ICAROB 2023

PROCEEDINGS OF THE 2023 INTERNATIONAL CONFERENCE ON ARTIFICIAL LIFE AND ROBOTICS

February 9 to 12, 2023
on line Oita in Japan
28th AROB International Meeting Series

Editor-in-Chief
Masanori Sugisaka
Editors: Yingmin Jia, Takao Ito, Ju-Jang Lee
ISBN 978-4-9908350-8-8

Proceedings of The 2023 International Conference on

ARTIFICIAL LIFE AND ROBOTICS

(ICAROB2023)

February 9 to 12, 2023, on line in Japan, 2023
28th AROB International Meeting Series

Editor-in-Chief
Masanori Sugisaka
Editors: Yingmin Jia, Takao Ito, Ju-Jang Lee
ISBN 978-4-9908350-8-8

Contents

1 Organization, etc. 1

2 Messages 11

3 Time Table 15

4 Opening Ceremony 19

5 Technical paper index 20

6 Abstracts

6-1 PS abstracts 46
6-2 OS abstracts 48
6-3 GS abstracts 102
6-4 POS abstract 112

7 Authors index 113

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)

ADVISORY COMMITTEES

Adam Grzech (Wroclaw University of Technology, Poland)
Bruce Eisenstein (Drexel University, Former IEEE President, USA)
Hidenori Kimura (RIKEN & Wasada University, Japan)
Jerzy Świątek (Wroclaw University of Technology, Poland)
Joshua M. Epstein (The Johns Hopkins University, USA)
Kai-Tai Song (National Chiao Tung University, Taiwan)
Ken-ichi Tanaka (Executive Fellow, Mitsubishi Electric Corporation, Japan)
Masato Nakagawa (Executive Fellow, DENSO CORPORATION, Visiting Professor of Hiroshima University, Japan)
Masayoshi Tomizuka (University of California Berkeley, USA)
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)
Steen Rasmussen (University of Southern Denmark, Denmark)
Zengqi Sun (Tsighua Univesity, P. R. China)

GENERAL CHAIR

Masanori Sugisaka
(Alife Robotics Corporation Ltd., Japan)
(Visiting Professor, The Open University (UK))

CO-GENERAL CHAIRS

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

VICE GENERAL CHAIRS

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)

©ICAROB 2023 ALife Robotics Corp. Ltd.

Katia Passerini (Seton Hall University, USA)
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 COMMITTEES

Akira Fukuda (Kyushu University, Japan)
Akira Nakamura (Saitama Institute of Technology, Japan)
Akinori Abe (Chiba University, Japan)
Andrew Gadsden (University of Guelph, Canada)
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)
Hideyuki Suzuki (The University of Tokyo, 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)
Kevin Voges (Canterbury University, New Zealand)
Kunikazu Kobayashi (Aichi Prefectural University, Japan)
Kuo-Hsien Hsia (Far East University, Taiwan)
Kyungho Park (U.S. ARMY, Japan and USA)
Masao Kubo (National Defense Academy of Japan, Japan)
Masanao Obayashi (Yamaguchi University, Japan)
Maxim Sokolov (Innopolis University, Russia)

©ICAROB 2023 ALife Robotics Corp. Ltd.

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)
Shin Jisun (Pusan National University, Korea)
Singo Mabu (Yamaguchi University, Japan)
Takashi Kohno (The University of Tokyo, Japan)
Takashi Ogata (Iwate Prefectural University)
Teruhisa Hochin (Kyoto Prefectural University, Japan)
Tetsuro Hattori (Kagawa University, Japan)
Thi Thi Zin (University of Miyazaki, Japan)
Thomas S. Ray (University of Oklahoma, USA)
Toru Yamamoto (Hiroshima University, Japan)
Toshihro Inukai (DENSO WAVE INCORPORATED, 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)
Zengqiang Chen (Nankai University, P. R. China)

INTERNATIONAL PROGRAM COMMITTEES

Abdul Rahman bin Dullah (Universiti Teknikal Malaysia Melaka, Malaysia)
Akinori Abe (Chiba University, Japan)
Akira Utsumi (The University of Electro-Communications, 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 (Kyushu Institute 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)

©ICAROB 2023 ALife Robotics Corp. Ltd.

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)
Dongmei Ai (Universty of Science and Technology Beijing, 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)
Junping Du (Beijing University of Posts and Telecommunications, 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)
Junping Du (Beijing University of Posts and Telecommunications, P. R. China)
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)

©ICAROB 2023 ALife Robotics Corp. Ltd.

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)
Noriaki Kuwahara (Kyoto Prefectural University, Japan)
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)

©ICAROB 2023 ALife Robotics Corp. Ltd.

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 (Yamaguchi University, Japan)
Takayoshi Yamada (Gifu University, Japan)
Takuya Fuginaga (Kyushu Institute of Technology, Japan)
Takuya Nanami (The University of Tokyo. Japan)
Taishiro Kishimoto (Keio University, Japan)
Taisuke Akimoto (Kyushu Institute of Technology, Japan)
Takashi Iwamoto (Mitsubishi Electric Corporation, Japan)
Tarik bin Abd Latef (Universiti Malaya, Malaysia)
Taro Shibanoki (Ibaraki University, Japan)
Teruhisa Hochin (Kyoto Prefectural 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 (Tokyo Institute of Technology, Japan)

©ICAROB 2023 ALife Robotics Corp. Ltd.

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 COMMITTEES

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)(vol.1-8) & Journal of Robotics,

©ICAROB 2023 ALife Robotics Corp. Ltd.

Networking and Artificial Life(vol.9-) & 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

©ICAROB 2023 ALife Robotics Corp. Ltd.

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

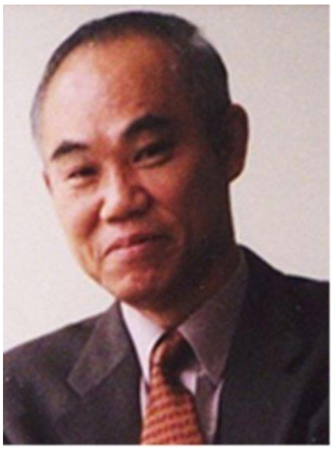
COPYRIGHTS

Accepted papers will be published in the proceeding of The 2022 International Conference on Artificial Life and Robotics (ICAROB2022) 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 (JRNAL)(vol.1-8) & Journal of Robotics, Networking and Artificial Life(vol.9-) & 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.

ICAROB Office
ALife Robotics Corporation Ltd.
4-Go, 8-Ban, Higshi 2Cyome, Handadai, Oita 870-1108, JAPAN
TEL/FAX：+81-97-597-7760
E-MAIL：icarob@alife-robotics.co.jp Home Page：https://alife-robotics.co.jp/

©ICAROB 2023 ALife Robotics Corp. Ltd.

MESSAGES



Masanori Sugisaka
General Chair
(President, ALife Robotics Co., Ltd., Japan)
(Visiting Professor, The Open University, UK)

Masanori Sugisaka

Masanori Sugisaka

General Chair of ICAROB

It is my great honor to invite you all to The 2023 International Conference on Artificial Life and Robotics (ICAROB 2023). 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 in SCOPUS and CPCI-Web of Science now. From now on, ALife Robotics Corporation Ltd. is in charge of management. This year we have The 2023 International Conference on Artificial Life and Robotics (ICAROB2023) (28th 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 (ICAROB2023) 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.

©ICAROB 2023 ALife Robotics Corp. Ltd.

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

Yingmin Jia

Co-General Chair of ICAROB

It is my great pleasure to invite you to The 2023 International Conference on Artificial Life and Robotics (ICAROB 2023), Oita, Japan, February 9-12, 2023. Because of the influence of COVID-19, The ICAROB2023 will be held on-line again, and 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 28th birthday in 2023. 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 ICAROB2023.

©ICAROB 2023 ALife Robotics Corp. Ltd.

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

Takao Ito

Takao Ito

Co-General Chair of ICAROB

It is my great honor and pleasure to invite you all to the 2023 International Conference on Artificial Life and Robotics (ICAROB 2023) which will be held online from February 9 to 12, 2023.

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.

This conference is online. I eagerly look forward to personally meeting you in online, during the ICAROB 2023 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.

©ICAROB 2023 ALife Robotics Corp. Ltd.

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

JJLee

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 10th ICAROB2023 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 10th ICAROB.

I'm looking forward to meeting you at the 10th ICAROB in on line and wishing you all the best.

©ICAROB 2023 ALife Robotics Corp. Ltd.

GENERAL SESSION TOPICS

GS1 Machine Learning & Neural Network & Artificial Life (3)	GS2 Image Processing I (5)
GS3 Image Processing II (4)	GS4 Robotics I (4)
GS5 Robotics II (4)	GS6 Applications I (5)
GS7 Applications II (3)	Poster (1)

ORGANIZED SESSION TOPICS

OS1 Human-Machine Interface Application (7)	OS2 Intelligent Life and Cybersecurity (6)
OS3 Advanced Information Processing Applications (4)	OS4 Artificial Life and Intelligent Systems (4)
OS5 Intelligent Systems and Control (5)	OS6 Modelling and Simulation in Robotics (8)
OS7 Advanced Control Systems (7)	OS8 Intelligent Control (5)
OS9 Intelligent Systems and Robotics - 1 (4)	OS10 Intelligent Systems and Robotics - 2 (4)
OS11 Intelligent Systems and Robotics - 3 (4)	OS12 Theory and Implementation of Neuromimetic Systems (2)
OS13 Industrial Artificial Intelligence Robotics (9)	OS14 Software Development Support Method (4)
OS15 Robot Control (5)	OS16 Robotic Manipulation (3)
OS17 Artificial Intelligence for Embedded Systems and Robotics (6)	OS18 Mathematical Informatics (7)
OS19 Natural Computing (2)	OS20 Robot Competitions and Education(7)
OS21 Advances in Field Robotics and Their Applications (7)	OS22 Bio-inspired Artificial Vision -Algorithms and Systems (5)
OS23 STEM Education for fostering Innovators (4)	OS24 Applications of Unmanned Aerial System (3)
OS25 Emerging Technologies in Industrial Revolution 4.0 (IR 4.0) Era (8)	OS26 Signals Processing and Automation I (6)
OS27 Research Towards the Sustainable Development Goals (SDG's) (6)	OS28 Is That Narratology 'Post'? -Computational and Cognitive Approaches and Beyond (5)
OS29 Signals Processing and Automation II (6)	OS30 Embedded systems projects (3)
OS31 Image Processing and Reconstruction (5)	OS32 AR, VR and Digital Contents (4)

©ICAROB 2023 ALife Robotics Corp. Ltd.

