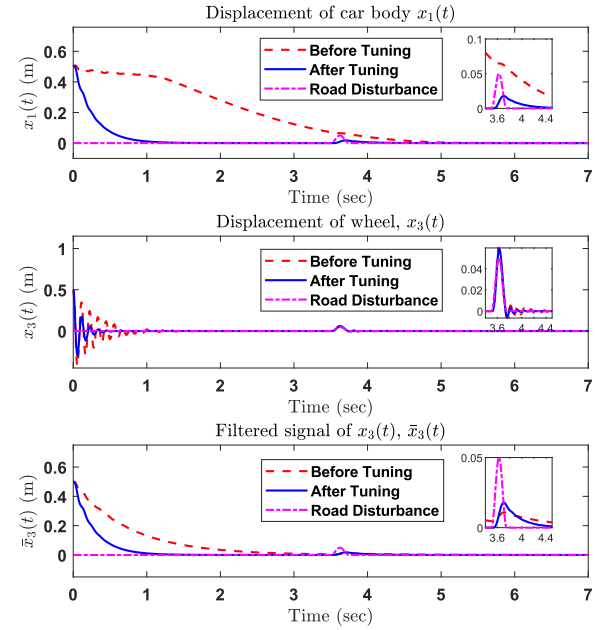

Fig. 2. Comparison of state response x_1, x_3, and $\bar{x}_3$ of the control system before and after DA tuning
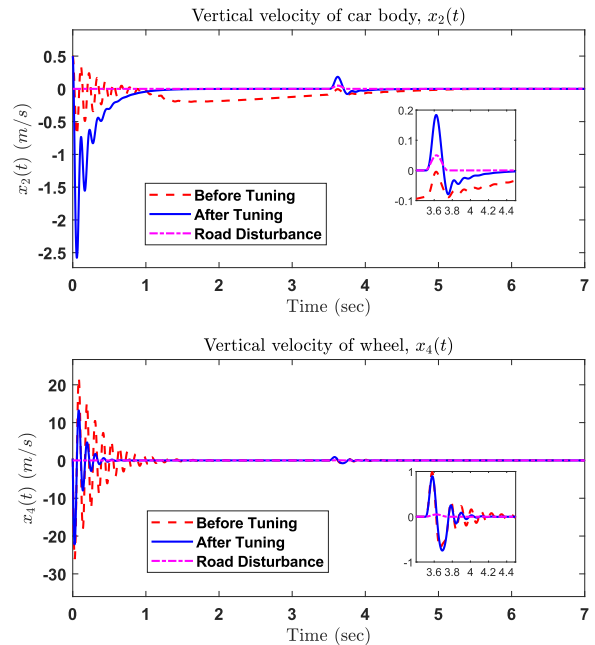

Fig. 3. Comparison of state response x_2 and x_4 of the control system before and after DA tuning
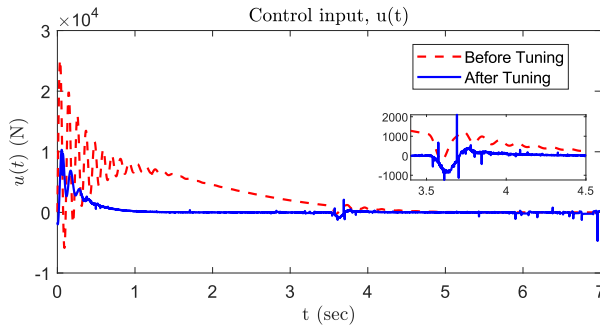
Fig. 4. Comparison of state response of the control input before and after DA tuning

4. Conclusion

In this study the TSC method was applied for controlling of the active suspension system. The controller parameters are selected optimally based on the algorithm of DA. The proposed controller effectively stabilizes the active suspension system by Lyapunov stability analysis. Simulation of the control system was carried out to validate the efficacy of the proposed controller. The simulation results demonstrate that the proposed controller enhances convergence properties. Additionally, the observed decrease in chattering in the control input signals, a significant issue in sliding mode control, further supports the potential application of the proposed TSC method in enhancing the overall performance of the suspension system.

References

1. A.A. Kolesnikov, "Introduction of synergetic control," Proceedings of the 2014 American Control Conference, Portland, OR, USA, Jun. 4-6, 2014, pp 3013-3016
2. A.A. Kolesnikov, *Modern Applied Control Theory: Synergetic Approach in Control Theory*. Moscow, Russia: Integracia-TSURE publ, 2000.
3. Y. Feng, X. Yu, and Z. Man, "Non-singular terminal sliding mode control of rigid manipulators," Automatica, vol. 38, no. 12, pp. 2159–2167, Dec. 2002
4. S. Mirjalili, "Dragonfly algorithm: a new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems," Neural Computing and Applications, vol. 27, no. 4, pp. 1053–1073, May 2015
5. A. Boonyaprapasorn, S. Kuntanapreeda, P.S. Ngiamsunthorn, T. Kumsaen, and T. Sethaput, "Time-varying sliding mode controller for heat exchanger with dragonfly algorithm," International Journal of Power Electronics and Drive Systems, vol. 13, no. 4, p. 3958, Aug. 2023.
6. J.S. Lin and I. Kanellakopoulos, "Nonlinear design of active suspensions," IEEE Control Systems Magazine, vol. 17, no. 3, pp. 45–59, Jun. 1997.
7. J.S. Lin and I. Kanellakopoulos, "Nonlinear design of active suspensions," Proceedings of the 34th IEEE Conference on Decision and Control, New Orleans, LA, USA, 1995, pp. 3567-3569.
8. A. Alkyne and J.K. Hedrick, "Nonlinear Control of a Quarter Car Active Suspension," Proceedings of the 1992 American Control conference, Chicago IL, pp. 21-25.
9. A.G. Alleyne, P. Neuhaus, and J.K. Hedrick, "Application of nonlinear control theory to electronically controlled suspensions," Vehicle System Dynamics, vol. 22, no. 5–6, pp. 309–320, Jan. 1993.

Authors Introduction

Dr. Tinnakorn Kumsaen

He received his B.S degree in chemical engineering from Khon Kaen University, Thailand, M.S. degree in system engineering from Case Western Reserve University (CWRU), and Ph.D. in chemical engineering from CWRU. He is currently a faculty member at chemical engineering, Khon Kaen University. His research interests are control system, process design simulations & integrations, and renewable energy.

Mr. Sorn Simatrang

He received a B.S. in electrical engineering from Chulalongkorn University, Thailand, and an M.S. in systems and control engineering from Case Western Reserve University, Ohio, USA in 2006 and 2015, respectively. Currently, he works as a control systems engineer at Nacres Co., Ltd, Bangkok, Thailand. His main research interests include output feedback control problems, adaptive control problems for nonlinear systems, and applications of nonlinear control theory to systems biology and mechanical systems.

Dr. Arsit Boonyaprapasorn

He received a B.S. in Mechanical Engineering from King Mongkut 's University of Technology (KMUTT), Thailand, in 1998. He continued his postgraduate studies at Case Western Reserve University, Cleveland, Ohio, U.S.A., obtaining an M.S.in Systems and Control Engineering in 2003 and a Ph.D. in Mechanical Engineering in 2009.Since then, he has served as a lecturer at Chulachomklao Royal Military Academy, Nakhon Nayok, Thailand. His research interests revolve around nonlinear feedback control and robotics.

Dr. Thunyaseth Sethaput

He received his B.S. degree in mechanical engineering from Sirindhorn International Institute of Technology, Thammasat University, Thailand and Ph.D. in systems and control engineering, from Case Western Reserve University, Ohio, USA. He is an assistant professor and chairperson of mechanical engineering program at Sirindhorn International Institute of Technology, Thailand. His research interests are control systems, systems biology, biomedical mechanics, simulation modeling.

Rehabilitating Flood-Damaged Cars for
Sustainable Car Rental Services: A Web-Based System

Pon Xiao Qi
Institute of Computer Science and Digital Innovation, UCSI University, Kuala Lumpur, Malaysia

Abdul Samad bin Shibghatullah
Institute of Computer Science and Digital Innovation, UCSI University, Kuala Lumpur, Malaysia

Kasthuri Subaramaniam
Institute of Computer Science and Digital Innovation, UCSI University, Kuala Lumpur, Malaysia
Email: 1001749088@ucsiuniversity.edu.my, abdulsamad@ucsiuniversity.edu.my,
kasthurisuba@ucsiuniversity.edu.my

Abstract

Reusing flood-damaged cars gives a unique opportunity in the car rental market in today's fast-paced world, where sustainable solutions are crucial. This study represents the result of efforts to design and construct a web-based car rental system that utilizes flood-damaged cars. The system will be developed using Rapid Application Development (RAD) methodology. Surveys and interviews are conducted to collect potential users' feedback. The results show that due to global economic recession and the increased awareness of global warming, citizens will prefer transportation that is eco-friendly and cost effective if the vehicle is in good condition and properly fixed. This project intends to transform the car rental environment by leveraging a strong technological stack and a user-centric design. With a strong emphasis on user experience and system efficiency, the web-based platform promises not only to ease seamless vehicle rentals but also to promote sustainability in the automotive sector. The favorable results of user acceptability testing confirm the system's success in reaching its objectives. GPS features and transparent vehicle history disclosure will be included for future enhancements.

Keywords: Car Rental System, Eco-Friendly, Flood

1. Introduction

The car industry has long struggled with the issues of sustainability and environmental friendliness. Simultaneously, natural calamities such as floods have resulted in many flood-damaged cars that, despite being salvageable, are frequently underutilized. Recognizing the potential of combining these two factors, this project will create a web-based automobile rental system that will reuse flood-damaged vehicles. This system solves the environmental challenges connected with underutilized cars while also catering to the increasing demands of modern consumers seeking economical and sustainable transportation alternatives by providing a digital platform for vehicle renters.

2. Methodology

2.1 Latest Statistic on Website Usage

In 2023, there are approximately 1.13 billion websites on the internet. Only a small portion of the world's 1.13 billion websites are regularly updated and used. Only 200,121,724 of the 1.13 billion websites are regularly maintained and viewed, a startling 82% of which are inactive. Each three seconds, a brand-new website is created.

In 2023, 71% of companies will have websites. This is a rise over prior years, owing primarily to the influence of the Covid-19 epidemic. Companies have realized the value of having a web site to reach a bigger audience and stay competitive in today's digital market because of the move to online commerce and remote employment. The trend of companies creating a website has been hastened by the Covid-19 pandemic, and it is probable that this trend will continue.

The latest research shows that 28% of all company activity is now done online. This transition to online shopping reflects the internet's growing significance in today's business environment and the expanding trend of e-commerce.

Recent studies show that 43% of small firms intend to spend money in 2023 to improve the functionality of their websites. This phenomenon highlights the growing significance of having a strong online presence for companies of all sizes. A business' success in today's digital environment depends on having a well-designed and optimized website, as consumers are increasingly turning to the internet to make purchasing decisions [1].

Moreover, 5.18 billion people used the internet at the start of the second quarter of 2023. The number of

Pon Xiao Qi, Abdul Samad bin Shibghatullah, Kasthuri Subaramaniam

individuals using the internet is increasing; according to the latest recent data, there were 147 million more users globally connected in the year before April 2023. Due to growing internet availability globally, the number of people who are "unconnected" to the internet has fallen to 2.85 billion, with the majority of these people living in Southern and Eastern Asia as well as Africa.

Mobile phones now account for approximately 60% of all web traffic and are used by 95% of internet users globally to access the internet at least periodically as shown in Fig 1. People spend more than 57 percent of their online time on mobile devices. However, in the world's major economies, more than 6 in 10 internet users still utilize laptops and desktop computers for at least some of their online activity [2].

Fig 1. Statistics of internet usage

2.2 History of Flood in Malaysia

According to studies conducted specifically in Malaysia, the fast-rising number of flood events in recent decades are the result of unregulated human activity in rivers, the cutting down of forests, and extreme weather events brought on by climate change [3]. Since 2001, there have been 143 floods on average annually, more than 90% of which have been flash floods. Such often recurrent floods pose a serious risk to the lives and property of Malaysian inhabitants, severely damaging the nation's infrastructure [4].

According to information from the Country Disaster and Risk Profile of the United Nations Office for Disaster Risk Reduction (UNDRR), floods in Malaysia account for 98% of average yearly damage between 1990 and 2014. According to a report from the Malaysian Department of Irrigation and Drainage published in 2003, flooding affects an average of 29 000 km2, or 9 percent of the nation's total land area, and more than 4.82 million people (22 percent of the population) per year [5].

Two different forms of rainfall generate floods. The first kind of rainfall is of a moderate intensity and lasts for a long time over a large region. The second kind is localized heavy rain that lasts for a short time. Floods in Malaysia are predictable to happen. While the west coast of Peninsular is primarily plagued by thunderstorms from September to November because of the inter-monsoon

season, the east coast and eastern Malaysia are typically affected by floods during December to January when the northeast monsoon sweeps. Most floods in Malaysia typically originate from persistent, heavy rains that cause runoff because of surplus water supplies that exceed the capacity of streams and rivers. Many states in Peninsular Malaysia, including Kelantan, Johor, Pahang, Perak, Kuala Lumpur, and Selangor, as well as east Malaysia's Sabah and Sarawak, are home to the majority of flood-prone areas [6].

2.3 Overview of Car Rental System

The development of technology caused fierce competition among businesses to win over more clients [7]. Technology aids in decision-making by allowing the owner of the car rental to concentrate on the usage of IT to gain a competitive edge [8]. Car rental companies frequently struggle with transactions that are inefficient. The tenant must make individual calls to the rental provider to inquire on the car's availability [9]. Distance is a problem for communication and information technologies, yet a website allows us to access a wide range of information from around the globe. Online transactions are those when the buyer and seller conduct their business over the Internet. Web-based Information Systems were created because of the quick advancement of information technology. Research on general models for the study and design of web-based information systems has been done in numerous papers [10]. Due to the current rapid growth of information technology, it is essential to access information quickly and accurately to facilitate decision-making. Due to the accessibility of information and the effective and efficient processing of order data, using web design as a tool to promote the rental of vehicles is highly beneficial in attracting more consumers and making transactions easier.

2.4 Problem Statement

Natural calamities, such as floods, have caused severe damage to automobiles in Malaysia in recent years, bringing huge financial damages to individuals. Depending on the type and level of damage, it might cost the automobile owner between RM20,000 and RM40,000. Repair costs will be higher for luxury vehicles such as BMW, Porsche, Range Rover, Peugeot, Mini Cooper, and so on. Due to the amount of damage in the car, some car owners may be unable to afford the repair costs or may opt to sell it and acquire a new car. These cars will almost certainly wind up in landfills, contributing to pollution. This is because not all vehicles have flood insurance. Driving a flood-damaged car that has not been adequately repaired might be hazardous to the driver's health and threaten the safety of other vehicles on the road. Mold can form in flood-damaged cars if wet conditions are not handled, causing respiratory difficulties for the driver and passengers. A web-based car rental system for flood-damaged vehicles can provide a convenient option for

these car owners to sell their flood-damaged vehicles through a reputable platform. The owner can utilize the money from selling the flood-damaged car to our platform to buy a new car. Our platform will then repair these flood-damaged vehicles and rent them out at a low cost.

Without a web-based car rental system, flood-affected car owners must either continue to use the vehicle, dispose it in landfills, or spend time contacting secondhand car dealers who are prepared to acquire their flood-damaged vehicle, which is a time-consuming process. The goal of this project is to reduce the number of vehicles that wind up in landfills, hence reducing environmental effect. Furthermore, reusing flood-damaged vehicles through rental businesses is a more sustainable option than scrapping them. Furthermore, an affordable automobile rental system including flood-damaged vehicles would provide transportation options for people who might not be able to purchase a car otherwise. This accessibility would allow people to commute to job, education, or other important places, making their lives easier. Furthermore, these flood-damaged vehicles might be used in the film business. Our platform ensures that vehicles are thoroughly inspected and repaired prior to rental, ensuring renters' safety.

3. Implementation

3.1 Method

The proposed car rental website is built using the Rapid Application Development (RAD) methodology. The RAD methodology is divided into four phases: user requirements, user design, prototype, and implementation. During the user requirements phase, system requirements will be gathered through surveys and questionnaires. Once the requirements are defined, we will move into the user design phase. The emphasis here was on creating the high-level structure of the web-based car rental system. We envisioned the system's components, their interrelationships, and the general architecture that would support its functionality.

The prototyping step marked the move from concept to actual representation. At this point, we'll start working on a workable model of the system. This prototype served as proof of concept as well as a visual depiction of the system's primary functions. It enabled my supervisors and potential customers the opportunity to interact with a simplified version of the system and obtain a hands-on grasp of how it would work. This phase included continuous testing and validation to ensure that the system remained reliable and of high quality throughout development.

The final phase, implementation, marked the web-based automobile rental system's preparedness for deployment. To ensure a successful launch, several critical procedures will be followed. Extensive testing will be carried out to ensure that the system's performance, usability, and functionality are satisfactory.

3.2 User Functionality

Users should be able to quickly register on the suggested website. Important information such as their name and contact information will need to be obtained throughout the registration process. Users must be able to securely log in using their selected credentials (username and password) once registered, allowing them to tailor their interactions with the system and log in or out as needed.

Access to a comprehensive list of available cars is an important component of the user experience. This list should include specific information about each car, such as its model, year, rental fee, and a brief description. A search option should be included to improve user ease, allowing consumers to filter available cars based on particular parameters such as location and rental term. This enables users to quickly select the appropriate car that suits their individual requirements.

Furthermore, reserving a car for a specific time should be a simple process. Users should be able to enter their preferred pick-up and return dates, allowing for smooth booking procedures. Users should also be able to alter or cancel previous bookings to accommodate changes in plans if the request is made at least three days before the scheduled pick-up date.

The Touch & Go payment system, which was recommended by 96% of poll respondents, will be included in the system. This strategy offers quick and easy payment processing, which improves the entire customer experience.

Additionally, the system will highlight how important user feedback is. Users should be able to provide ratings and comments about their rental experience after completing a booking. Future users can benefit greatly from these user reviews, which offer insights and support well-informed decision-making. A thorough booking history that provides a summary of prior reservations for simple access and tracking of rental activity should also be available to users. To safeguard our clients, our business will offer a warranty on every vehicle. However, GPS features and transparent vehicle history disclosure will be included for future enhancements (Fig 2).
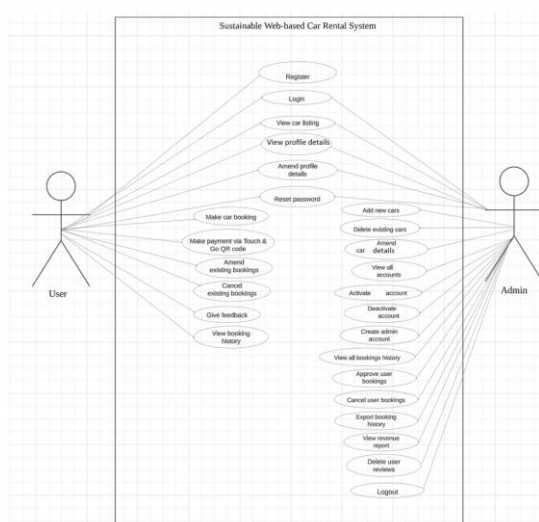
Fig 2. Use Case Diagram

4. Conclusion

Administrators should be equipped with powerful tools to efficiently manage automobile listings. This involves having the power to update current listings, remove out-of-date ones, and add new cars to the inventory. This guarantees the accuracy and currentness of the vehicle inventory.

Besides this, administrators need to be able to supervise and manage every user account. This includes having the power to activate and deactivate admin and user accounts as necessary. Admins also can make new admin accounts, which will improve platform integrity and security.

All bookings must be managed efficiently. Administrators should be able to approve or cancel reservations based on availability and particular criteria. Admin should be able to manage user reviews as well. The administration of user reviews is equally important in maintaining the platform's positive and trustworthy reputation.

Moreover, admins should be able to export booking history records to aid in record-keeping and analysis. This information is crucial in accounting and auditing. Admins should also be able to track the platform's financial performance over time by viewing revenue data, which will help them assess the platform's viability and growth.

Lastly, both users and administrators should have access to a visual booking calendar that displays all bookings made on certain days. This calendar tool aids in efficient planning and management, improving the system's overall functionality.

References

1. Haan, K. (2023) Top website statistics for 2023, Forbes. Available at: https://www.forbes.com/advisor/business/software/website-statistics/#:~:text=There%20are%20about%201.13%20billion,are%20actively%20used%20and%20updated. (Accessed: 30 May 2023).
2. Digital around the world - datareportal – global digital insights (no date) DataReportal. Available at: https://datareportal.com/global-digital-overview (Accessed: 30 May 2023).
3. Aliagha, G.U., Mar Iman, A.H., Ali, H.M., Kamaruddin, N. and Ali, K.N., 2015. Discriminant factors of flood insurance demand for flood‑hit residential properties: a case for M alaysia. Journal of Flood Risk Management, 8(1), pp.39-51.
4. Nasiri, H. and Shahmohammadi-Kalalagh, S., 2013. Flood vulnerability index as a knowledge base for flood risk assessment in urban area. Journal of Novel Applied Science, 2(8), pp.269-272.
5. D'Ayala, D., Wang, K., Yan, Y., Smith, H., Massam, A., Filipova, V. and Pereira, J.J., 2020. Flood vulnerability and risk assessment of urban traditional buildings in a heritage district of Kuala Lumpur, Malaysia. Natural Hazards and Earth System Sciences, 20(8), pp.2221-2241.
6. Safiah Yusmah, M.Y., Bracken, L.J., Sahdan, Z., Norhaslina, H., Melasutra, M.D., Ghaffarianhoseini, A., Sumiliana, S. and Shereen Farisha, A.S., 2020. Understanding urban flood vulnerability and resilience: a case study of Kuantan, Pahang, Malaysia. Natural Hazards, 101, pp.551-571.
7. Adiyanto, N., 2021. Customer Relationship Management (CRM) Based On Web To Improve The Performance Of The Company. IAIC Transactions on Sustainable Digital Innovation (ITSDI) The 1st Edition Vol. 1 No. 1 October 2019, 32.
8. Soegoto, E.S., 2018, November. Information Technology Based Entrepreneurship Education in University. In International Conference on Business, Economic, Social Science and Humanities (ICOBEST 2018) (pp. 316-320). Atlantis Press.
9. Rozi, N.F., Ruswiansari, M., Rachman, A., Wardhana, S.R. and Istiyanto, L., 2019. The Development of LIDI: A Web-Based Car Rent Marketplace Application in Sidoarjo, Indonesia. In IOP Conference Series: Materials Science and Engineering (Vol. 462, No. 1, p. 012052). IOP Publishing.
10. Welzer, T., Eder, J., Podgorelec, V., Robert, W., Ivanovic, M., Gamper, J., Morzy, M., Tzouramanis, T., Darmont, J. and Kamišalić, A., 2019. New Trends in Databases and Information Systems. In 23rd European Conference on Advances in Databases and Information Systems (ADBIS 2019) (Vol. 1064). Springer.

Authors Introduction

Ms. Pon Xiao Qi

She is currently pursuing B. Sc. (Hons) Business Information Systems at Institute of Computer Science and Digital Innovation (ICSDI), UCSI University, Kuala Lumpur, Malaysia.

Dr. Abdul Samad bin Shibghatullah

He received the bachelor accounting degree from Universiti Kebangsaan Malaysia, Bangi, Malaysia in 1999, the M.Sc. degree in computer science from the Universiti Teknologi Malaysia, Skudai, Malaysia in 2002, and the Ph.D. degree in computer science from the Brunel University of Uxbridge, United Kingdom. He is currently an Associate Professor at Institute of Computer Science etc.

Dr. Kasthuri Subaramaniam

She is currently an assistant professor at Institute of Computer Science and Digital Innovation (ICSDI), UCSI University, Kuala Lumpur, Malaysia. She earned both her bachelor's degree in computer science and a master's degree in computer science from the University of Malaya. Her research interests include human-computer interaction, etc.

Optimizing E-Invoicing Rollout: Adaptive E-Invoicing Rollout (AER) Framework for Navigating Malaysia's Digital Transformation

Koh Chee Hong
UCSI University, Kuala Lumpur, Malaysia
Email: 1002265962@ucsiuniversity.edu.my

Abdul Samad Bin Shibghatullah
UCSI University, Kuala Lumpur, Malaysia
Email: AbdulSamad@ucsiuniversity.edu.my

Abstract

In the dynamic landscape of Southeast Asia's digital economy, Malaysia is gearing up for a transformative shift with the mandatory rollout of electronic invoicing (e-Invoicing) beginning June 2024. Recognizing the critical need for a structured approach to this transition, this study presents the design and evaluation of an Adaptive E-Invoicing Rollout (AER) Framework, which is aimed at facilitating a successful e-Invoicing implementation across Malaysia's diverse business sectors. The framework is developed through a qualitative methodology that synthesizes insights from an extensive literature review, expert interviews, case studies, and the Delphi technique. The AER Framework stands out for its comprehensive approach, incorporating a theoretical predictive model that assesses various variables influencing e-Invoicing success. By employing the correlation coefficient, the model quantifies the relationships between key variables, providing a structured mechanism to predict and enhance the successful adoption of e-Invoicing. This model is particularly designed to be adaptable and dynamic, evolving with ongoing insights and changes in the e-Invoicing landscape. The study's findings offer a significant contribution to the digital economic literature and practice, providing policymakers, businesses, and stakeholders with a robust tool for strategic planning and effective decision-making. As Malaysia navigates its digital transformation, the AER Framework serves as a critical guide, ensuring that the e-Invoicing rollout is not only compliant with regulations but also conducive to enhancing fiscal transparency, economic efficiency, and business operations. Furthermore, the framework's adaptability and predictive nature make it a relevant model for other countries considering similar digital transitions, marking it as an asset in the broader context of digital finance evolution.