TIME TABLE (2/9)

in Japan time

2/9(Thu.) 17:30-18:30	Group meeting for the conference (Conference Room ZOOM ID:894 3889 7319)
2/12(Sun) 15:45-16:45	Group meeting for the next conference (Conference Room ZOOM ID:894 3889 7319)

TIME TABLE (2/10)

2/10(Fri.)	**Meeting Room 1** (ZOOM ID:853 5059 5830)	**Meeting Room 2** (ZOOM ID:878 7748 3158)	**Meeting Room 3** (ZOOM ID:821 6117 1740)	**Meeting Room 4** (ZOOM ID:897 3935 1940)
8:40-	Registration			
9:00-10:00	OS4 Artificial Life and Intelligent Systems (4) Chair: Chung-Wen Hung	OS12 Theory and Implementation of Neuromimetic Systems (2) Chair: Takashi Kohno	OS14 Software Development Support Method (4) Chair: Tetsuro Katayama	POS (1) Chair: Yong Hun Kim Poster PDF
10:00-10:20	Coffee break			
10:20-10:50	Chair: Makoto Sakamoto (Conference Room ZOOM ID:894 3889 7319) Opening Ceremony			
11:00-12:00	Chair: Norrima Mokhtar (Conference Room ZOOM ID:894 3889 7319) Plenary Speech PS1 Kaoru Sumi			
12:00-13:00	Lunch			
13:00-14:30	OS2 Intelligent Life and Cybersecurity (6) Chair: I-Hsien Liu	GS 1 Machine Learning & Neural Network & Artificial Life (3) Chair: Masato Nagayoshi	OS17 Artificial Intelligence for Embedded Systems and Robotics (6) Chair: Hakaru Tamukoh	OS22 Bio-inspired Artificial Vision -Algorithms and Systems- (5) Chair: Shinsuke Yasukawa
14:30-14:50	Coffee break			
14:50-16:35	OS5 Intelligent Systems and Control (5) Chair: Kuo-Hsien Hsia	OS7 Advanced Control Systems (7) Chair: Takuya Kinoshita	OS 20 Robot Competitions and Education (7) Chair: Kazuo Ishii	OS29 Signals Processing and Automation II (6) Chair: Norrima Mokhtar
16:35-16:50	Coffee break			
16:50-17:50	Chair: Kuo-Hsien Hsia (Conference Room ZOOM ID:894 3889 7319) Plenary Speech PS2 Evgeni Magid			

©ICAROB 2023 ALife Robotics Corp. Ltd.

TIME TEBLE (2/11)

2/11(Sat.)	**Meeting Room 1** (ZOOM ID:853 5059 5830)	**Meeting Room 2** (ZOOM ID:878 7748 3158)	**Meeting Room 3** (ZOOM ID:821 6117 1740)	**Meeting Room 4** (ZOOM ID:897 3935 1940)
8:40-	Registration			
9:00-10:30	OS9 & OS10 & OS11 Intelligent Systems and Robotics 1,2,3 (12)->6 implementations Chair: Hucheng Wang, Fangyan Li, Jiaxin Li	OS27 Research Towards the Sustainable Development Goals (SDG's) (6) Chair: Ammar A.M. Al Talib	OS26 Signals Processing and Automation I (6) Chair: Norrima Mokhtar	OS15 Robot Control (5) Chair: Yizhun Peng
10:30-10:45	Coffee break			
10:45-12:00	OS3 Advanced Information Processing Applications (4) Chair: Toru Hiraoka	OS8 Intelligent Control (5) Chair: Yingmin Jia	OS23 STEM Education for fostering Innovators (4) Chair: Hiroyuki Y. Suzuki	GS3 Image Processing II (4) Chair: Joo Kooi Tan
12:00-13:00	Lunch			
13:00-14:00	Chair: Makoto Sakamoto (Conference Room ZOOM ID:894 3889 7319) Plenary Speech PS3 Xiaoyan Chen			
14:00-14:20	Coffee break			
14:20-15:20	OS25-1 Emerging Technologies in Industrial Revolution 4.0 (IR 4.0) Era (4) Chair: Wei Hong Lim	OS13-1 Industrial Artificial Intelligence Robotics (4) Chair: Eiji Hayashi	OS16 Robotic Manipulation (3) Chair: Kensuke Harada	OS21-1 Advances in Field Robotics and Their Applications (4) chair: Keisuke Watanabe
15:20-15:40	Coffee break			
15:40-16:55	OS25-2 Emerging Technologies in Industrial Revolution 4.0 (IR 4.0) Era (4) Chait: Wei Hong Lim	OS13-2 Industrial Artificial Intelligence Robotics (5) Chair: Eiji Hayashi	OS6-1 Modelling and Simulation in Robotics (5) Chair: Evgeni Magid	OS21-2 Advances in Field Robotics and Their Applications (3) chair: Keisuke Watanabe
16:55-17:10	Coffee break			
17:10-17:55	OS19 Natural Computing (2) Chair: Marion Oswald	GS7 Applications II (3) Chair: Hiroaki Wagatsuma	OS6-2 Modelling and Simulation in Robotics (3) Chair: Evgeni Magid	OS30 Embedded systems projects (3) Chair: Mastaneh Mokayef

©ICAROB 2023 ALife Robotics Corp. Ltd.

TIME TABLE (2/12)

2/12(Sun.)	**Meeting Room 1** (ZOOM ID:853 5059 5830)	**Meeting Room 2** (ZOOM ID:878 7748 3158)	**Meeting Room 3** (ZOOM ID:821 6117 1740)	**Meeting Room 4** (ZOOM ID:897 3935 1940)
8:40-	Registration			
9:00-10:00	OS24 Applications of Unmanned Aerial System (3) Chair: Hazry Desa	OS18-1 Mathematical Informatics (4) Chair: Takao Ito	OS32 AR, VR and Digital Contents (4) Chair: R.P.C. Janaka Rajapakse	GS4 Robotics I (4) Chair: Amornphun Phunopas
10:00-10:30	Coffee break			
10:30-11:30	Chair: Takao Ito (Conference Room ZOOM ID:894 3889 7319) Plenary Speech PS4 Masato Nakagawa			
11:30-12:30	Lunch			
12:30-14:15	OS1 Human-Machine Interface Application (7) Chair: Norrima Mokhtar	OS18-2 Mathematical Informatics (3) Chair: Takao Ito	OS28 Is That Narratology 'Post'? - Computational and Cognitive Approaches and Beyond (5) Chair: Jumpei Ono	GS5 Robotics II (4) Chair: Takayoshi Yamada
14:15-14:30	Coffee break			
14:30-15:45		OS31 Image Processing and Reconstruction (5) Chair: Xiaoyan Chen	GS2 Image Processing I (5) Chair: Watcharin Tangsuksant	GS 6 Applications I (5) Chair: Minoru Kumano
15:45-16:45	Group meeting for the next conference (Conference Room ZOOM ID:894 3889 7319)			

©ICAROB 2023 ALife Robotics Corp. Ltd.

The 2023 International Conference on
ARTIFICIAL LIFE AND ROBOTICS
(ICAROB2023)

February 9 (Thursday)

17:30-18:30

Group meeting for the conference

February 10 (Friday)

10:20-10:50

Opening Ceremony

Chair: Makoto Sakamoto (University of Miyazaki, Japan)

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)
4-1. Vice-General Chairman of ICAROB	Katia Passerini (Seton Hall University, USA)
4-2. Vice-General Chairman of ICAROB	Henrik Hautop Lund (Technical University of Denmark, Denmark)
4-3. Vice-General Chairman of ICAROB	Norrima Mokhtar (University of Malaya, Malaysia)

February 12 (Sunday)

15:45-16:45

Group meeting for the next conference

©ICAROB 2023 ALife Robotics Corp. Ltd.

TECHNICAL PAPER INDEX

February 10 (Friday)

8:40-Registration

Conference Room
10:20-10:50 Opening Ceremony
Chair: Makoto Sakamoto (University of Miyazaki, Japan)

11:00-12:00
Plenary Speech PS1
Chair: Norrima Mokhtar (University of Malaya, Malaysia)

PS1 *Human to Human Interaction using Virtual Agents Posing as Another Person*
Kaoru Sumi (Future University Hakodate, Japan)

16:50-17:50
Plenary Speech PS2
Chair: Kuo-Hsien Hsia (National Yunlin University of Science and Technology, Taiwan)

PS2 *Simulation Tools for Urban Search and Rescue Robotics*
Evgeni Magid (Kazan Federal University, Russia)

Meeting Room 1
9:00-10:00 OS4 Artificial Life and Intelligent Systems (4)
Chair: Chung-Wen Hung (National Yunlin University of Science and Technology, Taiwan)
Co-Chair: Kuo-Hsien Hsia (National Yunlin University of Science and Technology, Taiwan)

OS4-1 *Interactive Beating Drum Unity Game*
Chung-Wen Hung, Cheng-Lung Ko, Wen-Huei Chou
(National Yunlin University of Science and Technology, Taiwan)

OS4-2 *Error Backpropagation Neural Network Based Image Identification for a Foot Massage Machine and Its Mechanism Design*
Chun-Chieh Wang (National Yunlin University of Science and Technology, Taiwan)

OS4-3 *Cross-domain sharing of robots in the community caring and practice of university social responsibility*
Jia-Ming Hsiao (National Yunlin University of Science and Technology, Taiwan)

©ICAROB 2023 ALife Robotics Corp. Ltd.

OS4-4 *Research on Design of Implementation Mechanism for Similar Production Line*
Chia-Nan Ko, Yi-Yu Li, Ting-Ru Ko, and Ting-Yi Chen
(Nan Kai University of Technology, Taiwan)

13:00-14:30 OS2 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: Chuan-Kang Liu (National Chin-Yi University of Technology)

OS2-1 *The Dam Gate Cybersecurity Testbed*
Chen-Yu Lee, I-Hsien Liu, Meng-Wei Chang, Jung-Shian Li
(National Cheng Kung University, Taiwan)

OS2-2 *Domain Name Infringement in Taiwan*
Shih-Chin Lin (Ming Chuan University, Taiwan)

OS2-3 *Device's Operation Tracking using Blockchain in Industrial Control System*
Chien-Hsin Wu, I-Hsien Liu, Jung-Shian Li (National Cheng Kung University, Taiwan)
Chu-Fen Li (National Formosa University, Taiwan)

OS2-4 *Strengthen the security of the Industrial Control System using SDN technology*
Min-Wei Huang, I-Hsien Liu, Hsin-Yu Lai, Meng-Huan Lee, Jung-Shian Li
(National Cheng Kung University, Taiwan)

OS2-5 *Fake Base Stations threats in 5G Standalone Networks*
Meng-Huan Lee, I-Hsien Liu, Jung-Shian Li (National Cheng Kung University, Taiwan)

OS2-6 *Cyber-Physical Security Testbed for River Basin Gate Control System*
Meng-Wei Chang, I-Hsien Liu, Jung-Shian Li
(National Cheng Kung University, Taiwan)

14:50-16:05 OS5 Intelligent Systems and Control (6)
Chair: Kuo-Hsien Hsia (National Yunlin University of Science and Technology, Taiwan)
Co-Chair: Chian C. Ho (National Yunlin University of Science and Technology, Taiwan)

OS5-1 *Using Multithreaded Load Balancer to Improve Connection Performance in Container Environment*
Pang-Wei Tsai, Hong-Yu Wei, Yu-Chi Hsu (National Central University, Taiwan)

OS5-2 *Electronic Biometric Detector and Body Composition Index in Predicting Disease Risk*
Wen-Fu Yang, Chung-Te Ting (Chang Jung Christian University, Taiwan)

©ICAROB 2023 ALife Robotics Corp. Ltd.