Keywords: Electronic Invoicing, Digital Transformation, Adaptive Framework, Predictive Model, Qualitative Methodology, Malaysia, Correlation Coefficient, Stakeholder Strategy, Fiscal Transparency

1. Introduction

In the burgeoning digital economy of Southeast Asia, Malaysia stands on the precipice of a significant transformation with the impending mandatory rollout of e-Invoicing beginning June 2024. The introduction of e-Invoicing is not merely a technological shift, but a strategic maneuver poised to streamline tax compliance, bolster fiscal transparency, and catalyze economic efficiency. However, despite the clear benefits and the comprehensive guidelines issued by the Inland Revenue Board (IRB), the successful adoption and implementation of e-Invoicing across diverse business ecosystems presents a complex challenge.

The significance of this study lies in its potential to bridge the gap in understanding the multifaceted dimensions of e-Invoicing adoption, from the perspective of compliance [1],[2] to the readiness of technological infrastructure [3],[4]. As noted by Chan [5], the digital economy collaboration and the integration of digital payment systems are crucial areas of development in ASEAN, with Malaysia being no exception. Yet, a comprehensive framework that encompasses all aspects of the transition to e-Invoicing is absent, underscoring the necessity for

this research. This study aims to contribute to the body of knowledge by providing an empirical evaluation of the factors influencing e-Invoicing success in Malaysia, thereby guiding policymakers and businesses in their decision-making processes. Moreover, it will offer a novel predictive model to anticipate the trajectory of e-Invoicing adoption and its ripple effects on the digital economy, drawing on the insights from the ASEAN Digital Economy framework [6].

The novelty element of this study is encapsulated in its comparative analysis approach, juxtaposing the Malaysian e-Invoicing rollout with global best practices [7], [8] and its predictive nature, which seeks to forecast the long-term implications of e-Invoicing on the digital transformation of the Malaysian economy [9], [10]. By examining the intersection of digital payments, data regulations, and e-Invoicing [5], [11], along with the critical factors for cloud-based e-invoice service adoption [12] and the effects of e-government measures on 2. taxpayer compliance [13], this research will carve out a novel paradigm in the digital economic literature, specifically tailored to the Malaysian context. The proposed research is poised to offer a robust, scalable, and adaptive framework for e-Invoicing implementation in Malaysia, contributing significantly to the literature

and providing practical implications for stakeholders across the board.

3. Methodology

The methodology for the research study, " Optimizing E-Invoicing Rollout: Adaptive E-Invoicing Rollout (AER) Framework for Navigating Malaysia's Digital Transformation" will be composed of a qualitative approach, given the pre-implementation state of e-invoicing in Malaysia.

3.1. *Literature Review*

A comprehensive literature review was conducted to synthesize global and local e-invoicing practices, identifying critical themes and practices that inform a predictive model for e-invoicing success. The key findings are summarized in Table 1.0: Key Findings from Literature Review.

3.2. *Expert Interviews*

Insights were gathered from industry experts to understand the operational, technical, and compliance aspects of e-invoicing implementation. The synthesized perspectives are documented in Table 2: Table 2.0: Theoretical Predictive Model Variables for E-Invoicing Success.

3.3. *Case Study Analysis*

An analysis of various regional e-invoicing implementations provided criteria for success, highlighting parallels and distinctions across ASEAN countries. Findings are presented in a comparative format.

3.4. *Delphi Technique*

A Delphi study established consensus on predictive parameters for e-invoicing success, integrating expert opinions into the AER Framework's development criteria. Parameters such as system integration readiness, compliance rates, and user adoption were discussed, echoing the approach suggested by Wagiman et al. [14] in applying the Delone & McLean Information Systems Success Model to assess net benefits.

3.5. *Theoretical Predictive Analysis*

The theoretical predictive model was developed from an integration of literature reviews, expert interviews, and case study findings. It evaluated various variables impacting the success of e-invoicing, each assigned a rating from 1 to 5 for its predictive importance.

Application of the Correlation Coefficient:

The correlation coefficient, a crucial metric in this model, was calculated to measure the relationship between pairs of variables, such as regulatory compliance and e-invoicing success rate. It indicated both the strength and direction of their relationship.

Pearson's Correlation Formula Simplified:

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

Where:

- **r (correlation coefficient):** This was the desired outcome of the analysis. It is a value between -1 and 1 that indicated how strongly two variables were related.
- **n (number of pairs):** This represented the total number of observations or data points available. Each pair represented a set of two variables, such as a particular e-invoicing feature's implementation and its observed impact on overall success.
- **Σxy (sum of the product of paired scores):** For each pair of variables, they were multiplied together and then all these products were summed up. This sum helped understand the joint variability of the two variables.
- **Σx and Σy (sum of x scores and y scores):** These were simply the totals of all the first and second variables in the pairs respectively.
- **Σx² and Σy² (sum of the squares of x scores and y scores):** Similar to Σx and Σy, but each variable was squared before they were summed up. This component helped in understanding the spread or distribution of each variable.

The results of this predictive analysis, including correlation coefficients and predictive importance ratings, were summarized in Table 2.0: Theoretical Predictive Model Variables for E-Invoicing Success. This table served as an essential tool for stakeholders, offering insights into the influential factors of e-invoicing success. It enabled stakeholders to anticipate potential outcomes more effectively and align strategies accordingly.

3.6. *Documentation and Reporting*

The meticulous documentation and reporting of findings throughout the research process ensured that the framework would be dynamic and subject to continuous improvement. This living document approach allows the framework to evolve as new insights emerge, maintaining its relevance and effectiveness.

4. Result/Discussion

4.1. *Development of the AER Framework*

The "Adaptive E-Invoicing Rollout (AER) Framework" is the outcome of a methodical approach that amalgamates a comprehensive literature review, expert interviews, comparative case studies, Delphi consensus,

and theoretical predictive analysis. This blend has formulated a detailed and adaptable guide suitable for the nuanced needs of Malaysia's e-invoicing initiative.

4.2. *Literature Review Outcome*

The outcome of the literature review of e-invoicing related journal papers were consolidated in the Table 1 below:

Table 1 Key Findings from Literature Review

Theme	Key Findings	Predictive Model Relevance	Reference
Digital Economy & E-Invoicing	Studies note the rapid growth of e-invoicing and regulations in ASEAN	Highlights the significance of Cintegration and Aadoption in the predictive model.	[5], [6]
Implementation & Adoption Challenges	Identifies technology, operational, and compliance barriers to e-invoicing adoption	Reinforces the importance of Rreadiness and Ccompliance	[1], [9], [5], [16], [18], [20]
Benefits & Outlook	Discusses the potential benefits of e-invoicing and forecasts its adoption trajectory	Suggests Uusage as a critical outcome measure in the model.	[7], [8], [14], [17]
Regulatory Aspects & Compliance	Focuses on the role of regulations and legal frameworks in facilitating or hindering e-invoicing	Underscores Ccompliance as a vital predictor in the model.	[3], [11], [13], [19]

4.3. *Predictive Analysis for E-Invoicing Success*

The critical component of the AER Framework is the theoretical predictive model, which was designed to forecast and evaluate the success of the e-invoicing rollout. Drawing insights from the comprehensive literature review, this model incorporates various variables identified as significant predictors of e-invoicing success. Each variable, substantiated by the literature, is accompanied by a correlation coefficient indicating its impact and relevance in forecasting the successful implementation of e-invoicing, as consolidated in Table 2 below.

Table 2 Theoretical Predictive Model Variables for E-Invoicing Success

Variable (Factor)	Predictive Importance (Scale: 1-5)	Correlation Coefficient	Description
Cintegration (System & Data Integration Capability)	4	0.85	Technical integration capability [10]
Aadoption (Rate of Adoption Among Users)	3	0.65	User acceptance rate [9]
Rreadiness (Organizational & Technological Readiness)	4	0.75	Business and infrastructure preparedness [16]
Ccompliance (Compliance with Legal & Tax Requirements)	5	0.90	Regulatory adherence [13]
Uusage (Actual Usage & Utilization Rates Post-Implementation)	4	0.80	Utilization post-implementation [1]

In this table:

- **Predictive Importance** is quantified on a scale of 1-5, where 1 is the lowest and 5 is the highest level of importance. This allows a more objective comparison of factors.
- **Correlation Coefficient** remains a numeric value indicating the strength and direction of the relationship between each variable and the success of e-invoicing implementation.
- **Description** is a concise statement of what each variable represents.
- **References** are provided for each variable, offering direct access to the underlying studies or sources.

The predictive model is a robust, theoretical construct based on identified variables derived from qualitative insights and empirical evidence from similar initiatives. Each variable is assigned a predictive importance and a correlation coefficient, indicating how strongly it's expected to influence the success of the e-invoicing rollout. This model not only aids in forecasting outcomes but also assists in identifying areas requiring strategic focus and resource allocation, making it a grounded and practical tool for stakeholders. A conceptual model was defined to guide future quantitative studies. The variables identified for the predictive model are enumerated in Table 5: Predictive Model Variables for E-Invoicing Success. Informed by Hernandez-Ortega's work [10] and other seminal studies, this model was conceptualized to guide stakeholders in anticipating and preparing for the

potential outcomes of the e-invoicing rollout. The formula for the theoretical predictive model is as follows:

$$Psuccess=$$
$$f(Cintegration, Aadoption, Rreadiness, Ccompliance, Uusage)$$

Where:

- *Psuccess* is the probability of successful e-invoicing rollout.
- *Cintegration* represents system and data integration capability.
- *Aadoption* indicates the rate of adoption
- *Rreadiness* reflects organizational and technological readiness.
- *Ccompliance* is the degree of compliance with legal and tax requirements.
- *Uusage* measures actual usage and utilization rates post-implementation.

This model, as illustrated in Fig. 1, is significantly grounded in empirical evidence and industry insights, making it a novel and practical contribution to the field of e-invoicing and digital economy strategies.
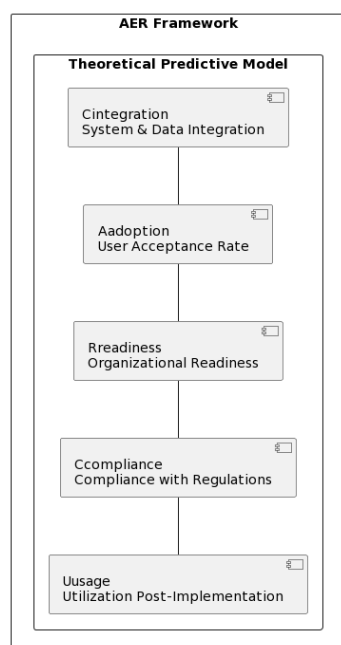

Fig.1 AER Framework - Theoretical Predictive Model

4.4. *Conclusion and Strategic Implications*

The AER Framework emerges as a strategic, comprehensive guide for e-invoicing adoption in Malaysia. It is not just a theoretical construct but a practical, dynamic tool reflecting the complexities and requirements of real-world implementation. The framework's adaptability, combined with its predictive capability, positions it as a model for digital transformation, not only for Malaysia but potentially for other countries embarking on a similar journey. As the landscape of digital finance continues to evolve, the AER Framework is poised to remain a relevant, effective tool for navigating the future of e-invoicing.

5. Conclusion

Malaysia's transition to mandatory e-Invoicing reflects a crucial shift in its digital economy. This study developed an adaptive framework, utilizing a qualitative methodology to integrate insights from various sources, culminating in a theoretical predictive model. This model offers stakeholders a robust tool for strategic e-Invoicing implementation. The AER emerges as a practical guide, significant for Malaysia's digital transformation and relevant for similar global contexts, ensuring continued efficacy in evolving digital finance landscapes.