OS5-3 *Exploring Consumers' Intention to Use Mobile Payment APPs Based on Technology Acceptance Models - Taking Line Pay as an Example*
Chun -Nan Chen, Yu-Sheng Huang (Chang Jung Christian University, Taiwan)

OS5-4 *Optimization of Robot Path and IoT Communication Path Based on Artificial Intelligence*
Jr-Hung Guo, Kuo-Hsien Hsia
(National Yunlin University of Science and Technology, Taiwan)

OS5-5 *Key Success Factors Affecting Family Members' Intention to Withdraw from Life-sustaining Treatment for Long-term Ventilator-dependent patients: Nursing Professionals' Perspective*
Hsiao-Fang Chen, Jin-Yuan Chern
(Chang Jung Christian University, Taiwan)

Meeting Room 2
9:00-9:30 OS12 Theory and Implementation of Neuromimetic Systems (2)
Chair Takashi Kohno (University of Tokyo, Japan)
Co-Chair Takuya Nanami (University of Tokyo, Japan)

OS12-1 *Adaptive STDP Learning with Lateral Inhibition for Neuromorphic Systems*
Ashish Gautam, Takashi Kohno (University of Tokyo)

OS12-2 *Spike pattern detection with close-to-biology spiking neuronal network*
Takuya Nanami, Takashi Kohno (University of Tokyo)

13:00-13:45 GS1 Machine Learning & Neural Network & Artificial Life (3)
Chair: Masato Nagayoshi (Niigata College of Nursing, Japan)

GS1-1 *Defect Solder Classification in Print Circuit Boards using Machine Learning*
Watcharin Tungsuksan, Hattapat Silangren, Phattharaporn Iamcharoen, Wisanu Jitviriya
(King Mongkut's University of Technology North Bangkok, Thailand)

GS1-2 *A dynamic nurse scheduling using reinforcement learning: Dealing with various sudden absences of a nurse*
Masato Nagayoshi (Niigata College of Nursing, Japan)
Hisashi Tamaki (Kobe University, Japan)

GS1-3 *Classification of Time Series Data Obtained by the Satellite by Using Rule-Based and Machine-Learning Methods*
Satoko Saita[1], Mariko Teramoto[2], Kentarou Kitamura[2]
([1]National Institute of Technology, Kitakyushu College, Japan)
([2]Kyushu Institute of Technology, Japan)

©ICAROB 2023 ALife Robotics Corp. Ltd.

14:50-16:35 OS7 Advanced Control Systems (7)
Chair: Takuya Kinoshita (Hiroshima University, Japan)
Co-Chair: Takao Sato (University of Hyogo, Japan)

OS7-1 *Predictive Functional Controller Design with Disturbance Observer and Its Application*
Syota Yoshida, Shin Wakitani (Hiroshima University)

OS7-2 *Study on a Construction of Velocity Perception Model and Kansei Feedback Control System in Active Behavior*
Sota Takemura, Takuya Kinioshita, Toru Yamamoto (Hiroshima University, Japan)

OS7-3 *Design of a Database-Driven Control for a Web Conveyor*
Atsushi Takatani, Takuya Kinioshita, Toru Yamamoto (Hiroshima University, Japan),
Tomohiro Hirakawa, Hiroki Hamamoto, Takashi Ochiiwa, Hideki Tomiyama
(Japan Steel Works, Ltd.)

OS7-4 *Development of IoT self-tuning control device using Wi-Fi*
Shinichi Imai (Tokyo Gakugei University, Japan)

OS7-5 *Consensus Control for Dual-rate Multi-agent Systems*
Takaya Tanaka, Natsuki Kawaguchi, Takao Sato (University of Hyogo, Japan)

OS7-6 *Data-driven Control Experiments of a Quadrotor Drone*
Tomonori Yao, Natsuki Kawaguchi, Takao Sato (University of Hyogo, Japan)

OS7-7 *Design of Data-driven Multi-agent Systems*
Kenta Nagao, Natsuki Kawaguchi, Takao Sato (University of Hyogo, Japan)

Meeting Room 3
9:00-10:00 OS14 Software Development Support Method (4)
Chair: Tetsuro Katayama (University of Miyazaki, Japan)
Co-Chair: Tomohiko Takagi (Kagawa University, Japan)

OS14-1 *Automated Random Simulation for Checking a Behavioral Model of Systems Based on Extended Place/Transition Net with Attributed Tokens*
Sho Matsumoto[1], Tetsuro Katayama[2], Tomohiko Takagi[1]
([1] Kagawa University, Japan), ([2] University of Miyazaki, Japan)

OS14-2 *Training of Software Formal Modeling Using Visual Blocks for Actions and Guards of Extended Place/Transition Net*
Akio Usuda, Ryoichi Ishigami, Tomohiko Takagi (Kagawa University, Japan)

©ICAROB 2023 ALife Robotics Corp. Ltd.

OS14-3 *Proposal of a Framework to Improve the Efficiency of the Implementation Step in Test Driven Development (TDD)*
Takeaki Miyashita[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)

OS14-4 *Continuance Intention Factor of Online Learning Management System in Case on Faculty of Computer Science at Brawijaya University in Indonesia*
Mochamad Chandra Saputra[1], Tetsuro Katayama[2], Yoshihiro Kita[3], Hisaaki Yamaba[2], Kentaro Aburada[2], Naonobu Okazaki[2]
([1]Brawijaya University, Indonesia), ([2]University of Miyazaki, Japan), ([3]University of Nagasaki, Japan)

13:00-14:30 OS17 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)

OS17-1 *Pose Detection for Flexible-Indefinite Objects using Pseudo-Bone Data*
Yuma Yoshimoto[1], Hakaru Tamukoh[2]
([1]National Institute of Technology, Kitakyushu College, Japan), ([2]Kyushu Institute of Technology, Japan)

OS17-2 *Object Search and Empty Space Detection System for Home Service Robot*
Tomoya Shiba[1], Tomohiro Ono[1,2], Hakaru Tamukoh[1]
([1]Kyushu Institute of Technology, Japan), ([2]JSPS Research Fellow, Japan)

OS17-3 *Robust Classification Model with Multimodal Learning for Home Service Robots*
Ikuya Matsumoto, Daiju Kanaoka, Hakaru Tamukoh
(Kyushu Institute of Technology, Japan)

OS17-4 *Flexible Human-Robot Interaction in Domestic Environment Using Semantic Map*
Yuga Yano[1], Yukiya Fukuda[1], Tomohiro Ono[1,2], Hakaru Tamukoh[1]
([1]Kyushu Institute of Technology, Japan), ([2]JSPS Research Fellow, Japan)

OS17-5 *Impact of PS Load on FPGA Object Detection System Performance*
Yusuke Watanabe[1,2], Hakaru Tamukoh[2]
([1]CRAFTWORK Co. Ltd, Japan,), ([2]Kyushu Institute of Technology, Japan)

OS17-6 *An Effective Method for Minimizing Domain Gap in Sim2Real Object Recognition Using Domain Randomization*
Tomohiro Ono[1,2], Akihiro Suzuki[1], Hakaru Tamukoh[1]
([1]Kyushu Institute of Technology, Japan), ([2]JSPS Research Fellow, Japan)

©ICAROB 2023 ALife Robotics Corp. Ltd.

14:50-16:35 OS20 Robot Competitions and Education (7)
Chair: Kazuo Ishii (Kyushu Institute of Technology, Japan)
Co-Chair: Yasunori Takemura (Nishinippon Institute of Technology, Japan)

OS20-1 *Report on Underwater-Robot-Festival Junior Division Aiming at Marine Debris Clean up*
[1]Takayuki Matsuo, [2]Masanori Sato, [3]Masayoshi Ozawa, [4]Seiichiro Miura, [5]Masakazu Arima, [6]Kazuo Ishii
([1]National Institute of Technology, Kitakyushu College, Japan),
([2]Nagasaki Institute of Applied Science, Japan),
([3]Kobe City College of Technology, Japan),
([4]National Institute of Technology, Tokuyama College, Japan)
([5]Osaka Metropolitan University), ([6]Kyushu Institute of Technology)

OS20-2 *Image-based navigation of Small-size Autonomous Underwater Vehicle "Kyubic" in International Underwater Robot Competition*
Yusuke Mizoguchi, Daiki Hamada, Riku Fukuda, Irmiya R. Inniyaka, Kaito Kuwata, Keisuke Nishimuta, Akihiro Sugino, Rikuto Tanaka, Yoshiki, Tanaka, Yuya Nishida, Kazuo Ishii
(Kyushu Institute of Technology, Japan)

OS20-3 *Development of Harvesting Robot for Tomato Robot Competition 2022 and Its Evaluation*
Takeru Oshige[1], Moeko Tominaga[2], Takuya Fujinaga[3], Yasunori Takemura[2], Jonghyun Ahn[1]
([1]Hiroshima Institute of Technology, [2]Nishinippon Institute of Technology, [3]Fukuoka University, Japan)

OS20-4 *Development of a Tomato Harvesting Robot for Farm Field*
Shunsuke Oda, Ryuma Fukumoto, Kensuke Hirata, Shu Tahara, Keisuke Yoshida, Shinsuke Yasukawa, Kazuo Ishii (Kyushu Institute of Technology, Japan)

OS20-5 *Tomato-Harvesting Robot Competition: Developed Robots and Results of 9th Competition*
[1]Kazuo Ishii, [2]Takayuki Matsuo, [3]Yasunori Takemura, [3]Takashi Sonoda, [3]Atsushi Sanada, [1]Yuya Nishida, [1]Shinsuke Yasukawa, [4]Takuya Fujinaga, [3]Moeko Tominaga, [5]Daisaku Arita, [6]Kazushi Kawajiri, [7]Kenich Ohshima, [7]Masayuki Okada, [1]Kanako Shirahashi
([1]Kyushu Institute of Technology, Japan), ([2]National Institute of Technology, Kitakyushu College, Japan), ([3]Nishinippon Institute of Technology, Japan), ([4]Fukuoka University, Japan), ([5]University of Nagasaki, Japan), ([6]Hibikinada Greenhouse, Japan), ([7]Kyushu Polytechnic College, Japan)

OS20-6 *A Modeling of Sphere Considering Slipping Adapted Three-Rollers*
[1]Kenji Kimura, [2]Kouki Ogata,[3]Hiroyasu Hirai, [3]Takumi Ueda, [3]Kazuo Ishi
([1]National Institute of Technology, Matsue College, Japan), ([2]Saga University, Japan)
([3]Kyushu Institute of Technology, Japan)

©ICAROB 2023 ALife Robotics Corp. Ltd.

OS20-7 *Roller Arrangement Problem of Omnidirectional Mobil Robot Adapted Three Omni Rollers*
[1]Kenji Kimura, [2]Yuki Shigyo,[3]Kazuo Ishi
([1]National Institute of Technology, Matsue College, Japan), ([2]Fujitsu, Japan)
([3]Kyushu Institute of Technology, Japan)

Meeting Room 4
9:00-10:00 Poster1 (1)
Chair: Yong Hun Kim (Sejong University, Republic of Korea)

POS *Magnetic Anomaly-Matched Trajectory and Dead Reckoning Fusion Mobile Robot Navigation*
Yong Hun Kim, Bo Sung Ko, Jin Woo Song (Sejong University, Republic of Korea)

13:00-14:15 OS22 Bio-inspired Artificial Vision -Algorithms and Systems-(5)
Chair: Shinsuke Yasukawa (Kyushu Institute of Technology, Japan)
Co-Chair: Yuki Hayashida (Mie University, Japan)

OS22-1 *Increasing selectivity to a feature combination using inhibitory synaptic plasticity in a spiking neural network*
Mahiro Ikeda, Hirotsugu Okuno (Osaka Institute of Technology, Japan)

OS22-2 *A binocular disparity estimation algorithm using multiple spatial frequency information and a neural network*
Ryoka Sato, Hirotsugu Okuno (Osaka Institute of Technology, Japan)

OS22-3 *A figure-ground discrimination algorithm inspired by border-ownership selective cells*
Tomoya Kobayashi, Hirotsugu Okuno (Osaka Institute of Technology, Japan)

OS22-4 *Event-Driven Particle Filter for Tracking Irregularly Moving Objects*
Yuki Kawasaki[1], Masahiro Ohtani[2], Shinsuke Yasukawa[1]
([1]Kyushu Institute of Technology, [2]Mie University, Japan)

OS22-5 *A simulation model for analyzing the spatiotemporal receptive field of retinal ganglion cells in the presence of fixational eye movements*
Hiroyuki Yokota[1], Yuki Hayashida[2], Shinsuke Yasukawa[1]
([1]Kyushu Institute of Technology, [2]Mie University, Japan)

©ICAROB 2023 ALife Robotics Corp. Ltd.

14:50-16:20 OS29 Signals Processing and Automation II (6)

Chair: Norrima Mokhtar (University of Malaya, Malaysia)
Co-Chair: Heshalini Rajagopal (UCSI University, Malaysia)

OS29-1 Weather Forecast System for Mobile Devices
Hu Jiahao, Neesha Jothi, Chloe Thong (UCSI University, Malaysia)

OS29-2 *Development of a Novel E-Learning System for Improved Usability*
Soo Yang Yew, Shaik Shabana Anjum, Shayla Islam (UCSI University, Malaysia)

OS29-3 *A Development of a Prototype-based Mobile Pet Care Application*
Gan Ai Leen, Shaik Shabana Anjum, Chloe Thong Chee Ling
(UCSI University, Malaysia)

OS29-4 *Modeling of an Environmentally Independent and Contactless Speed Sensor for Measuring the Speed of Ships, Submarines, and Aircraft in Relation to the Ground Development of Image*
Jakaria Mahdi Imam, Mohammad Aminul Islam, Norrima Binti Mokhtar,
S. F. W. Muhamad Hatta (Universiti Malaya, Malaysia)

OS29-5 On Correcting Luminosity and Contrast of Retinal Images with Reflectance
Mofleh Hannuf AlRowaily[1], Hamzah Arof[1], Imanurfatiehah Ibrahim[1], Wan Amirul Wan Mohd Mahiyiddin[1], Norrima Mokhtar[1], Haniza Yazid[2]
([1]Universiti Malaya, Malaysia), ([2]Universiti Malaysia Perlis, Malaysia)

OS29-6 *Rate Adaptation for Quality of Service Improvements in IEEE 802.11ax*
Hazwani Zawawi, Wan Norsyafizan W Muhamad, Nani Fadzlina Naim
(Universiti Teknologi, MARA)

February 11 (Saturday)

8:40-Registration

Conference Room

13:00-14:00

Plenary Speech PS3

Chair: Makoto Sakamoto (University of Miyazaki, Japan)

PS3 *Enhancement methodology for low light image*

Xiaoyan Chen (Tianjin University of Science and Technology, China)

©ICAROB 2023 ALife Robotics Corp. Ltd.

Meeting Room 1
9:00-10:30 OS9 Intelligent Systems and Robotics – 1 (4) & OS10 Intelligent Systems and Robotics – 2 (4) & OS11 Intelligent Systems and Robotics – 3 (4)

OS9 Intelligent Systems and Robotics – 1 (4)
Chair: Hucheng Wang (Tianjin University of Science and Technology, China)
Co-Chair: Fengzhi Dai (Tianjin University of Science and Technology, China)

OS9-1 *A Survey of Target Detection Based on Deep Learning*
Hucheng Wang, Fengzhi Dai, Min Zhao
(Tianjin University of Science and Technology, China)

OS9-2 *A Design of New Air Ground Cooperative Unmanned Transportation System*
Hucheng Wang, Min Zhao (Tianjin University of Science and Technology, China)

OS9-3 Attitude Solution of Quadrotor UAV
Siyuan Liu, Zhihao Zhao, Haoran Gong
(Tianjin University of Science and Technology, China)

OS9-4 *A four-dimensional conservative chaotic system and its application in image encryption*
Hongyan Jia, Wei Li, Jingwen Liu
(Tianjin University of Science and Technology, Tianjin, China)

OS10 Intelligent Systems and Robotics – 2 (4)
Chair: Fangyan Li (Tianjin University of Science and Technology, China)
Co-Chair: Yande Xiang (Tianjin University of Science and Technology, China)

OS10-1 *A Research on Image Defogging Algorithm Based on Enhancement*
Fangyan Li, Haokang Wen, Chang Sheng, Min Zhao
(Tianjin University of Science and Technology, China)

OS10-2 *Autonomous Microcontroller-Based Aerial Water Sampling Device*
Weifang Wang[1], Mingxia Kang[2], Ruming Kang[3]
([1]North Minzu University, China), ([2]Tianjin University of Commerce, China), ([3]Chongqing University, China)

OS10-3 *Intelligent Electronic Guide Dog*
Mengyu Liu, Yande Xiang, Zhi Qiao, Tao Zhu
(Tianjin University of Science and Technology, China)

OS10-4 *Research on Chaos Synchronization of Qi System and Lü System with Different Structures*
Hanyuan Wang, Yiting Gao
(Tianjin University of Science and Technology, China)

©ICAROB 2023 ALife Robotics Corp. Ltd.

OS11 Intelligent Systems and Robotics – 3 (4)

Chair: Jiaxin Li (Tianjin University of Science and Technology, China)
Co-Chair: Fengzhi Dai (Tianjin University of Science and Technology, China)

OS11-1 *Generation and Analysis of a Multi-scroll Conservative Chaotic System*
Jiaxin Li, Yong Liu, Min Zhao (Tianjin University of Science and Technology, China)

OS11-2 *A Design of Fire Detection Device Based on YOLOv5*
Zhiyang Li, Yande Xiang, Haoyu Guo, Yu Chen, Wenxuan Pan, Fengzhi Dai
(Tianjin University of Science and Technology, China)

OS11-3 *Application of Convolutional Neural Network in Accurate Breast Cancer Identification*
Zhiyang Li, Haoyu Guo, Yande Xiang, Wentao Kuang, Lu Chen
(Tianjin University of Science and Technology, China)

OS11-4 *Intelligent Infusion Service Based on Open MV*
Haoran Gong, Zongyi Li, Qi Chu, Siyuan Liu, Feiyang Qu, Lu Wang
(Tianjin University of Science and Technology, China)

10:45-11:45 OS3 Advanced Information Processing Applications (4)

Chair: Toru Hiraoka (University of Nagasaki, Japan)
Co-Chair: Masaharu Hirota (Okayama University of Science, Japan)

OS3-1 *Generation of Arbitrarily-Oriented Ripple Images Using Smoothing Filter with Translated Window*
Shogo Noma, Toru Hiraoka (University of Nagasaki, Japan)

OS3-2 *Generation of Moire-Like Videos from RGB-D Videos*
Sho Enomoto, Toru Hiraoka (University of Nagasaki, Japan)

OS3-3 *A Proposal of Shoulder-surfing Attack Countermeasure Method with Improved Usability*
Yoshihiro Kita, Shingo Nakamura (University of Nagasaki, Japan)

OS3-4 *User-movement Estimation in Social Media Sites Based on Seq2Seq Model*
Masaharu Hirota (Okayama University of Science, Japan)

©ICAROB 2023 ALife Robotics Corp. Ltd.

14:20-16:40 OS25 Emerging Technologies in Industrial Revolution 4.0 (IR 4.0) Era (8)
Chair: Takao Ito (Hiroshima University, Japan)
Co-Chair Wei Hong Lim (UCSI University, Malaysia)

OS25-1 *Chaotic African Vultures Optimization Algorithm for Feature Selection*
Wy-Liang Cheng[1], Li Pan[1], Mohd Rizon Bin Mohamed Juhari[1], Chin Hong Wong[2,3], Abhishek Sharma[4], Tiong Hoo Lim[5], Sew Sun Tiang[1], Wei Hong Lim[1]
([1]UCSI University, Malaysia), ([2]Maynooth University, Ireland), ([3]Fuzou University, China), ([4]Graphic Era University, India), ([5]Universiti Teknologi Brunei, Indonesia)

OS25-2 *Multi Chaotic Flow Directional Algorithm for Feature Selection*
Wy-Liang Cheng[1], Li Pan[1], Mohd Rizon Bin Mohamed Juhari[1], Abhishek Sharma[2], Hameedur Rahman[3], Chun Kit Ang[1], Sew Sun Tiang[1], Wei Hong Lim[1]
([1]UCSI University, Malaysia), ([2]Graphic Era University, India), ([3]Air University, Pakistan)

OS25-3 *Performance Comparison of Convolutional Neural Networks for COVID-19 Diagnosis*
Suhaim Parvez Wadekar[1], Koon Meng Ang[1], Chin Hong Wong[2,3], Abhishek Sharma[4], Tiong Hoo Lim[5], Chun Kit Ang[1], Sew Sun Tiang[1], Wei Hong Lim[1]
([1]UCSI University, Malaysia), ([2]Maynooth University, Ireland), ([3]Fuzou University, China), ([4]Graphic Era University, India), ([5]Universiti Teknologi Brunei, Indonesia)