References

1. Suendra, A. A. P. (2023). Analysis of the Implementation of E-Invoice 3.0 in Increasing Taxpayer Compliance at Kpp Pratama Bandar Lampung Dua. Data: Journal of Information Systems and Management, 1(1), 50-62.
2. Irawati, I., Darmawan, H., Sofyan, M., & Serebryakova, T. (2022). The effect of the implementation of E-Invoice 3.0 and taxable person compliance to value added tax revenue at KPP Pratama Pasar Rebo. Ilomata International Journal of Tax and Accounting, 3(1), 1-12.
3. Balakrishnan, V., & Shuib, N. L. M. (2021). Drivers and inhibitors for digital payment adoption using the Cashless Society Readiness-Adoption model in Malaysia. Technology in Society, 65, 101554.
4. Ooi, K. B., Lee, V. H., Tan, G. W. H., Hew, T. S., & Hew, J. J. (2018). Cloud computing in manufacturing: The next industrial revolution in Malaysia?. Expert Systems with Applications, 93, 376-394.
5. Chan, M. (2024). Malaysia: Digital Payments, Data Regulations, and AI as Most Promising Areas for Digital Economy Collaboration. In The ASEAN Digital Economy (pp. 76-96). Routledge.
6. Cheung, P., & Taojun, X. (Eds.). (2023). The ASEAN Digital Economy: Towards an Integrated Regional Framework. Taylor & Francis.
7. Olaleye, S. A., Sanusi, I. T., Dada, O. A., & Agbo, F. J. (2023). A bibliometric review of global visibility, impact and adoption of electronic invoicing: The past and the future. Heliyon.
8. Harris, H., Guru, B. K., & Avvari, M. V. (2011). Evidence of firms' perceptions toward Electronic Payment Systems (EPS) in Malaysia. International journal of business and information, 6(2), 226.
9. Nguyen, A. H., Nguyen, T. P., & Dang, G. T. T. (2020). Determinants Of E-Invoice Adoption: Empirical evidence from Vietnam. The Journal of Asian Finance, Economics and Business (JAFEB), 7(7), 311-321.
10. Hernandez-Ortega, B. (2012). Key factors for the adoption and subsequent use of e-invoicing. Academia. Revista Latinoamericana de Administración, (50), 15-30.
11. Kamaruzaman, K. N., Handrich, Y. M., & Sullivan, F. (2010). E-commerce adoption in Malaysia: Trends, issues and opportunities. ICT strategic review, 11.

12. Lian, J. W. (2015). Critical factors for cloud based e-invoice service adoption in Taiwan: An empirical study. International Journal of Information Management, 35(1), 98-109.

13. Rokhman, A., Handoko, W., Tobirin, T., Antono, A., Kurniasih, D., & Sulaiman, A. (2023). The effects of e-government, e-billing and e-filing on taxpayer compliance: A case of taxpayers in Indonesia. International Journal of Data and Network Science, 7(1), 49-56.

14. Wagiman, A. N., Aspasya, G. S., & Prawati, L. D. (2023). Net Benefit on E-Invoice Implementation: Applying the Delone & McLean Information Systems Success Model. In E3S Web of Conferences (Vol. 388). EDP Sciences.

15. Hannah, B., & Nani, G. (2021). Improving procurement performance in the public sector with the implementation of E-procurement: a study of selected Metropolitan and Municipal Assemblies in the Ashanti region of Ghana.

16. Liew, S. Y. (2022). The determinants of Malaysia's revenue: the role of e-filing, tax administration and tax compliance (Doctoral dissertation, UTAR).

17. Groznik, A., & Manfreda, A. (2015, October). E-Invoicing and e-government–impact on business processes. In DIEM: Dubrovnik International Economic Meeting (Vol. 2, No. 1, pp. 204-217). Sveučilište u Dubrovniku.

18. Alaweti, M. F., Azizan, N., & Faryadi, Q. (2013). The Impact of Experience and Perception on the Efficiency of E-procurement in Malaysia. Journal of Communication and Computer, 10, 1475-1483.

19. Panayiotou, N. A., & Stavrou, V. P. (2021). Government to business e-services–A systematic literature review. Government Information Quarterly, 38(2), 101576.

20. Waithaka, R. K., & Kimani, J. G. (2021). Effect of e-procurement practices on supply chain performance. Global Journal of Purchasing and Procurement Management, 1(1), 32-42.

Authors Introduction

Mr. Koh Chee Hong

A seasoned financial technologist specializing in SAP FICO, SAP BPC/BW, and Oracle NetSuite, he is currently pursuing a PhD in Computer Science at UCSI University. He holds a First Class Honors B.Tech in Business Information System from Universiti Teknologi PETRONAS.

Dr Abdul Samad Bin Shibghatullah

He is a Senior Lecturer in UCSI with academic qualifications including a PhD and MSc in Computer Science, along with a Bachelor's degree in Accounting. His areas of focus include Modelling, Simulation, Optimization, and Artificial Intelligence.

App Alert System for Smart Phones

Chee Kin Hoe
Institute of Computer Science & Digital Innovation, UCSI University,
UCSI Heights, 1 Jalan Puncak Menara Gading, Kuala Lumpur, WP Kuala Lumpur 56000, Malaysia

Kasthuri Subaramaniam
Institute of Computer Science & Digital Innovation, UCSI University,
UCSI Heights, 1 Jalan Puncak Menara Gading, Kuala Lumpur, WP Kuala Lumpur 56000, Malaysia

Abdul Samad bin Shibghatullah
Institute of Computer Science & Digital Innovation, UCSI University,
UCSI Heights, 1 Jalan Puncak Menara Gading, Kuala Lumpur, WP Kuala Lumpur 56000, Malaysia
E-mail: 1002163440@ucsiuniersity.edu.my, kasthurisuba@ucsiuniversity.edu.my,
abdulsamad@ucsiuniversity.edu.my

Abstract

In this project the researcher chosen area is the development of an App for forgetful users to take their phones from their cars it functions as a medium to take their phones from their cars. the target users will be senior adults who may be careless or may have memory deteriorating diseases such as dementia or Alzheimer. The researcher used surveys to gather information and used various online resources. The research methodology used by the researcher is modified waterfall model for its simplicity and ability to reverse to different phases. The research methodology used by the researcher is modified waterfall model for its simplicity and ability to reverse to different phases. Overall, the project can be deemed as success as the researcher has successfully developed an App for users to remind them to take their phone before exiting their personal vehicle.

Keywords: Development of an App, Senior Adults, Surveys, Modified Waterfall Model

1. Introduction

Since the introduction of the first smartphone, made by IBM, it was called the IBM Simon it had a calendar, address book, clock, notepad, keyboard, touchscreen and email capability. Since then, humans have been using this device to benefit themselves in daily tasks such as navigating, working and using it as a form of entertainment [3]. But with good the bad always come. Below are the problem that users are facing:

- Elderly users who have dementia may forget to take their smartphones.
- The underground basement parking is too dark and did not notice their smartphones in the holder.
- Buying a new smartphone is affordable but repairing or buying a new car is a liability.

With the problems identified, we propose to build a notification App for forgetful users when the car engine has been shut off. The app will notify the user when the ignition has been turned off. The App will be user friendly and simple. he objectives of the research include:

- To identify why users always forget to take their phones from their cars.
- To design a notifying system
- To develop a prototype Notifying App
- To evaluate the system

2. Methodology
2.1 *Impact of Smartphones in Malaysia*

According to Malaysian Communications and Multimedia Commission (MCMC) the use of smartphones and their apps has dramatically increased since the introduction of affordable 5G smartphones in 2021. According to the survey conducted by MCMC

the number of smartphone users in 2021 is 94.8% compared to a mere 78.0% in 2018 [1], [2]. The higher usage of smartphones indicates people are much more digitally connected. As shown in (Fig.1) below.

2.2 What kind of Apps are used in Malaysia

According to the HPUS 2021, MCMC has found out that communication remained the top activity among smartphone users [1]. The survey has found out that 82.9% of the respondents use text messaging. On which it found that 97.3% of users performed at least once a day. Not only that, 78.9% of users communicated via social media and statistics showed 94.4% users performing it at least once in a day [4].

While in the realm of voice communication, the survey has found that 78.6% of respondents use
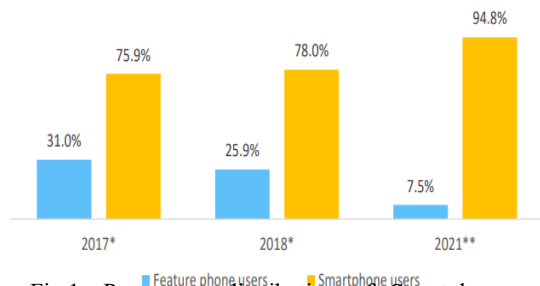
smartphones for voice calls and 80.2% benefit from it once a day.

2.3 Methodology that had been used in app development

For the researcher's case it is developing an App for forgetful users to take their smartphones from their cars. As for this the researcher have discussed with his/her supervisor the choices the researcher has for developing the App. The researcher may have some factors to consider such as suitability of the methodology, the confidence of the researcher using that methodology, and factoring the time constraint [5] [6]. After considering all the factors, the researcher has decided to go with the modified waterfall model as can see in (Fig.2).

According to the survey created by the researcher titled "FYP Survey: Forgetfulness, Phones & Cars" has received a total of 40 respondents. This is all possible without the use of various distribution channels such as Facebook, Word of Mouth, Instagram and family members.

3.1 Survey Results

In the first question the researcher has asked what age is the respondent. According to the survey, 18 respondents are in between 31-60 years old while 15 are young adults in between 18-30 years of age and another 7 which are seniors in the golden years. According to the survey conducted by the researcher. The number of respondents that currently possessed diseases that deteriorates memory are as follow. Respondent that has dementia are 10 respondents while 6 have Alzheimer disease.

3.2 Use Case Diagram

Shown below in (Fig.3) is the Use Case Diagram which represents the methodology used during the development of the App. The Use Case Diagram is used to organize and identify the requirements required for the Forgetful Alarm App.
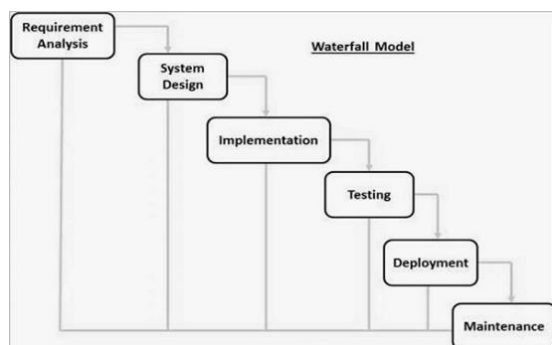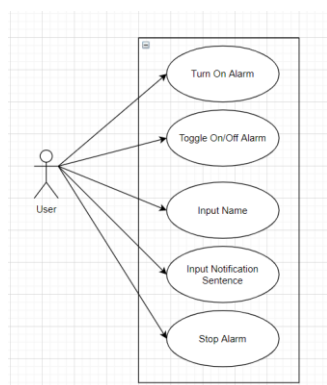

Fig.1. Percentage distribution of Smartphone and feature phone users, 2017 to 2021.

Fig.2. Modified Waterfall model

3. Results and Discussion

Fig.3.Use Case Diagram

Fig.4. Usefulness of the App.

3.3 Implementation

In this section, the implementation of the App is discussed. To develop the App the researcher used Java as the backend programming language and use Extensible Markup Language (XML) to develop the frontend of the App. Android Studio was also used in developing the App [7].

3.4 Evaluation

After successfully developed the App. The researcher will conduct testing to determine whether the App satisfy the requirements that had been identified [8], [9]. To ensure the apps functionality, the researcher has conducted various ways of testing to test the App. The various ways are Unit Testing and User Acceptance Testing.

3.5 User Acceptance Testing

Before the user acceptance test, the researcher would ask the participants which would be selected from the previous survey to download the App and use the app in their daily routine. Participants are asked to commute to work as usual but with a twist of putting their phones in their phone holders and charging it using the cars cigar lighter that can be function as a charging port. In (Fig.4) majority of the participants has agreed that the app is useful.

4. Conclusion

In conclusion all the objectives of the App development had been met and users who used the app had shown interest and can see the potential of this app. Based on the results of this researcher it can be concluded that this app can benefit users in helping them to remember to take their phones from their car. Some recommendations had been elevated from the users during their time using the app. One of the recommendations received during the evaluation is to have reminders to take medication or take an insulin jab. To improve the overall usability of the app the researcher has found some ways to improve the app.

References

1. "Hand Phone Users Survey 2021 (HPUS 2021)," Malaysian Communications And Multimedia Commission, Cyberjaya, 2021.

2. Z. Lazzari, "Technology & How We Communicate," 27 August 2018. [Online]. Available: https://smallbusiness.chron.com/role-cell-phone-communication-today-31479.html. [Accessed 25 May 2023].

3. B. Jung, "What Are the Benefits of Smartphone Technology," 12 February 2019. [Online]. Available: https://smallbusiness.chron.com/role-cell-phone-communication-today-31479.html. [Accessed 25 May 2023].

4. Indeed Editorial Team, "What Is an App? Types of Apps and Examples," 11 March 2023. [Online]. Available: https://www.indeed.com/career-advice/career-development/what-is-an-app. [Accessed 25 May 2023].

5. "Mayo Clinic," 12 October 2022. [Online]. Available: https://www.mayoclinic.org/diseases-conditions/dementia/symptoms-causes/syc-20352013. [Accessed 22 May 2023].

6. "Dementia cases set to ruse 312 per cent by 2050: Is Malaysia prepared?," 8 June 2022. [Online]. Available:

https://adfm.org.my/dementia-cases-set-to-rise-312-per-cent-by-2050-is-malaysia-prepared/. [Accessed 25 May 2023].

7. "Mayo Clinic," 02 February 2023. [Online]. Available: https://www.mayoclinic.org/diseases-conditions/alzheimers-disease/symptoms-causes/syc-20350447. [Accessed 25 May 2023].