OS25-4 *Classification of Wafer Defects with Optimized Deep Learning Model*
Koon Hian Ang[1], Koon Meng Ang[1], Mohd Rizon Bin Mohamed Juhari[1], Chin Hong Wong[2,3], Abhishek Sharma[4], Chun Kit Ang[1], Sew Sun Tiang[1], Wei Hong Lim[1]
([1]UCSI University, Malaysia), ([2]Maynooth University, Ireland), ([3]Fuzou University, China), ([4]Graphic Era University, India)

OS25-5 *Compact Wearable Antenna for Millimeter-Wave (mm-Wave) Fifth Generation (5G)*
Wai Kiat Wong[1], Sew Sun Tiang[1, *], Wei Hong Lim[1], Mastaneh Mokayef[1], Chin Hong Wong[2,3] ([1]UCSI University, Malaysia), ([2]Fuzou University, China),
([3]Maynooth University, Ireland)

OS25-6 *Driver's Fatigue Recognition using Convolutional Neural Network Approach*
Samer Abbas[1], Sew Sun Tiang[1], Wei Hong Lim[1], Li Sze Chow[1], Chin Hong Wong[2,3]
([1]UCSI University, Malaysia), ([2]Fuzou University, China),
([3]Maynooth University, Ireland)

OS25-7 *Deep Residual Neural Network for Efficient Traffic Sign Detection System*
Hanlin Cai[1,2], Jiaqi Hu[1,2], Zheng Li[1,2], Wei Hong Lim[3], Sew Sun Tiang[3],
Chin Hong Wong[1,2]
([1]Fuzou University, China), ([2]Maynooth University, Ireland), ([3]UCSI University, Malaysia)

©ICAROB 2023 ALife Robotics Corp. Ltd.

OS25-8 *Wall Crack Detection based on Adaptive Double Threshold Grayscale Transform*
Mingrui Lin[1,2], Xin Xu[1,2], Tengxiang Li[1,2], Yuhang Hong[1,2], Weiqin Wang[1,2], Shilin Chen[1,2], Wei Hong Lim[3], Chin Hong Wong[1,2]
([1]Fuzou University, China), ([2]Maynooth University, Ireland), ([3]UCSI University, Malaysia)

17:10-17:55 OS19 Natural Computing (2)
Chair: Marion Oswald (Technische Universität Wien, Austria)
Co-Chair: Yasuhiro Suzuki (Nagoya University, Japan)

OS19-1 *Changes in the Behavior of a Small Number of Molecular Systems*
Yasuhiro Suzuki (Nagoya University, Japan)

OS19-2 *Retrieval by Sensory Information*
Yasuhiro Suzuki (Nagoya University, Japan)

Meeting Room 2
9:00-10:30 OS27 Research Towards the Sustainable Development Goals (SDG's) (6)
Chair: Ammar A.M. Al Talib (UCSI University, Malaysia)
Co-Chair: Takao Ito (Hiroshima University, Japan)

OS27-1 *Solar Powered Seed Sprayer Machine*
Ammar A.M. Al Talib, Yap Chee Xian, Ain Atiqa, Nor Fazilah Abdullah
(UCSI University, Malaysia)

OS27-2 *Solar Powered Outdoor Air Purifier With Air Quality Monitoring*
Ammar A.M.Al Talib, IK Chu Aung, Noor Idayu M. Tahir (UCSI University, Malaysia)

OS27-3 *Design and Fabrication of a Mutual Control Electronic Circuit for Solar and Electrical Water Heating*
Ammar A.M.Al Talib[1], Sarah Atifah Saruchi[2]
([1]UCSI University, Malaysia), ([2]UTP, Malaysia)

OS27-4 *Optimization of the Major Factors Affecting The Home Recycling of Disposed (LDPE) Plastics*
Ammar A.M.Al Talib, Chua Ray Sern Grayson, Ain Atiqa (UCSI University, Malaysia)

OS27-5 *Investigation of the Mechanical Properties and Applicability of Recycled Plastic Bags*
Ammar A.M.Al Talib, Wan Wai Kit, Nor Fazilah Binti Abdullah
(UCSI University, Malaysia)
OS27-6 *Design and Fabrication of Power Generating Treadmill*
Ammar A. M. Al-Talib[1], They Kai Yang[1], Sarah Atifah[2], Noor Idayu M.Tahir[1]
([1]UCSI University, Malaysia), ([2]UTP, Malaysia)

©ICAROB 2023 ALife Robotics Corp. Ltd.

10:45-12:00 OS8 Intelligent Control (5)

Chair: Yingmin Jia (Beihang University, P.R.China)
Co-Chair: Weicun Zhang (University of Science and Technology Beijing, P.R.China)

OS8-1 *A Self-triggering Control Based on Adaptive Dynamic Programming for Nonzero-sum* Game Systems
Yibo Shi, Chaoli Wang (University of Shanghai for Science and Technology, China)

OS8-2 *Harmony of Agent System with Heterogeneity*
Yunzhong Song[1], Weicun Zhang[2], Fengzhi Dai[3], Huimin Xiao[4], Shumin Fei[5], Jichao Zhao[6]
([1]Henan Polytechnic University, P.R.China) ([2]University of Science and Technology Beijing, P.R.China), ([3]Tianjin University of Science and Technology, P.R.China) ([4]Henan University of Economics and Law, P.R.China) ([5]South East University, P.R.China), ([6]Tianjin University of Science and Technology, P.R.China)

OS8-3 *Apple Grading based on IGWO Optimized Support Vector Machine*
Yi Zhao[1], Qunpo Liu[1,2], Yuxi Zhao[1], Yueqin Sheng[1]
([1]Henan Polytechnic University, P.R.China), ([2]Henan International Joint Laboratory of Direct Drive and Control of Intelligent Equipment, P.R.China)

OS8-4 *Cartesian Space Coordinated Impedance Control of Redundant Dual-Arm Robots*
Yang Zhang, Yingmin Jia (Beihang University (BUAA), China)

OS8-5 *Disturbance Observer-based Anti-unwinding Control for Flexible Spacecrafts*
Qian Sun[1], Yingmin Jia[1], Weicun Zhang[2]
([1]Beihang University (BUAA), China)
([2]University of Science and Technology Beijing, P.R.China)

14:20-16:55 OS13 Industrial Artificial Intelligence Robotics (9)

Chair: Eiji Hayashi (Kyushu Institute of Technology, Japan)
Co-Chair: Sakmongkon Chumkamon (Kyushu Institute of Technology, Japan)

OS13-1 *Object Status Detection in clutter environment for robot grasping using Mask R-CNN*
Kasman, Eiji Hayashi (Kyushu Institute of Technology, Indonesia, Japan)

OS13-2 *Deep-Learning-Based Designed Weight Picking Noodle-like Object*
Nattapat Koomklang, Sakmongkon Chumkamon, Prem Gamolped, Tomofumi Tsuji, Eiji Hayashi (Kyushu Institute of Technology, Japan)

OS13-3 *Research on grasping of string foods in the home meal replacement industry*
Akihiro Ooya, Sakmongkon Chumkamon, Prem Gamolped, Tomofumi Tsuji, Eiji Hayashi (Kyushu Institute of Technology, Japan)
Abbe Mowshowitz (The City College of New York, USA)

©ICAROB 2023 ALife Robotics Corp. Ltd.

OS13-4 *Development of Drifting Debris Detection System using Deep Learning on Coastal Cleanup*
Shintaro Ogawa, Sackmongkon Chumkamon, Eiji Hayashi (Kyushu Institute of Technology, Japan),
Ayumu Tominaga (National Institute of Technology, Kitakyushu College, Japan)

OS13-5 *Soft Object Dexterous Manipulation Using Deep Reinforcement Learning*
Sornsiri Promma, Sakmongkon Chumkamon, Eiji Hayashi (Kyushu Institute of Technology, Japan)

OS13-6 *Practical Implementation of FastSLAM for Forestry Robot*
Sylvain Geiser[1], Sakmongkon Chumkamon[1], Ayumu Tominaga[2], Takumi Tomokawa[1], Eiji Hayashi[1]
([1]Kyushu Institute of Technology, Japan)
([2]National Institute of Technology, Kitakyushu College, Japan)

OS13-7 *Research on AR system for industrial robot introduction*
Takuya Matsumoto[1], Eiji Hayashi[1], Sakmongkon Chumkamon[1], Tomofumi Tsuji[1],
Ayumu Tominaga[2], Abbe Mowshowitz[3]
([1]Kyushu Institute of Technology, Japan), ([2]National Institute of Technology Kitakyushu, Japan),
([3]The City College of New York, USA)

OS13-8 *The BCRobo dataset for Robotic Vision and Autonomous Path Planning in Outdoor Beach Environment*
Tan Chi Jie[1], Takumi Tomokawa[1], Sylvain Geiser[1], Shintaro Ogawa[1], Ayumu Tominaga[2],
Sakmongkon Chumkamon[1], Eiji Hayashi[1]
([1]Kyusyu Institute of Technology, [2]National Institute of Technology (Kitakyushu College), Japan)

OS13-9 *The research about editing system of performance information for player piano.*
-Develop inference methods using machine learning -
Takaaki Ueno, Sakmongkon Chumkamon, Eiji Hayashi (Kyushu Institute of Technology, Japan)

17:10-17:55 GS7 Applications II (3)

Chair: Hiroaki Wagatsuma (Kyushu Institute of Technology, Japan)

GS7-1 *An Accuracy Evaluation of Multibody Dynamics for the Knee Support Exoskeleton Model with Respect to Implicit Methods for Numerical Integration*
Shintaro Kasai, Hiroaki Wagatsuma (Kyushu Institute of Technology, Japan)

GS7-2 *A Basic Concept of the Nonlinear Oscillator-Based Hough Transform Implementation to Improve the Voting Procedure in the Scheme of Continuous Dual Spaces*
Amarbold Purev, Hiroaki Wagatsuma (Kyushu Institute of Technology, Japan)

GS7-3 *Survey on Harness Design for CubeSats: Understanding the Constraints of CubeSats Design and Toward an Optical Wireless Bus for CubeSats*
Masahiro Tokumitsu[1], Masatoshi Tsuji[2], Jun Nakaya[3]
([1]National Institute of Technology (KOSEN), Yonago College, Japan)
([2]National Institute of Technology (KOSEN), Kagawa College, Japan)
([3]Aichi Institute of Technology, Japan)

©ICAROB 2023 ALife Robotics Corp. Ltd.