8. R. A. Armstrong, "Folia Neuropathological," What causes Alzheimer's disease, p. 20, 2013.

9. E. Ng, "A brief history of cars in Malaysia," 9 August 2022. [Online]. Available: https://mytukar.com/blog/a-brief-history-of-cars-in-malaysia/. [Accessed 27 May 2023].

Authors Introduction

Chee Kin Hoe

He is currently an undergraduate student, pursuing his study in Diploma in Information Technology from Institute of Computer Science and Digital Innovation (ICSDI) at UCSI University, Malaysia.

Dr. Kasthuri Subaramaniam

She is currently an assistant professor at Institute of Computer Science and Digital Innovation (ICSDI), UCSI University, Kuala Lumpur, Malaysia. She earned both her bachelor's degree in computer science and a master's degree in computer science from the University of Malaya. Her research interests include human-computer interaction, etc

Dr. Abdul Samad bin Shibghatullah

He received the bachelor accounting degree from Universiti Kebangsaan Malaysia, Bangi, Malaysia in 1999, the M.Sc. degree in computer science from the Universiti Teknologi Malaysia, Skudai, Malaysia in 2002, and the Ph.D. degree from the Brunel University of Uxbridge, United Kingdom. He is currently Associate Professor at UCSI University, Kuala Lumpur, Malaysia. His current research interests include optimization, etc.

Developing a Prototype Hand Gesture Recognition System in Interpreting American Sign Language

Kong Seh Chong
Institute of Computer Science and Digital Innovation, UCSI University, Malaysia

Kasthuri Subaramaniam
Institute of Computer Science and Digital Innovation, UCSI University, Malaysia

Ismail Ahmed Al-Qasem Al-Hadi
Institute of Computer Science and Digital Innovation, UCSI University, Malaysia
Email: 1002058020@ucsiuniversity.edu.my, kasthurisuba@ucsiuniversity.edu.my, ismailAhmed@ucsiuniversity.edu.my

Abstract

Hand gestures of sign language is a form of non-verbal communication which have been used by most people in their daily life. Sign language is not only used by people with speaking issues but it is also unconsciously used by normal people during their daily interaction with others. This is because it is a way to express their current feelings or the meaning they wanted to convey to others. On other hand, sign language is an important alternative used by people with hearing impairment or speaking obstacles so they can communicate with others. However, not everyone from all walks of life has learned sign language so there will be problems of interaction between them and people with speaking issues. Thus, this research focuses on developing a hand gesture recognition system for accurately interpreting American Sign Language (ASL) so that it can deliver a message that can be understood by others and enable efficient communication. In our system, it will utilize computer vision techniques to analyze hand postures and movements which will include hand sign recognition, finger tracking, and motion estimation. With the pre-developed libraries like OpenCV and MediaPipe are employed in the system so it can recognize and classify ASL gestures based on extracted features. Extensive datasets of ASL hand gestures are collected and annotated to enhance the system's accuracy and robustness. The developed system aims to improve human-computer interaction, enabling seamless communication between deaf individuals and technology. The potential applications include real-time interpretation of ASL gestures for enhanced accessibility and inclusivity.

Keywords: Human-Computer Interaction, American Sign Language, Hand Gesture Recognition System.

1. Introduction

In this age of technology, there are many IT related tools or systems that provide various benefits to all walks of life which include Hand Gesture Recognition System (HGRs). The Hand Gesture Recognition System is a type of technology that uses computer vision and machine learning to interpret the motion and sign of individuals [1]. Therefore, this technology can contribute a significant transformation to society nowadays by providing a particular group of people huge benefits who are born with hearing impairment or muted. This is because people with hearing impairment or mute usually rely on sign language, and it is an important tool of communication for them. There is a specific meaning from each gesture in sign language which can construct a complex meaning with the combination of various gestures. In American sign language, where each hand sign will represent different alphabets from A to Z which then can form a specific word. As a result, this project purpose is about developing a Hand Gesture Recognition System that can help in reading sign language. With this system, it can help people with hearing impairment or speaking disabilities to communicate with others more efficiently as well as allow non-signers to get know more about what actually a sign language is [7], [8].

Hand gestures are something that are typically used in daily communication including communicating with people having hearing problems. According to the research, hand gestures can be defined as a type of body language that can be expressed through the center of the palm, finger position and the shape constructed by the hand [9]. There are two classifications of hand gestures which will be the static and dynamic. The static

gestures mostly have to do with the shape of the hand whereas the dynamic gestures deal with a variety of hand movements [9]. For instance, we can find static gestures from the American Sign Language (ASL) at the figure (Fig. 1) in which most of its hand gestures are in a particular shape or pattern that represent an alphabet excluding j and z. Dynamic gestures are mostly related to body language or hand gestures that require movement such as waving, shaking other hands, or even showing the J and Z letter in ASL [11]. It is hard to track the dynamic gestures so in our project we will be focusing more on the ASL static gestures and think of future improvement for the dynamic hand gestures recognition system.

Fig. 1 American Sign Language [10].

2. Methodology

2.1 Comparison Of Different Methodology in Hand Gestures Recognition System

A comparison of methodology is shown in the table below (Table 1).

Table 1 Comparison Table of methodology
in Hand Gesture Recognition System

Sr. No	Paper Name	Year of Public	Techniques	Advantages
1	Hand Gesture Recognition via Lightweight VGG16 and Ensemble Classifier [4]	2022	1.CNN 2.Random Forest Classifier	• Random forest is great in handling large input variables. • Having the highest classification rate.
2	Indian Sign Language Recognition Using Random Forest Classifier [2]	2021	1.Random Forest Classifier	• High accuracy even on complex datasets. • Robust to missing data as users may not always perform gestures perfectly.
3	A Vision-based System for Recognition of Words used in Indian Sign Language Using MediaPipe [6]	2022	1.Random Forest Classifier 2.MediaPipe	• Able to learn complex relationships between the features of the sign language gestures. • Achieving high accuracy in 97.4%.
4	Communicating with the Deaf and Hard of Hearing through Automatic Arabic Sign Language Translator [5]	2021	1.Feature Extraction 2.Random Forest Classifier	• Achieving best recognition accuracy out of all other classification algorithms.
5	Smart Communication System Using Sign Language Interpretation [3]	2022	1.MediaPipe 2.Light Weight Random Forest Classifier	• Achieving good accuracy on recognition about 94.69% • Ability in adapting to different sign language datasets by using different feature extraction techniques and hyperparameter tuning

Fig. 2 Proposed concept of Hand Gesture Recognition System
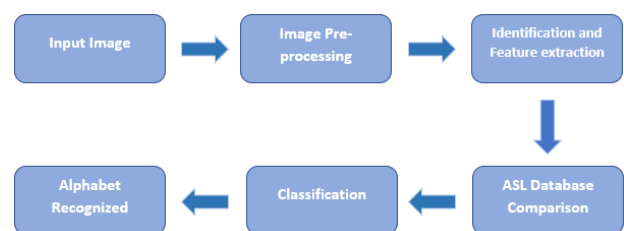
2.2 Analysis

We will be revealing the outcomes and analyzing the information about the Hand Gesture Recognition System in interpreting ASL from the survey organized that was distributed on the network through emailing and messaging (Fig. 2). Based on the outcomes of the survey, we will know about whatever this project can provide a system in interpreting ASL for

communication between normal people and deaf or people with hearing problems.

2.3 Methodology

In this report, we have employed a quantitative method to gather information regarding the Hand Gesture Recognition System in interpreting ASL. The chosen approach for data collection was a survey form created using Google Forms. The survey form consisted of 2 parts, the first part is about the demographics of the respondents and the second part is about the system inquiries. There are a total of 16 questions that focus on which the 12 questions from the total are mostly asking about the Hand Gesture Recognition System and ASL. The questions included in the survey form were carefully selected to extract essential information required for our research. We aimed to understand more about the Hand Gesture Recognition system so that we can successfully develop the system in interpreting ASL.

The survey form consisted of multiple-choice questions that allowed respondents to select multiple answers from a given list of options. This format enabled us to capture a range of basic understanding of the respondents in the Hand Gesture Recognition System. In addition to multiple-choice questions, we also included rating questions in the survey form. These questions employed a scaling approach to measure respondents' agreement or disagreement with specific statements related to the Hand Gesture Recognition System. This format allowed us to assess the level of agreement towards certain measures or opinions about the Hand Gesture Recognition System in interpreting ASL.

By utilizing the survey form with these question formats and employing various distribution methods, we intended to gather comprehensive insights into the Hand Gesture Recognition System. For reference, the appendix of this report contains 15 survey questions pertaining to the Hand Gesture Recognition System in interpreting ASL.

3. Results and Discussion

In this section, it will be discovering all the validation and testing that are done after the development of the Hand Gesture Recognition system. This chapter is a very significant part that needs to be taken in the entire process of the project development as it will be the sources to know what are the weaknesses and limitations of the current system and the improvement that it needs to enhance in the future. The results are shown in Table 2 and Table 3.

Table 2. Results of Unit Testing

Procedure Description	Expected Result	Actual Result	Test Result
Click Start Camera Button	The camera is started after button was clicked	Worked as expected	Pass
Click Stop Camera Button	The camera frame is stopped after the button was clicked	Worked as expected	Pass
Recognizing the ASL hand gesture	All the static ASL hand gesture is recognized successfully on both right and left hand	Worked as expected	Pass
Click Exit Button	Close the entire program	Worked as expected	Pass
Click Save Character Button	Able to save the recognized ASL hand gesture to text file	Worked as expected	Pass
Click Clear Texts Button	Able to clear the textbox with the recognized ASL hand gesture	Worked as expected	Pass
Enable to save message as history	Click the Save Messages button to store it as history at history message page	Worked as expected	Pass
Click History Message Button	Open up a new frame that redirect to the History Message page	Worked as expected	Pass
Click ASL Display Button	Allow users to view on the ASL gestures information on new frame	Worked as expected	Pass
Click Testing Page Button	Open up a new frame that redirect to the testing page	Worked as expected	Pass
On History Message page, double click on message	Allow long message to be view in full on the label of "Full Message:"	Worked as expected	Pass
On History Message page, click clear button	Able to clear the full shown messages on the below the label of "Full Message:"	Worked as expected	Pass
On History Message page, click clear history button	All the history on the page will be able to clear	Worked as expected	Pass
Click Text to Speech	Able to hear the voice of messages to speech	Worked as expected	Pass
Inside message box of History Message page	Information of saved message will be recorded such as Date, Time and Message contents	Worked as expected	Pass
On testing page, click start camera button	It will activate the camera frame for hand gesture recognition process	Worked as expected	Pass
On testing page, click stop camera button	It will stop the operation of camera frame	Worked as expected	Pass
On testing page, information box display	After camera operated, it will show different information such as recognized alphabets, hands, interpretation and FPS	Worked as expected	Pass

Table 3. Results of System Testing

Procedure Description	Expected Result	Actual Result	Test Result
Recognizing ASL hand gestures through webcam	To ensure the users were able to recognize the ASL hand gesture through the start camera.	Worked as expected	Pass
Saving the ASL hand gestures	To ensure that the users are able to save the recognized ASL alphabets to the textbox.	Worked as expected	Pass
ASL information display	To ensure that the users are able to view types of ASL hand gestures at the page.	Worked as expected	Pass
Saving the text in Message History	To ensure that the users are able to save and view back the history of text.	Worked as expected	Pass
Testing feature for users	To ensure that the users are able to familiarize themselves with the ASL hand gesture system before directly using it.	Worked as expected	Pass
Converting Text to Speech	To ensure the messages from the deaf or people with hearing problem can be heard by the others	Worked as expected	Pass
Stop the system operation	To ensure that the users are able to quit the system by clicking the exit button.	Worked as expected	Pass

4. Conclusion

In conclusion, the project is researched and analyzed comprehensively in order to ensure that the objectives and aims that have been decided at the proposal phase are accomplished. After the development and analysis of the hand gesture recognition system, it can be stated and concluded that all objectives that were planned in the proposal phase have all been accomplished. Since the hand gesture recognition system is specially created for people with hearing problems to have an interaction with others through the gesturing and recording characters of ASL which form a text to bring the meaning they are expressing. At the same time, it may also serve as an educational tool to help people who might be interested in American Sign Language and want to learn it by using the system. This is how the research project's key goals will be accomplished. However, there are no perfect system exist in the world which include our hand gesture recognition system. For instance, one of the limitations of our hand gesture recognition system are unable to recognize in dark environment or classify dynamic gestures. Therefore, there will be future works to be accomplish in order to conquer these limitations stated in this project. Nevertheless, our hand gesture recognition system has already attained the objectives and solve some of the problems stated in the project.