Meeting Room 3
9:00-10:30 OS26 Signals Processing and Automation I (6)
Chair: Norrima Mokhtar (University of Malaya, Malaysia)
Co-Chair: Heshalini Rajagopal (UCSI University, Malaysia)

OS26-1 *Noise Filtering of Hyperspectral Data of Oil Palms by Median-Mean Projection Filtering*
Imanurfatiehah Ibrahim[1], Hamzah Arof[1], Mohd Izzuddin Anuar[2],
Mohamad Sofian Abu Talip[1]
([1]Universiti Malaya, Malaysia), ([2]Malaysian Palm Oil Board, Malaysia)

OS26-2 *Smart Telehealth Appointment System – WI Care*
Siah Cheong Lin, Heshalini Rajagopal, Chloe Thong Chee Ling
(UCSI University, Malaysia)

OS26-3 *Web-based Stocktaking application in Businesses*
Cheah Dei Xuan, Heshalini Rajagopal, Shayla Islam (UCSI University, Malaysia)

OS26-4 *Pharmacy Warehouse Management System*
Gan Jhui Ken, Heshalini Rajagopal, Shaik Shabana Anjum (UCSI University, Malaysia)

OS26-5 *A Development of an Automatic Allocation Parking System*
Ng Wai Lam, Neesha Jothi, Shayla Islam (UCSI University, Malaysia)

OS26-6 *Healthcare Mobile Application*
Yee Chee Hong, Neesha Jothi, Javid Iqbal (UCSI University, Malaysia)

10:45-11:45 OS23 STEM Education for fostering Innovators (4)
Chair: Hiroyuki Y. Suzuki (Hiroshima University, Japan)
Co-Chair: Kazuo Kawada (Hiroshima University, Japan)

OS23-1 *Junior High School Rescue Robot Challenge for Fostering Problem-Solving Skills*
Kazuo Kawada, Keita Murai, Yuta Susawa, Hiroyuki Y. Suzuki
(Hiroshima University, Japan)

OS23-2 *A Study of Experiential Learning Activities using Model Materials for the Kicking Motion*
Teruyuki Tamai, Shoki Takeuchi, Yoshihiro Ohnishi (Ehime University, Japan)
Kazuo Kawada (Hiroshima University, Japan)

OS23-3 *Making High Precision Single Balance in Active Learning Seminar for Hiroshima Univ. Monozukuri Junior Doctor Special Educational Program*
Hiroyuki Y. Suzuki, Kazuo Kawada, Masayasu Nagamatsu
(Hiroshima University, Japan)

©ICAROB 2023 ALife Robotics Corp. Ltd.

OS23-4 *Capstone Class of "Mechatronics Innovation Project" as STEM Educational Curriculum for Teacher Training Course*
Hiroyuki Y. Suzuki, Masayasu Nagamatsu, Kazuo Kawada
(Hiroshima University, Japan)

14:20-15:05 OS16 Robotic Manipulation (3)
Chair: Kensuke Harada (Osaka University, Japan)
Co-Chair: Tokuo Tsuji (Kanazawa University, Japan)
Co-Chair: Akira Nakamura (Saitama Institute of Technology, Japan)

OS16-1 *Acquisition of Synergy for Low-dimensional Control of Multi-fingered Hands by Reinforcement Learning*
Kazuki Higashi, Tomohiro Motoda, Akiyoshi Hara, Kensuke Harada (Osaka University, Japan)

OS16-2 *Error Recovery Techniques 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)

OS16-3 *Flexible assembly system with stiffness switching joint*
He Maike, Tokuo Tsuji, Naoki Ichikawa, Takuro Sawada, Tatsuhiro Hiramitsu, Hiroaki Seki
(Kanazawa University, Japan)

15:40-17:55 OS6 Modelling and Simulation in Robotics (8)
Chair: Evgeni Magid (Kazan Federal University, Russia)
Co-Chair: Kuo-Hsien Hsia (National Yunlin University of Science & Technology, Taiwan)

OS6-1 *Android Based Educational Mobile Robot Design and Pilot Evaluation*
Elvira Chebotareva, Maksim Mustafin (Kazan Federal University, Russia)

OS6-2 *Virtual Collaborative Cells Modelling for UR3 and UR5 Robots in Gazebo Simulator*
Ramir Sultanov, Shifa Sulaiman, Tatyana Tsoy, Elvira Chebotareva
(Kazan Federal University, Russia)

OS6-3 *New Features Implementation for Servosila Engineer Model in Gazebo Simulator for ROS Noetic*
Alexandra Dobrokvashina[1], Shifa Sulaiman[1], Timur Gamberov[1], Kuo-Hsien Hsia[2], Evgeni Magid[1]
([1]Kazan Federal University, Russia)
([2]National Yunlin University of Science & Technology, Taiwan)

OS6-4 *Features of Interaction Between a Human and a Gestures-controlled Collaborative Robot in an Assembly Task: Pilot Experiments*
Maksim Mustafin[1], Elvira Chebotareva[1], Hongbing Li[2], Martínez-García Edgar A[3], Evgeni Magid[1,4]
([1]Kazan Federal University, Russia) ([2]Shanghai Jiao Tong University, China)
([3]The Autonomous University of Ciudad Juarez, Mexico) ([4]HSE University, Russia)

©ICAROB 2023 ALife Robotics Corp. Ltd.

OS6-5 *Modeling of Human Actions in a Collaborative Robotic Space Using AR601M Humanoid Robot: Pilot Experiments in the Gazebo Simulator*
Aidar Zagirov, Artem Apurin , Elvira Chebotareva (Kazan Federal University, Russia)

OS6-6 *Modern Methods of Map Construction Using Optical Sensors Fusion*
Ramil Safin[1], Tatyana Tsoy[1], Roman Lavrenov[1], Ilya Afanasyev[1], Evgeni Magid[1,2]
([1]Kazan Federal University, Russia) ([2]HSE University, Russia)

OS6-7 *Omniwheel Chassis' Model and Plugin for Gazebo Simulator*
Artem Apurin[1], Bulat Abbyasov[1], Alexandra Dobrokvashina[1], Yang Bai[2], Mikhail Svinin[2], Evgeni Magid[1]
([1]Kazan Federal University, Russia) ([2]Ritsumeikan University, Japan)

OS6-8 *Using Optical Sensors for Industrial Human-Robot Interaction in Gazebo Environment*
Elaman Kidiraliev, Roman Lavrenov (Kazan Federal University, Russia)

Meeting Room 4
9:00-10:15 OS15 Robot Control (5)
Chair: Yizhun Peng (Tianjin University of Science and Technology, China)

OS15-1 *Design of Intelligent Crutch System Based on STM32 and Raspberry Pie*
Zongxuan Zhang, Jianhao Hu, Yizhe Sun, Yizhun Peng
(Tianjin University of Science and Technology, China)

OS15-2 *Design of Intelligent Fish Box Based on Machine Vision and Internet of Things Technology*
Suqing Duan, Jiangyu Wu, Shuai Chen, Yizhun Peng
(Tianjin University of Science and Technology, China)

OS15-3 *Design and Implementation of Internet of Things Planting System Based on esp32 MCU*
Cuiying Ji, Yizhun Peng (Tianjin University of Science and Technology, China)

OS15-4 *A Customized Dispensing Robot Based on OpenMV Visual Recognition*
Hongze Liu, Yizhun Peng (Tianjin University of Science and Technology, China)

OS15-5 *An Intelligent Guide Hat Based on The Internet of Things*
Suqing Duan, Yizhun Peng (Tianjin University of Science and Technology, China)

©ICAROB 2023 ALife Robotics Corp. Ltd.

10:45-11:45 GS3 Image Processing II (4)

Chair: Joo Kooi Tan (Kyushu Institute of Technology, Japan)

GS3-1 *3D Point Cloud Registration and Segmentation of Reflective Metal Objects Using Go-ICP and Improved RANSAC*
Kairi Morita, Ziyue Liu, Jing Cao, Seiji Ishikawa, Masuhiro Nitta, Joo Kooi Tan
(Kyushu Institute of Technology, Japan)

GS3-2 *Development of a Safe Walking Assistive System for Visually Impaired Person Using MY VISION — Estimation of a Safe Passage from Sidewalk Information Based on Transfer Learning of VGG-16 Network*
Takumi Yokote, Joo Kooi Tan (Kyushu Institute of Technology, Japan)

GS3-3 *Detection of Fallen Persons and Person Shadows from Drone Images*
Taisei Ono, Haruka Egawa, Seiji Ishikawa, Joo Kooi Tan (Kyushu Institute of Technology, Japan)

GS3-4 *A Systematic Literature Review on Emotion Recognition System in Malaysia*
Muhammad Nadzree Mohd Yamin, Kamarulzaman Ab. Aziz, Tan Gek Siang,
Nor Azlina Ab. Aziz (Multimedia University, Malaysia)

14:20-16:25 OS21 Advances in Field Robotics and Their Applications (7)

Chair: Keisuke Watanabe (Tokai University, Japan)
Co-Chair: Kazuo Ishii (Kyushu Institute of Technology, Japan)

OS21-1 *Underwater Live Video Streaming Experiment Using Radio Frequency Communication for AUVs*
Raji Alahmad, Yuya Nishida, Kazuo Ishii, Yukihiro Fukumoto
(Kyushu Institute of Technology, Japan)

OS21-2 *Control strategy to change the locomotion mode of a reconfigurable wheel/track robot based on the soil conditions*
Supaphon Kamon, Enrico di Maria, Kazuo Ishii (Kyushu Institute of Technology, Japan)

OS21-3 *Sea-floor Image Restoration with Variable Absorbance Coefficient*
Irmiya R. Inniyaka, Yuya Nishida, Kazuo Ishii (Kyushu Institute of Technology, Japan)

OS21-4 *Design of A Parameter Update Method of the Database-Driven PID Controller Considering Norm of the System*
Takumi Ueda, Kazuo Ishii (Kyushu Institute of Technology, Japan)

OS21-5 *Occluded Objects Detection by Ultrasonic Sensors*
Ryuugo Mochizuki, Yuya Nishida and Kazuo Ishii (Kyushu Institute of Technology, Japan)

OS21-6 *Development of a Variable Stiffness Function for a New Multifunctional Wire Driven Joint Mechanism*
Katsuaki Suzuki, Yuya Nishida, Kazuo Ishii (Kyushu Institute of Technology)

©ICAROB 2023 ALife Robotics Corp. Ltd.