References

1. A. Kumar C, C. Venkatesh, L. Vaishnavi D A, and H. S, Computer vision based Hand gesture recognition system, vol. 20, no. 7, pp. 2859–2866, Jul. 2022. doi:doi: 10.14704/nq.2022.20.7.NQ33365

2. A. S, A. Potluri, S. M. George, G. R, and A. S, "Indian sign language recognition using random forest classifier," 2021 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), 2021. doi:10.1109/conecct52877.2021.9622672

3. D. Bisht, M. Kojage, M. Shukla, Y. P. Patil, and P. Bagade, "Smart Communication System Using Sign Language Interpretation," 2022 31st Conference of Open Innovations Association (FRUCT), May 2022. doi:10.23919/fruct54823.2022.9770914

4. E. L. Ewe, C. P. Lee, L. C. Kwek, and K. M. Lim, "Hand gesture recognition via lightweight VGG16 and ensemble classifier," Applied Sciences, vol. 12, no. 15, p. 7643, 2022. doi:10.3390/app12157643

5. G. Latif, J. Alghazo, N. Mohammad, and R. Alghazo, "Communicating with the deaf and hard of hearing through Automatic Arabic Sign Language translator," Journal of Physics: Conference Series, vol. 1962, no. 1, p. 012055, Mar. 2021. doi:10.1088/1742-6596/1962/1/012055

6. S. Adhikary, A. K. Talukdar, and K. Kumar Sarma, "A Vision-based system for recognition of words used in indian sign language using MediaPipe," 2021 Sixth International Conference on Image Information Processing (ICIIP), Feb. 2021. doi:10.1109/iciip53038.2021.9702551

7. Thwe, P.M., (2019) 'Hand Gesture Detection and Recognition System: A Critical Review'. International Journal of Computer (IJC), vol. 32, no. 1, pp. 64-72.

8. Li, G. et al. (2017) 'Hand gesture recognition based on convolution neural network', Cluster Computing, 22(S2), pp. 2719–2729. doi:10.1007/s10586-017-1435-x.

9. M. Oudah, A. Al-Naji, and J. Chahl, "Hand gesture recognition based on Computer Vision: A review of techniques," Journal of Imaging, vol. 6, no. 8, p. 73, 2020. doi:10.3390/jimaging6080073

10. M. Yasen and S. Jusoh, "A systematic review on hand gesture recognition techniques, challenges and applications," PeerJ Computer Science, vol. 5, 2019. doi:10.7717/peerj-cs.218

11. N. Bargellesi, M. Carletti, A. Cenedese, G. A. Susto, and M. Terzi, "A random forest-based approach for hand gesture recognition with wireless wearable motion capture sensors," IFAC-PapersOnLine, vol. 52, no. 11, pp. 128–133, 2019. doi:10.1016/j.ifacol.2019.09.129

Authors Introduction

Mr. Kong Seh Chong

He is a student majoring in Bachelor of Computer Science (Honours) at Institute of Computer Science and Digital Innovation (ICSDI), UCSI University, Kuala Lumpur, Malaysia.

Dr. Kasthuri Subaramaniam

She is currently an assistant professor at Institute of Computer Science and Digital Innovation (ICSDI), UCSI University, Kuala Lumpur, Malaysia. She earned both her bachelor's degree in computer science and a master's degree in computer science from the University of Malaya. Her research interests include human-computer interaction, etc.

Dr. Ismail Ahmed Al-Qasem Al-Hadi

He is currently an assistant professor at Institute of Computer Science and Digital Innovation (ICSDI), UCSI University, Kuala Lumpur, Malaysia. He has a PhD in Intelligent System. His research interests include AI and Data Mining.

Minature Resource Planner

Adim Khalid Aldireejah
Institute of Computer Science and Digital Innovation, UCSI University, Kuala Lumpur, Malaysia

Kasthuri Subaramaniam
Institute of Computer Science and Digital Innovation, UCSI University, Kuala Lumpur, Malaysia

Ghassan Saleh Hussein Al-Dharhani
Institute of Computer Science and Digital Innovation, UCSI University, Kuala Lumpur, Malaysia
Email: 1002061045@ucsiuniversity.edu.my, kasthurisuba@ucsiuniversity.edu.my,
ghassan@ucsiuniversity.edu.my

Abstract

Due to the complexity and high cost of typical Enterprise Resource Planning (ERP) systems, medium-sized enterprises (SMEs) often face challenges in their implementation. To address this, the article develops a tailored, cost-effective ERP system specifically designed for SMEs, considering their unique needs and limitations. The research analyzes existing ERP systems in large organizations and conducts surveys and interviews to identify common issues faced by SMEs when adopting standard ERP systems. The paper's main goal is to create an affordable and user-friendly ERP system that empowers SMEs to overcome financial obstacles in an increasingly digital world. The system addresses key business processes such as financial accounting, employer management, and procurement. Additionally, the report highlights the interest homeowners have in using ERP systems for household tasks, indicating a substantial market and demand. Emphasizing the importance of a flexible Home Resource Planning (HRP) system, the study aims to design a cost-effective ERP solution suitable for both homes and SMEs. The proposed ERP system's impact on SME operating performance, productivity, and decision-making is thoroughly examined. The conclusion underscores the significance of customized ERP systems for homes and SMEs in the current digitized environment. The developed small-scale ERP system proves successful in helping SMEs and homes overcome financial challenges and reach their full potential.

Keywords: SMEs, Resource Planner, ERP

1. Introduction

Organizations nowadays are trying to attain optimal efficiency, get an edge over their competitors, and accomplish their objectives as quickly as possible. ERP, or enterprise resource planning, aids in achieving previous corporate goals. As a definition, ERP is a piece of software that contains all the components necessary for a corporation to be managed and operated with maximum effectiveness, avoiding the most possible issues. Enterprise resource planning works by integrating technology from many industries— including human resources, finance, manufacturing, supply chain, and more—in one environment. In other words, it automates numerous business processes with the least amount of effort while remaining in the same setting. Most large businesses globally have already embraced ERP, and more and more small and medium-sized businesses are trying to do the same to compete; they find it both cost-effective and necessary to keep pace [1]. Although the scope and close integration of ERP has only just become available, the market for ERP has been expanding quickly, with more customers and users on a daily basis.

Any firm may gain a lot from using an ERP system. For instance, it can improve and automate the majority of daily tasks an organization must do by removing the need for human data entry, saving time and money [2]. Second, ERP keeps everything organized and in a single system, which removes the need for numerous technologies. Third, ERP can quickly analyse and identify the saved data to provide managers with operational insights. Overall, enterprise resource planning provides seamless data flow between any organization's sectors while maintaining customizability so that it can follow the business needs and strategies.

It has become a must for the future of any establishment that is looking to improve and remain on the popular side of the market [3]. Throughout this proposal, a deeper analysis of the significant characteristics, challenges of implementation, and difficulties related to ERP systems will be done to ultimately gain a thorough understanding of their

function in promoting organizational success. also mentioning some issues and their solutions, using my approach to resolve them, and expanding the accessibility of ERPs to everyone, without exception.

2. Methodology

2.1 Brief History

ERP history traces back to the 1960s with material requirements planning (MRP) systems, pioneered by J.I. Case in collaboration with IBM. The increasing complexity of workplaces led to the need for cross-functional data flow in areas like decision-making, product ordering, inventory management, accounting, human resources, and distribution. This complexity gave rise to the first ERP system, aimed at consolidating diverse functions into a unified platform. ERP systems emerged commercially in the late 1980s and early 1990s, initially designed for the demands of large, complex organizations. Recognized for their complexity and high costs, ERP systems have evolved over time [4].

2.2 Existing System In The Market

SAP ERP, by SAP, a leading corporate software solutions provider, automates key business operations like accounting, sales, and human resources. SAP ECC, the main component, handles transactions, reporting, and integration. SAP S/4HANA, the latest version, offers real-time data processing and improved analytics. Known for a well-explained UI, SAP ERP is popular among giants like PayPal and NBA. Recently integrated with cloud capabilities, SAP ERP provides a customizable cloud solution for global business models and revenue growth [5].

Oracle ERP, from Oracle Corporation, optimizes finance, supply chain management, and human resources. Renowned for dependability and scalability, it suits both small and large businesses. Oracle ERP is customizable to meet industry needs and offers extensive capabilities in data analytics, reporting, and system integration. With a solid reputation and broad support services, Oracle ERP has been a market leader for over a decade.

Microsoft Dynamics 365, part of Microsoft's business application package, is a popular ERP system. It covers accounting, supply chain management, sales, and customer support. Unique for seamless integration with Microsoft products like Office 365, it provides a consistent user experience. Microsoft Dynamics 365 ERP offers advanced analytics, AI-driven insights, and cloud-based deployment options, facilitating operational growth, insights, and automation.

Odoo ERP, an adaptable open-source system, is among the fastest-growing ERP systems. It offers tools for accounting, manufacturing, and sales. Odoo's uniqueness lies in its modular approach, allowing organizations to integrate only the required modules. With an approachable UI, community-driven ecosystem, and cloud-based options, Odoo is an affordable and scalable solution for businesses of all sizes and sectors. Known for fitting SMEs, Odoo is simple and free to test for new users.

2.3 ERP Benefits

Streamlined Operations: Streamlined operations are achieved through the reduction of delivery and cycle times, strategically minimizing delays in data retrieval and reporting. By optimizing these processes, the organization enhances overall efficiency and responsiveness, contributing to a more agile and competitive operational framework.

Cost Reduction: The realization of cost reduction stems from a comprehensive, enterprise-wide review of organizational decisions. This holistic approach not only saves valuable time but also enhances control over financial aspects. Through the identification and elimination of inefficiencies, the organization not only achieves monetary savings but also establishes a more robust and cost-effective foundation for future endeavours.

Adaptability and Flexibility: The system's inherent adaptability and flexibility empower the organization to undergo simple reorganization swiftly. This adaptability extends to accommodating modifications in business processes, ensuring that the enterprise remains agile in response to evolving market dynamics. This capability becomes a cornerstone for sustained innovation and resilience in an ever-changing business landscape.

E-Commerce and E-Business Integration: The system seamlessly supports online shopping experiences and fosters a collaborative environment conducive to effective teamwork. This collaboration, in turn, enables effortless integration with e-commerce systems. By embracing and facilitating digital interactions, the organization not only capitalizes on the growing trend of online commerce but also positions itself for enhanced connectivity and synergy across various aspects of its e-business operations.

2.4 ERP Drawbacks

Time-consuming deployment, the time and resources required to accomplish an ERP deployment are frequently considerable [6]. The collecting of requirements, system configuration, data migration, testing, and user training all need resources from organizations.

Organizational disruption can result from the implementation phase, which forces staff members to become used to new procedures and systems. Employee

reluctance to change and temporary productivity declines may result from this shift.

Challenges with complexity and customization, because of their extensive nature, ERP systems may be very complicated. Adapting the system to a given organization's requirements can be difficult and may call for advanced technical knowledge.

2.5 ERP In Different Industries

Healthcare Industries: ERPs are pivotal in healthcare, streamlining processes like patient scheduling, inventory control, and invoicing. Key benefits include increased operational effectiveness and compliance with healthcare data protection laws like HIPAA.

Financial Sector: In finance, ERPs are indispensable for managing financial data, reducing data loss risks, and ensuring compliance with industry laws, emphasizing the importance of rapid and secure access to accounting information [7].

Education Sector: ERPs play a vital role in education, simplifying tasks like resource organization, grading, and student enrollment. They provide colleges and schools with seamless support for effective operations.

Manufacturing Industries: In manufacturing, ERP systems prove invaluable throughout the workday, efficiently managing logistics, invoicing, and ensuring data organization [8]. Beyond these functionalities, they also assist employees in maintaining accessible personal data, highlighting the depth of their capabilities.

2.6 Comparison Table

Table 1 shows the comparison of the ERP providers.

Table 1. ERP Comparison Table

ERP Provider	Odoo	Microsoft 365	Oracle	SAP
Deployment	On-premises, Cloud	On-premises, Cloud	On-premises, Cloud	On-premises, Cloud
Open source	Yes	No	No	No
Customization	Highly	Moderate	Moderate	Moderate
Scalability	Yes	Yes	Yes	Yes
User-Friendly	Yes	Yes	Yes	Yes
Cost-effective	Yes	Varies	Varies	Varies
Popularity	Growing	Popular	Popular	Popular
Community	open	Strong	Active	Active

2.7 ERP Implementation

When implementing an ERP system, a comprehensive plan is crucial. It begins with a detailed requirements analysis to identify organizational issues and objectives. Vendor selection follows, involving research to choose an ERP provider that aligns with the organization's

needs. Project planning outlines the implementation schedule, resource allocation, and key tasks after vendor selection. Configuration and customization of the ERP system to match unique requirements and procedures follow, including data transfer with accuracy. User training introduces the new system through sessions, workshops, and documentation. Testing and quality control ensure functionality, with various tests like user acceptability and functionality testing. Upon readiness, the system is deployed, with ongoing support to address user queries. Continuous improvement is emphasized, with regular monitoring and feedback collection for streamlined procedures.

2.8 Quantitative Methodology

Data collection involved in-depth interviews and organized surveys to gather user and stakeholder insights. Surveys targeted user groups, focusing on quantitative aspects like performance indicators, system usability, and user satisfaction. A comprehensive analysis of comparable ERP systems enriched contextual knowledge, dissecting and comparing various systems to understand ERP dynamics. The investigation explored the advantages, disadvantages, and unique features of existing ERPs for designing and developing the Home Resource Planning (HRP) system.

2.9 Survey Structure

The survey comprised two parts. The first assessed respondents' ERP knowledge, experience, and understanding of fundamental features. The second part delved into comprehensive testing methods, evaluating the system's stability, performance, and user satisfaction in challenging circumstances. Survey sections included Knowledge and Familiarity with ERP, HRP Interest, ERP Experience, Future ERP Use, and an Additional Comments section for open-ended feedback.

2.10 Qualitative Methodology

Qualitative methods, including interviews and personal comments, supplemented statistical approaches for a holistic understanding. One-on-one interviews covered participants' experiences with ERP systems, potential benefits of an HRP system for daily tasks, and desired features. Questions focused on familiarity with ERP systems, willingness to use an HRP system, valued features, anticipated impacts, and challenges faced with ERP systems in the past.

3. Implementation

3.1 Functional Requirements

User registration and authentication: Users are able to set up accounts with distinctive passwords and

Adim Khalid Aldireejah, Kasthuri Subaramaniam, Ghassan Saleh Hussein Al-Dharhani

usernames. During the login process, the system authenticates and verifies the user's credentials.

User-Specific Views: Based on their role and preferences, every user is shown a customized dashboard or view. The interface is customized by the system to provide appropriate information, including appointments, finances, people, and duties, for various users.

Service Selection: A variety of services are available for users to select from, such as appointment scheduling, money, personnel management, and chore organization.

Data Entry and Management: Within the service category they have chosen, users can enter and manage data. To ensure data accuracy, the system offers capabilities for data entry, modification, and deletion.

Notification and Reminder System: Users are able to choose how they want to be notified and reminded. Reminders and automatic notifications are sent by the system in accordance with user-specified schedules and actions.

Data Integration: Users may access and use information from several modules by means of the ERP system's smooth integration of data from the other services. Maintaining data consistency helps to guarantee accuracy and dependability.

3.2 Non-Functional Requirements

Security: To safeguard user data and preserve privacy, the ERP system complies with industry-standard security standards. Sensitive data is protected by robust encryption and authentication processes.

Scalability: Exhibits the capacity to increase with the number of users and volume of data. meets predetermined performance goals and maintains peak performance as the system grows.

Performance: Reduces downtime and guarantees quick user request responses. Supports several users concurrently and effectively handles data transfers without appreciable performance reduction.

Usability: The design of the user interface places a high value on simplicity and clarity, making it easy for users to navigate and finish activities.

Dependability: An ERP system with high dependability has fewer mistakes and interruptions.
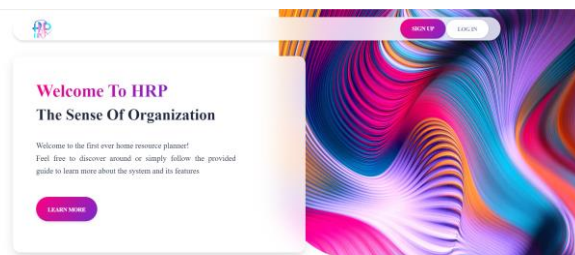
3.3 Web-based Resource Planner

Fig 1. HRP Home-Page

The Homepage in Home Resource Planning (HRP) is an intuitive user interface that offers a smooth sign-up or log-in process for current users (Fig. 1). The design places a high value on simplicity, guaranteeing users a seamless onboarding experience. Furthermore, the platform provides a brief yet useful feature summary, enabling customers to rapidly understand the possibilities it provides. A dedicated "Learn More" page offers a thorough examination of the web application for individuals looking for more specific details. This page functions as an in-depth manual, providing information on the features, and advantages.

Fig 2. HRP Home-Page Responsiveness

It's also important to note that the majority of its pages adjust to the screen size, even on mobile devices. Whether on a laptop, tablet, or desktop, HRP will try to adjust the size to fit the screen (Fig. 2).

Fig 3. HRP Main Page

Similar to the home page, the system home page is simple to navigate; users may select the service they want to use here as shown in Fig. 3.. Additionally, the header and footer are always there and allow users to

select from the services offered on each page. At this point, it is worth noting that the purpose of HRP is to provide similar services any ERP can but with easier navigation.

Fig 4. HRP Services

One of the features offered by the web-based app is the ability for users to add individuals to the database, modify or remove their information, or even look for a specific record as shown in Fig. 4.. Each record has a unique ID to help with identification, and each user will see their data in a unique way, guaranteeing that each user's data is private and personal.
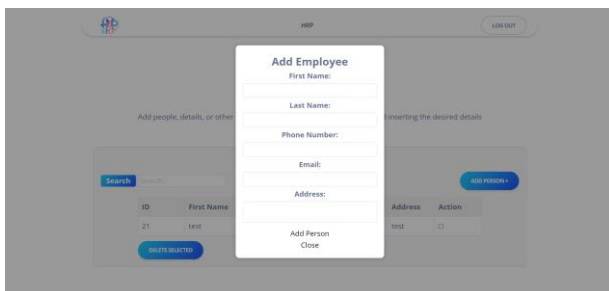

Fig 5. Pop-ups and Forms

The user will always be informed if an activity was successful or not with pop-up pages and messages that appear after any action (Fig. 5). In any online application, error management is Considered one of the most important features for a web-app/website to have.
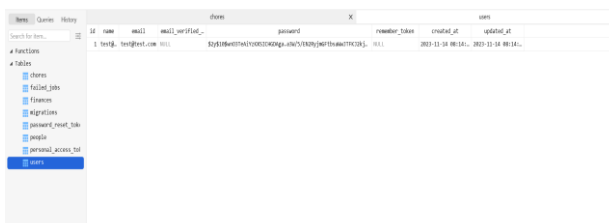

Fig 6. Database Schema

The ERP system's database is distinguished by advanced development and maximum functionality. With a strong performance, it effectively organizes and maintains large amounts of data, which adds to the system's overall dependability and efficiency. Because the database is designed to be scalable, data can be stored and retrieved quickly, guaranteeing a flexible and responsive system. Its challenging architecture creates a safe and organized space for the archiving and retrieval

of vital corporate data while accommodating the various needs of ERP modules (Fig. 6). The smooth functioning of the database is essential to the overall effectiveness and success of any ERP system.
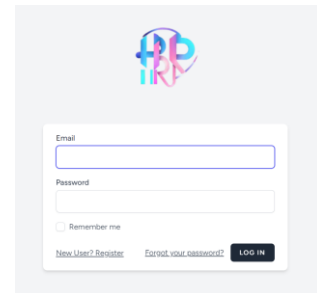

Fig 7. User Authentication

Since HRP uses user authentication (Fig. 7), that allows each user to have their own view which will allow data to be secure and no mix-ups can happen between users.

4. Conclusion

As a result of their ability to offer a centralized system that unifies several company processes such as supply chain management, manufacturing, finance, and human resources—resource planners are incredibly helpful to businesses. They have been becoming more popular by the day owing to their utilization in the business world. Small firms can increase their productivity and efficiency considerably by using an ERP system, but it does require careful design, sufficient funding, and good user support and training. A number of other implementation-related aspects are covered, such as small business owners' as well as homeowners' interest in utilizing such a system, their familiarity with ERP software, and their long-term goals for system adoption or upgrades. The analysis also showed that different small business owners have different long-term plans for adopting or upgrading their systems; some want to move to more simple solutions, while others would rather learn more about them in general.

Additionally, Homeowners' interest in utilizing an ERP system for household tasks is also emphasized. The research revealed a significant market and need for such systems, with over 80% of participants expressing a strong interest in an HRP system designed for duties connected to the household. In order to satisfy the needs of a diverse variety of customers with a wide range of user requirements and expectations, the research highlights the significance of having an adaptive and versatile HRP system.

Introducing HRP to the market and matching it with the current trends might have a lot of benefits for homes and company owners alike, as well as for anyone looking to improve efficiency and organization in their everyday lives. An HRP system is a promising addition to homes and small enterprises, with the potential to revolutionize resource management.

Adim Khalid Aldireejah, Kasthuri Subaramaniam, Ghassan Saleh Hussein Al-Dharhani

References

1. K. Hayes, "ERP industries: What industries use ERP software?," SelectHub raquo, 2023.
2. E. Watson, "Using ERP systems in education - researchgate," Using Erp Systems In Education, 1999.
3. S. Singh, S. Singh, and S. C. Misra, "Post-implementation challenges of ERP system in pharmaceutical companies," International Journal of Quality & Reliability Management, 2023.
4. B. Sampath and K. A. Mathappa, "Analyzing the Impact of ERP on Improving Business Operations Using Oracle," Latest Trends in Multidisciplinary Research & Development, vol. 35, 2023.
5. U. Malhotra, "Secure and Compatible Integration of Cloud-Based ERP Solution: A Review," International Journal of Intelligent Systems and Applications in Engineering, vol. 11, no. 9s, pp. 695-707, 2023.
6. S. L. Z. Zahra and T. Siswanto, "Implementation of Odoo-Based ERP in The Case Study of Micro, Small, and Medium Enterprises (MSME) 'Woody Moody Jakarta'," Intelmatics, vol. 3, no. 2, pp. 68-77, 2023.
7. B. Arifianto, S. S. Azhar, and D. F. Murad, "Evaluation and Recommendation of Odoo Enterprise Resource Planning System Operation & Maintenance Module," in 2023 8th International Conference on Business and Industrial Research (ICBIR), IEEE, May 2023, pp. 102-107.
8. S. Tongsuksai, S. Mathrani, and K. Weerasinghe, "Influential characteristics and benefits of cloud ERP adoption in New Zealand SMEs: a vendors' perspective," IEEE Access, vol. 11, pp. 23956-23979, 2023.

Authors Introduction

Mr. Adim Khalid Aldireejah

He is currently pursuing Bachelor of Computer Science (Hons) at Institute of Computer Science and Digital Innovation (ICSDI), UCSI University, Kuala Lumpur, Malaysia.

Dr. Kasthuri Subaramaniam

She is currently an assistant professor at Institute of Computer Science and Digital Innovation (ICSDI), UCSI University, Kuala Lumpur, Malaysia. She earned both her bachelor's degree in computer science and a master's degree in computer science from the University of Malaya. Her research interests include human-computer interaction, human personality types, etc.

Dr. Ghassan Saleh Hussein Al-Dharhani

He is currently an assistant professor at Institute of Computer Science and Digital Innovation (ICSDI), UCSI University, Kuala Lumpur, Malaysia. His research interests include cybersecurity, ethical hacking and machine learning.

Developing Cloud-based Sportswear Website

Wei Yee Lim

Institute of Computer Science and Digital Innovation, UCSI University, Malaysia

Kasthuri Subaramaniam

Institute of Computer Science and Digital Innovation, UCSI University, Malaysia

Raenu Kolandaisamy

Institute of Computer Science and Digital Innovation, UCSI University, Malaysia
Email: weiyee55@gmail.com, kasthurisuba@ucsiuniversity.edu.my, raenu@ucsiuniversity.edu.my

Abstract

This research is going to develop a cloud-based sportswear website that can handle high volumes of traffic during events or seasonal sales, assist customers in selecting the right size for their sportswear, the interface should have options, size charts, fitting instructions, and recommendations. A study on the existing system will be conducted, design, develop and evaluate on the sportswear website also will be carried out. The system development approach used on this website is Rapid Application Development (RAD) model. The reason to choose RAD is because this method is time efficient, flexibility and adaptability. With RAD, functional software is delivered quickly. We can quickly build prototypes, iterate on them, and gather feedback from stakeholders with a website project. There are four of the steps in Rapid Application Development (RAD) methodology life cycle, which is define the requirements, prototypes, rapid construction and feedback gathering, and cutover.

Keywords: Cloud computing, Sportswear trend, E-commerce

1. Introduction

The Internet, especially with the advent of cloud computing, has brought about significant changes in our world. Cloud computing, a highly advanced technological product, has emerged as a key player in the 21st century's era of globalization, promoting social normalization and expanding social equality [1]. This innovation relies on the World Wide Web and the rapid growth of Internet technology, offering a solid and scalable security structure crucial for its real-world applications.

Distinguishing e-commerce from traditional commerce, e-commerce conducts transactions online, allowing users to shop, trade, auction, and access various services. The advantages of e-commerce include lower costs, increased convenience, and access to a broader range of products. Cloud computing, viewed as the next revolution in the science and technology industry, is expected to significantly impact enterprise operations when combined with e-commerce. It eliminates the complexities and costs associated with traditional business applications, providing shared infrastructure and scalability.