OS21-7 *Analyzing an OFDM using Cyclic Prefix to Improve the Underwater Communication System*
Alraie Hussam, Kazuo Ishii (Kyushu Institute of Technology, Japan)

17:10-17:55 OS30 Embedded systems projects (3)
Chair: Mastaneh Mokayef (UCSI University, Malaysia)
Co-Chair: Takao Ito (Hiroshima University, Japan)

OS30-1 *Data Transmission by Li-Fi in Coal Mining*
Myadadha Goutham Reddy[1], Tan Jia Wei[1], Mandigiri Golla Jagan Kumar[1], Vytla Poojitha[1], Kaki Harshavardhan Goud[1], Wang Ji[1], MHD Amen Summakieh[1], Atefeh Mohammadpour[2], Mastaneh Mokayef[1]*
([1]UCSI University, Malaysia) ([2]California State University, USA)

OS30-2 *Embedded Table Tennis Ball Launcher with a Trajectory Path Analyzer for Junior Players*
Mastaneh Mokayef[1], Lee Qi Jian[1], Bushra Naeem[2], Miad Mokayef[1], M.K.A Ahamed Khan[1], MHD Amen Summakieh[1]
([1]UCSI University, Malaysia) ([2]Balochistan University of Information Technology, Pakistan)

OS30-3 *Table Tennis Tournament Scores and Statistics Web Application*
Belal Khaled[1], Mastaneh Mokayef[1*], Chin Hong Wong[2], Sew Sun Tiang[1], Wei Hong Lim[1], MHD Amen Summakieh[1], Miad Mokayef[1]
([1]UCSI University, Malaysia) ([2] Fuzou University, China)

February 12 (Sunday)

8:40-Registration

Conference Room

10:30-11:30

Plenary Speech PS4

Chair: Takao Ito (Hiroshima University, Japan)

PS4 *A New Style of Research and Development from the EU Perspective*

Masato Nakagawa

(Denso Corporation, Fellow, Hiroshima University, Guest Professor, Japan)

©ICAROB 2023 ALife Robotics Corp. Ltd.

Meeting Room 1
9:00-9:45 OS24 Applications of Unmanned Aerial System (3)
Chair: Hazry Desa (Universiti Malaysia Perlis, Malaysia)
Co-Chair: Nurfadzillah Ishak (Universiti Malaysia Perlis, Malaysia)

OS24-1 *Effect of Spraying Dispersion Using UAV Spraying System with Different Height at Paddy Field*
Hazry Desa[1], Muhammad Azizi Azizan[1], Nik Noriman Zulkepli[1], Nurfadzillah Ishak[1], Teh Xi Hang[1], Siti Syuhaidah Yahya[1], Aisyah Arina Mohammad Shahrazel[1], Fakhrul Mukmin Mansor[1], Siti Zaleha Abdul Aziz[2], Abadal-Salam T Hussain[3]
([1]universiti Malaysia Perlis, Malaysia), ([2]Mara Japan Industrial Institute (MJII), Malaysia)
([3]Al-Kitab University, Iraq)

OS24-2 *Spraying Dispersion Analysis with Different Nozzle Types Using UAV Spraying System in a Paddy Field*
Hazry Desa[1], Muhammad Azizi Azizan[1], Nik Noriman Zulkepli[1], Nurfadzillah Ishak[1], Tan Yew Tian[1], Siti Syuhaidah Yahya[1], Aisyah Arina Mohammad Shahrazel[1], Fakhrul Mukmin Mansor[1], Siti Zaleha Abdul Aziz[2], Abadal-Salam T Hussain[3]
([1]universiti Malaysia Perlis, Malaysia), ([2]Mara Japan Industrial Institute (MJII), Malaysia)
([3]Al-Kitab University, Iraq)

OS24-3 *The capabilities and Readiness of Unmanned Aerial System (UAS) implementation in Construction Work Progression*
Nurfadzillah Ishak, Muhammad Azizi Azizan, Hazry Desa (universiti Malaysia Perlis, Malaysia)

12:30-14:15 OS1 Human-Machine Interface Application (7)
Chair: Norrima Mokhtar (University of Malaya, Malaysia)
Co-Chair: Heshalini Rajagopal (UCSI University, Malaysia)

OS1-1 *Arduino Based Smart IoT Food Quality Monitoring System*
Ashraf Ali Jamal Deen[1], Thivagar Chettiar Sarawanan[1], Heshalini Rajagopal[2], Devika Sethu[1], Neesha Jothi[2], Raenu Kolandaisamy[2]
([1]Manipal International University, Malaysia), ([2]UCSI University, Malaysia)

OS1-2 *Development of Image Quality Assessment (IQA) For Haze Prediction*
Heshalini Rajagopal[1], Sayanth Sudheer[2], Neesha Jothi[1], Keoy Kay Hooi[1], Norrima Mokhtar[3]
([1]UCSI University, Malaysia), ([2]Manipal International University, Malaysia),
([3]University of Malaya, Malaysia)

OS1-3 *Development of IoT based Key Finder*
Sayanth Sudheer[1], Heshalini Rajagopal[2], Azam Mohammed Al-Qussari[1], Norrima Mokhtar[3], Neesha Jothi[2], Raenu Kolandaisamy[2]
([1]Manipal International University, Malaysia), ([2]UCSI University, Malaysia),
([3]University of Malaya, Malaysia)

©ICAROB 2023 ALife Robotics Corp. Ltd.

OS1-4 *Quality assessment for microscopic parasite images*
Muhammad Amirul Aiman Bin Asri[1], Norrima Mokhtar[1], Heshalini Rajagopal[2], Wan Amirul Wan Mohd Mahiyiddin[1], Yvonne Ai Lian Lim[1], Masahiro Iwahashi[3], Anees ul Husnain[1]
([1]University of Malaya, Malaysia), ([2]UCSI University, Malaysia),
([3]Nagaoka University of Technology, Japan)

OS1-5 *A study on the Impact of Limitations in Multi-Rotor UAVs on Coverage Path Planning*
Anees ul Husnain[1], Norrima Mokhtar[1], Noraisyah Mohamed Shah[1], Mahidzal Dahari[1], Muhammad Syazni1 Ikmal Ramlee[1], Heshalini Rajagopal[2], Masahiro Iwahashi[3]
([1]University of Malaya, Malaysia), ([2]UCSI University, Malaysia),
([3]Nagaoka University of Technology, Japan)

OS1-6 *Blood Vessels Segmentation in Eye Fundus Using Image Processing Algorithms*
Obaid Al-quraan, Hiam Alquran, Mohammed Alsalatie, Wan Azani Mustafa, Wan Khairunizam (University of Malaysia Perlis, Malaysia)

OS1-7 *Automated Diagnosis of Eye Fundus Images*
Ala'a Zyout, Hiam Alquran, Wan Azani Mustafa, Mohammed Alsalatie, Aaa Al-Badarneh, Wan Khairunizam (University of Malaysia Perlis, Malaysia)

Meeting Room 2
9:00-10:00 OS18-1 Mathematical Informatics (4)
12:30-13:15 OS18-2 Mathematical Informatics (3)
Chair: Takao Ito (Hiroshima University, Japan)
Co-Chair: Makoto Sakamoto (University of Miyazaki, Japan)

OS18-1 *A Fundamental Study on Car Sickness Using Data Science*
Tsutomu Ito[1], Seigo Matsuno[1], Makoto Sakamoto[2], Satoshi Ikeda[2], Takao Ito[3]
([1]NIT, Ube College, Japan), ([2]University of Miyazaki, Japan), ([3]Hiroshima University, Japan)

OS18-2 *An Analysis of Quoridor by reusing the results of reduced version*
Satoshi Ikeda[1], Iwanaga Takuro[1], Makoto Sakamoto[1], Takao Ito[2]
([1]University of Miyazaki, Japan), ([2]Hiroshima University, Japan)

OS18-3 *Prototype Software for Designing Hula Accessories*
Takumi Nakahara[1], Satoshi Ikeda[1], Amane Takei[1], Kenji Aoki[1], Makoto Sakamoto[1*], Tsutomu Ito[2], Takao Ito[3]
([1]University of Miyazaki, Japan), ([2]NIT, Ube College, Japan), ([3]Hiroshima University, Japan)

©ICAROB 2023 ALife Robotics Corp. Ltd.

OS18-4 *A Basic Study on Museum Exhibition Support Using AR Technology*
Kakeru Takemura[1], Satoshi Ikeda[1], Amane Takei[1], Masahiro Yokomichi[1], Makoto Sakamoto[1], Shuichi Kurogi[2], Tsutomu Ito[3], Takao Ito[4]
([1]University of Miyazaki, Japan), ([2]Miyazaki Prefectural Museum of Nature and History, Japan), ([3]NIT, Ube College, Japan), ([4]Hiroshima University, Japan)

OS18-5 *Tourism Support for Bioluminescent Fungi Using Video Technology*
Bidesh Biswas Biki[1], Kodai Hasebe[1], Fumito Hamakawa[1], Satoshi Ikeda[1], Amane Takei[1], Makoto Sakamoto[1], Shuichi Kurogi[2]
([1]University of Miyazaki, Japan), ([2]Miyazaki Prefectural Museum of Nature and History, Japan)

OS18-6 *Parallel wave sound analysis based on hierarchical domain decomposition method*
Amane Takei[1], Akihiro Kudo[2], Makoto Sakamoto[1]
([1]University of Miyazaki, Japan), ([2]Tomakomai Collage, Japan)

OS18-7 *Parallel full-wave electromagnetic field analysis based on hierarchical domain decomposition method*
Amane Takei, Nanako Mizoguchi, Kento Ohnaka, Makoto Sakamoto
(University of Miyazaki, Japan)

14:30-15:45 OS31 Image Processing and Reconstruction (5)

Chair: Xiaoyan Chen (Tianjin University of Science and Technology, China)
Co-Chair: Shiming Wang (Tianjin University of Science and Technology, China)

OS31-1 *Small-Target Detection Based on YOLOX*
Keying Ren[1], Xiaoyan Chen[1*] (Tianjin University of Science and Technology, China)

OS31-2 *Super Resolution Reconstruction Model Based on Attention Mechanism and Generative Adversarial Network*
Xia Miao, Xiaoyan Chen (Tianjin University of Science and Technology,China)

OS31-3 *Pedestrian-vehicle detection based on YOLOv7*
Xiaoyan Chen, Zhihui Chen (Tianjin University of Science and Technology, China)

OS31-4 *A lightweight, faster, more generalized low light image enhancement network*
Yutao Jin, Xiaoyan Chen (Tianjin University of Science and Technology, China)

OS31-5 *An Improved Landweber Method for Electrical Impedance Tomography*
Qian Wang, Xiaoyan Chen (Tianjin University of Science and Technology, China)

©ICAROB 2023 ALife Robotics Corp. Ltd.