Cloud computing provides computer resources as a service over the internet, offering dynamic scalability and virtualization. Users' information is stored remotely, allowing accessibility anytime, anywhere. The three basic cloud computing service models are Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS), each serving distinct purposes in the cloud ecosystem [6].

The e-commerce landscape is evolving, and cloud computing is playing a pivotal role in shaping its technical architecture. Different types of e-commerce, such as B2B, B2C, C2C, C2B, B2A, and C2A, cater to diverse business models and transactions. This project specifically adopts the B2C model, focusing on selling athletic apparel and related items directly to individual users through a sportswear website [2].

Amazon EC2 (Elastic Compute Cloud) exemplifies how cloud computing can benefit e-commerce enterprises. Instead of purchasing hardware and software, businesses can rent them, reducing system building costs. EC2, a popular cloud system, allows users to rent cloud-based applications, providing flexibility and cost-effectiveness. In this project, EC2 is

chosen as the cloud server to store website data, leveraging the benefits of Amazon Web Services (AWS).

The impact of social media on customer buying behavior is noteworthy, with 3.6 billion users across various platforms. Social media platforms enhance customer connection, broaden reach, and provide businesses with opportunities to interact with and influence customer behavior. Building brand equity through social media is recognized as a valuable asset in the contemporary market [3].

Sportswear has become a fashion trend, influenced by changes in lifestyle, the demand for functional performance clothing, and the growing interest in fitness outerwear [7]. Activewear manufacturers thrive as fitness-conscious consumers seek appropriate gear. The activewear market, particularly in the United States, is robust, with Nike and Adidas leading in revenue. Women's sportswear sales surpass those of men's and children's, and this trend is expected to continue into the 2020s [4], [5], [8].

2. Methodology

The Rapid Application Development (RAD) approach revolutionizes software application development with its iterative and user-centric methodology. Primarily designed for projects requiring ongoing user participation, RAD's strength lies in its adaptability and responsiveness to changing requirements. Contrary to the traditional waterfall development paradigm, RAD, introduced in 1991, thrives on flexibility and efficiency, making it particularly effective for managing modest and medium-sized projects.

The RAD process unfolds through distinct phases. During the Planning and Define Requirements phase, meticulous research and requirement gatherings take place to grasp the project's objectives, audience, scope, and establish measurable goals. A questionnaire sent to website clients ensures a comprehensive understanding of their needs. Prototyping follows, where the project team collaboratively creates an early working model of the website, incorporating user feedback through iterative cycles and casual interviews to refine features and eliminate design errors.

Rapid Construction and Feedback Gathering phase translates user feedback into practical development tasks, including coding and testing. Adjustments are made to application prototypes based on recommendations, and thorough testing ensures effective functionality. Once the website is launched, the Maintenance phase kicks in, focusing on regular upkeep to keep the site current, secure, and operational. This involves content updates, software and plugin maintenance, issue resolution, and adapting to evolving user input or business requirements.

3. Results and Discussion

The questionnaire results reveal key insights into consumer behavior and preferences for online sportswear shopping. Regardless of age, respondents engaged in sports are more inclined to shop for sportswear online, with those in their 20s and 30s showing the highest likelihood. Older age groups, especially those between 51 to 60 years old, show a reluctance to adopt e-commerce, possibly due to limited familiarity with technology.

Device usage for online shopping is evenly split between desktop and mobile, with respondents less likely to use tablets. JD Sports emerges as more popular than Nike and Adidas, attributed to its diverse range of brands and products. User-friendly navigation menus and simple, clean website designs are highly valued by respondents.

Product preferences highlight running shoes and activewear as top choices, driven not only by sports engagement but also as fashion items. Eco-friendly and sustainable sportswear gains significant preference, reflecting a growing awareness of environmental impact among consumers.

In terms of information, sizing guides, material details, and product reviews are considered crucial. The website's content, including blog articles, styling tips, and athlete endorsements, is appreciated by half of the respondents.

Checkout preferences are divided, with half opting for guest checkout and the other half preferring account creation before purchase. Touch ′ n Go is the favored payment method, emphasizing the importance of seamless transactions and a secure checkout process.

Email is the preferred mode of communication for customer support, chosen by half of the respondents over phone and live chat. Privacy and data security are paramount, with customers expecting companies to safeguard their information.

Overall, respondents express satisfaction with their online shopping experiences, emphasizing the importance of trust, reasonable pricing, and quality products for a successful e-commerce venture.

3.1 Implementation

This section focuses on a dynamic use case as shown in Fig. 1, emphasizing the synthesis of various diagrams to provide a comprehensive understanding of a specific system. Starting with a clever use case scenario analysis, An Entity-Relationship (ER) diagram (Fig. 2) visualizes the database structure. Concrete observations, such as screenshots of the admin and customer website sections, are integrated to bridge conceptual and practical aspects,

offering a holistic overview of the system's design and operation (Fig. 3).
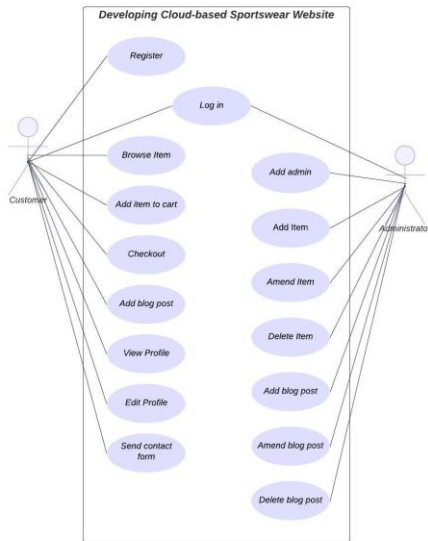

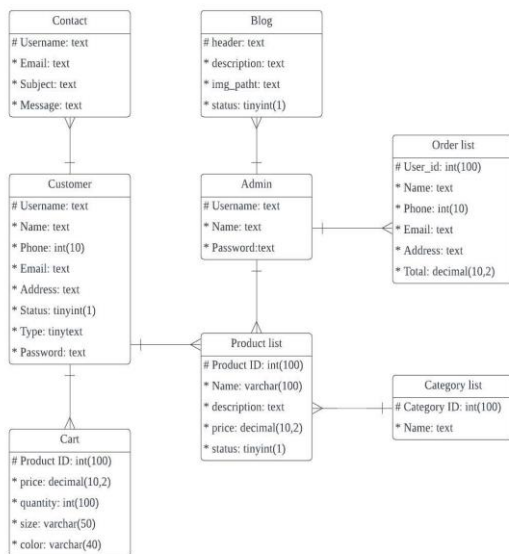Fig. 1 Use Case Diagram

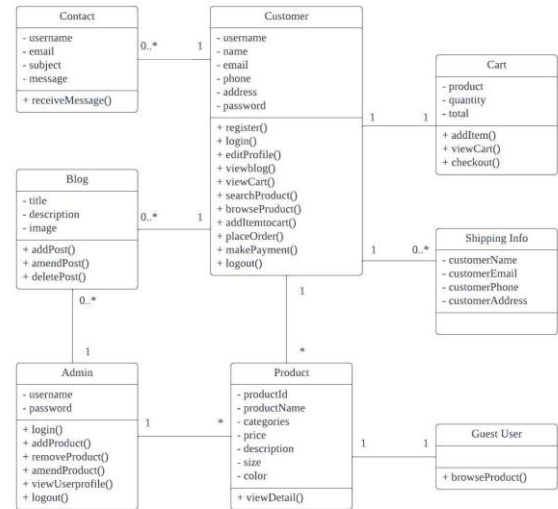Fig. 2 Entity Relationship Diagram

Fig. 3 Class Diagram

3.2 Evaluation

Evaluation begins with three crucial testing stages: unit testing, integration testing, and system testing. Together, these stages allow us to thoroughly assess the functionality and dependability of the proposed website.

Table 1, Table 2, Table 3 and Table 4 show the summary of the testing for each of the stages.

Table 1. Login Unit Test

Procedure Description	Expected Result	Actual Result	Test Result
Username text field	Allow user to key in their username.	Working as expected	Pass
Password text field	Allow user to key in their password.	Working as expected	Pass
Login button (Admin)	Allow admin to authenticate to access the website.	Working as expected	Pass
Login button (Customer)	Allow customer to authenticate to access the website.	Working as expected	Pass
Forget Password button	Allow user to reset their password.	Working as expected	Pass

Table 2. Integration Test

Procedure Description	Expected Result	Actual Result	Test Result
Login function should connect to database to validate the user.	The system able to connect to database and authenticate the user based on their username and password.	The system connects to database and verifies the user.	Pass
Customer place order function should connect to database to record customer order.	The system able to store customer order in the database.	The system connects to database and store customer order.	Pass
Admin side function should connect to database to update product detail.	The system able to store product detail in the database.	The system connects to database and store product detail.	Pass

Table 3. Admin side System Testing

Expected Result	Actual Result
Admin able to login to the system.	Admin can login successfully by insert username and password.
Admin able to view product on website.	All product is showed on the website.
Admin can add new product.	Product can be added by insert all the product detail.
Admin can amend product detail.	Product can be amended by insert all the amend detail.
Admin can delete product.	Product can be deleted by clicking the confirm delete button.
Admin can search user list.	User list is showed when search button is clicked.
Admin can add new admin.	New admin can be added by insert all the admin detail.
Admin can inactivate or activate user.	User can inactivate or activate when button is clicked.
Admin can logout from then system.	Admin can logout successfully when logout button is clicked.

Table 4. User Perception Test

Function	Yes/No
All the button are working perfectly	Yes
Able to login	Yes
All text field allow user to enter text	Yes
Dropdown menu allow user selection	Yes
Navigation menu allow user selection	Yes
All website functions are operational and working seamlessly	Yes
Error message when username or password incorrect	Yes
Alert successful message when action is performed	Yes
Able to logout	Yes

4. Conclusion

In summary, the construction and hosting of a cloud-based sportswear website on Amazon EC2 involved careful consideration of critical factors, particularly in response to the evolving landscape of e-commerce accelerated by the COVID-19 pandemic. The adoption of Amazon EC2, a reliable and scalable cloud computing solution, underscored the importance of leveraging technology for resilience and competitiveness in the e-commerce industry. The Rapid Application Development (RAD) methodology facilitated a collaborative and iterative development approach, aligning with the dynamic nature of e-commerce. Rigorous testing, including unit tests, integration tests, system tests, and user perception tests, ensured the website's reliability, functionality, and user satisfaction. This effective combination of digital technologies, cloud computing, RAD methodology, and comprehensive testing positions the sportswear website for success in the competitive e-commerce landscape, addressing challenges posed by the pandemic and establishing a robust foundation for sustained success in the dynamic online business environment.

References

1. Kysh L.M., "Adaptation Of B2c E-Commerce To The Conditions Of The Covid-19 Pandemic," East European Scientific Journal, vol. 12, no. 64, 2020.
2. S. Ramizulhasan, Dr. N. Wajid, and K. Azam, "Changing Dynamics of Online Trust and Retailers Ethics: A B2C study of high valued products in Pakistan from the COVID-19 perspective," International Journal of Business and Economic Affairs, vol. 6, no. 6, 2022. doi:10.24088/ijbea-2021-66003
3. Ahmed, I. (2019). A brief review: security issues in cloud computing and their solutions. Telkomnika,17(6), 2812-2817.
4. Hussainn, W., Hussainn, F.K., Hussainn, O.K., Damiani, E., & Chang, E. (2017). Formulating and managing viable SLAs in cloud computing from a small to medium service provider's viewpoint: A state-of-the-art review. Journal of Information System, 71(10), 240-259
5. Namburu, "Analysis Of Cloud Computing Services Using Amazon EC2," Course Hero, 2023.
6. S. Shokeen and A. Singh, Sci-Hub | Deploying an E-commerce website using Amazon Web Services ..., https://sci-hub.se/10.1109/IC3I46837.2019.9055586 (accessed Nov. 26, 2023).
7. What is Amazon EC2? - amazon elastic compute cloud, https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/concepts.html (accessed Nov. 27, 2023).

8. X Li, "Nike website analysis," Academia.edu, https://www.academia.edu/5711105/Nike_Website_Analysis (accessed Jun. 6, 2023).

Authors Introduction

Ms. Wei Yee Lim

She is a student majoring in B. Sc. (Hons) Business Information Systems at Institute of Computer Science and Digital Innovation (ICSDI), UCSI University, Kuala Lumpur, Malaysia.

Dr. Kasthuri Subaramaniam

She is currently an assistant professor at Institute of Computer Science and Digital Innovation (ICSDI), UCSI University, Kuala Lumpur, Malaysia. She earned both her bachelor's degree in computer science and a master's degree in computer science from the University of Malaya. Her research interests include human-computer interaction, e-commerce and etc.

Dr. Raenu Kolandaisamy

He is currently an assistant professor at Institute of Computer Science and Digital Innovation (ICSDI), UCSI University, Kuala Lumpur, Malaysia. He has a PhD in Computer Science. His research interests include Wireless Networking, VANET and IoTs.