Meeting Room 3
9:00-10:00 OS32 AR, VR and Digital Contents (4)
Chair: R.P.C. Janaka Rajapakse (Tainan National University of the Arts, Taiwan)
Co-Chair: Yoshimasa Tokuyama (Transworld University, Taiwan)

OS32-1 *A Kinect-based Augmented Reality Game for Arm Exercise*
Toshimasa Tokuyama (Transworld University, Taiwan)
R.P.C. Janaka Rajapakse (Tainan National University of the Arts, Taiwan)

OS32-2 *Development of EEG-based VR Application for Chakras Guided Meditation*
Chien-Tung Lin, R.P.C. Janaka Rajapakse (Tainan National University of the Arts, Taiwan)

OS32-3 *A Study on Flower Patterns of Temple Cut-and-Paste Decorations based on L-system*
Meng-Fan Huang, Tzu-Hsien Yuan, R.P.C. Janaka Rajapakse
(Tainan National University of the Arts, Taiwan)

OS32-4 *Pass-By: Development of Pedestrian Counts-based Art Installation for Passive Interaction*
Jianlin Zhong, R.P.C. Janaka Rajapakse
(Tainan National University of the Arts, Taiwan)

12:30-13:45 OS28 Is That Narratology 'Post'? -Computational and Cognitive Approaches and Beyond (5)
Chair: Jumpei Ono (Aomori University, Japan)
Co-Chair Hiroki Fxyma (Tainan University of Technology, Taiwan (ROC))
Co-Chair Yuki Hayashi (Chiba University / NINJAL, Japan)
Co-Chair Takashi Ogata (Iwate Prefectural University, Japan)

OS28-1 *Proposing YouTube-based vocabulary learning system*
Hiroki Fxyma (Tainan University of Technology, Taiwan (ROC))

OS28-2 *A Study on the Impact of "Flaming" on Content from "A Crocodile Who Will Die in 100 Days"*
Taiki Sugimoto, Jun Nakamura (Chuo University, Japan)

OS28-3 *Protocol analysis for constructing Verbalizing Support System*
Yuki Hayashi (Chiba University/NINJAL, Japan)

OS28-4 *Story generation during appreciating an artwork based on an actual tale (Ugetsu-monogatari)*
Akinori Abe (Chiba University, Japan)

OS28-5 *Designing a Narrative Generation Game Based on the Russian Invasion of Ukraine*
Jumpei Ono[1], Takashi Ogata[2]
([1]Aomori University, Japan), ([2]Iwate Prefectural University, Japan)

©ICAROB 2023 ALife Robotics Corp. Ltd.

14:30-15:45 GS2 Image Processing I (5)

Chair: Watcharin Tangsuksant (King Mongkut's University of Technology North Bangkok, Thailand)

GS2-1 *Microalgae Detection by Digital Image Processing and Artificial Intelligence*
Watcharin Tangsuksant, Pornthep Sarakon
(King Mongkut's University of Technology North Bangkok, Thailand)

GS2-2 *SmartCropPlanting: IoT-Based Mobile Application for Hydroponic System*
Sung Jun Kyu[1], Chit Su Mon[2], Kasthuri Subaramaniam[3]
([1,3]UCSI University, Malaysia) ([2]Heriot-Watt University Malaysia, Malaysia)

GS2-3 *Detection of Eye Misalignment Using an HMD with an Eye-tracking Capability*
Yoki Nagatomo, Noriyuki Uchida, Takuya Ikeda, Kayoko Takatsuka, Masayuki Mukunoki, Naonobu Okazaki (University of Miyazaki, Japan)

GS2-4 *Using OpenCV for real-time image recognition through augmented reality devices*
Gabdullina Dinara, Zykov Evgeniy, Kugurakova Vlada
(Kazan Federal University, Russia)

GS2-5 *A Structure Pattern Extraction by Using Morphological Component Analysis in the Aerial Image Edge Detection*
Wataru Oshiumi, Ankur Dixit, Hiroaki Wagatsuma
(Kyushu Institute of Technology, Japan)

Meeting Room 4

9:00-10:00 GS4 Robotics I (4)

Chair: Amornphun Phunopas (King Mongkut's University of Technology North Bangkok, Thailand)

GS4-1 *A Basic Study of Hand Eye Calibration using a Tablet Computer*
Junya Sato, Takayoshi Yamada, Kazuaki Ito (Gifu University, Japan)

GS4-2 *Human Detection with Uprisen Angle of a Camera for the Service Robot*
Watcharin Tangsuksant, Amornphun Phunopas, Pornthep Sarakon, Aran Blattler
(King Mongkut's University of Technology North Bangkok, Thailand)

GS4-3 *Synthesis of Drive Systems of Flapping and Feathering Motions for Bird-like Robot using Twist Drive Mechanism*
Jun Iwao, Hiroshi Ohtake (Kyushu Institute of Technology, Japan)

GS4-4 *Optimization Algorithm for Balancing QoS Configuration in Aggregated Robot Processing Architecture*
Abdul Jalil, Jun Kobayashi, Takeshi Saitoh (Kyushu Institute of Technology, Japan)

©ICAROB 2023 ALife Robotics Corp. Ltd.

12:30-13:30 GS5 Robotics II (4)

Chair: Takayoshi Yamada (Gifu University, Japan)

GS5-1 *Research on robotic assembly of gear motors (Stator recognition using keypoint matching and stator insertion using contact position estimation)*
Yasumoto Imai, Takayoshi Yamada, Junya Sato (Gifu University, Japan)
Toshiki Hayashi, Shota Aono (Tsubakimoto Chain Co., Japan)

GS5-2 *Robot Arm Operating Interface for Easy Grasping by Specifying the Gripping Width of End-effector*
Rio Takeuchi, Laijun Yang, Norihiko Kato (Mie University, Japan)

GS5-3 *D Real-Time Conversational Virtual Agents System: Do Facial Expressions and Camera Angles Persuade Human*
Pinkie Anggia, Kaoru Sumi (Future University Hakodate, Japan)

GS5-4 *Quasi-static Stability Analysis of Frictionless Planar Enveloping Grasps (Analysis of curvature effects at contact points)*
Takayoshi Yamada, Junya Sato, Kazuaki Ito, Hidehiko Yamamoto (Gifu University, Japan)

14:30-15:45 GS6 Applications I (5)

Chair: Minoru Kumano (University of Miyazaki, Japan)

GS6-1 *My Tally -A Personal Book Keeping Mobile Application*
Li Zhihan[1], Abdul Samad Shibghatullah[1], Nur Hazirah Hamdan[2]
([1]UCSI University, Malaysia), ([2]Universiti Teknologi, MARA)

GS6-2 *A Survey on Charitable Acts, Challenges and Using Charitable Mobile Application*
Nan Pepin[1], Abdul Samad Shibghatullah[1], Kasthuri Subaramaniam[1], Nur Hazirah Hamdan[2]
([1]UCSI University, Malaysia), ([2]Universiti Teknologi, MARA)

GS6-3 *Online Parcel Management System (PMS) for Small and Medium Company*
Chew Cheng Jin[1], Kasthuri Subaramaniam[1], Abdul Samad Shibghatullah[1], Salini Devi Rajendran[2], Nur Hazirah Hamdan[3]
([1]UCSI University, Malaysia), ([2]Taylor's University, Malaysia), ([3]Universiti Teknologi, MARA)

GS6-4 *Automatic Classification Method for Plastic Bottles and Caps Using Multi Attention Eff-UNet*
Shunsuke Moritsuka, Tohru Kamiya (Kyushu Institute of Technology, Japan)

©ICAROB 2023 ALife Robotics Corp. Ltd.

Group Meeting for the next Conference

©ICAROB 2023 ALife Robotics Corp. Ltd.

PS abstracts (4)

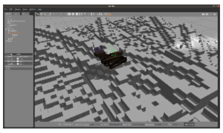
PS1 Human to Human Interaction using Virtual Agents Posing as Another Person

Kaoru Sumi (Future University Hakodate, Japan)

There have been ever expanding opportunities for online distance education in recent years, and agent-based interactions in virtual spaces have been attracting attention in this context. In a virtual space, we can communicate with others as if we were a virtual agent who is totally different from ourselves. Thanks to the recent advances in online software that enable us to communicate with virtual agents remotely, we will soon be able to have such experiences more easily. For example, a woman may become a male virtual agent, a man may become a female virtual agent, and someone could become a person of a different ethnicity. In this study, we investigate the various effects of interaction with virtual agents in a virtual space, such as how they feel about each other and what impacts there may be on a particular task.

PS2 Simulation Tools for Urban Search and Rescue Robotics

Evgeni Magid (Kazan Federal University, Russia)



Real world experiments are critical for validating performance of new concepts and lgorithms in robotics field. Yet, experiments tend to be too expensive in terms of time and resources of a research team. Moreover, it is not feasible to conduct thousands of complex experiments with a physical robot in a real environment. To check new ideas, preliminary evaluate new algorithms and interaction protocols, on first stages of a research project it is reasonable to start within a simulation. To produce relevant results, a simulator should provide adequate models of robots and environments with realistic physical properties. This paper presents an overview of our experience in using robot operating system (ROS) with Gazebo and Webots simulators for urban search and rescue robotics projects and considers constructing new models of mobile robots and complicated environments, algorithm validation and comparative analysis.

PS3 Enhancement methodology for low light image

Xiwen Liang, Xiaoyan Chen (Tianjin University of Science and Technology, China)

Raw Detection

Enhanced Detection

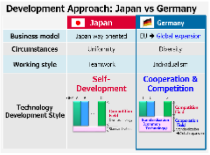
In order to solve the problems such as low brightness, high noise and poor contrast in under illumination images, there are several methods proposed to address this issue. Usually, these methods are categorized into two different ways. One is based on traditional light-based technology, the other is based on machine learning technology. The low-light image enhancement is often a challenging task because the noises in dark areas are amplified with the overall brightness and contrast of the image. With the development of machine learning techniques, deep learning networks are becoming the popular research topics recently to overcome the disadvantages of noisy dots. Based on the deep analysis of the current research work, we proposed a novel network and carried out lots of comparison experiments to analysis the performances of these methods. By training, validation and testing on the datasets, the evaluation critics are defined and utilized to analysis the efficiency of the methods. With the results, we draw the conclusion that the efficient low-light image method can make up for the shortcomings of the environment, bring better viewers' experience and provide preprocessing for subsequent high-level computer vision tasks, such as target recognition, face recognition, semantic segmentation, etc.

©ICAROB 2023 ALife Robotics Corp. Ltd.

PS4“A New Style of Research and Development from the EU Perspective”

Masato Nakagawa (Denso Corporation, Fellow, Hiroshima University, Guest Professor, Japan)

This paper introduces a new style of research and development with a unique process of engineering development based on the EU perspective. Two different fields will be separated in the new style: competition field and non-competition field. The former is a specific area encouraging companies to develop their unique technology as differentiation strategy, and the later stresses collaborations among different companies and organizations for spreading the standardization of common technologies. This new style with two different fields shows us a new direction of the engineering development in various engineering industry. In addition, this paper explains robot technologies in the manufacturing of automotive sector in terms of smart manufacturing concept.



Development Approach: Japan vs Germany

	Japan	Germany
Business model	Japan way oriented	EU → Global expansion
Circumstances	Uniformity	Diversity
Working style	Teamwork	Individualism
Technology Development Style	Self-Development	Cooperation & Competition

Competition Field
Cooperation Field

©ICAROB 2023 ALife Robotics Corp. Ltd.

OS abstracts

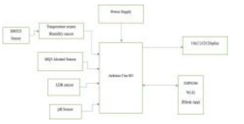
OS1 Human-Machine Interface Application (7)

Chair Norrima Mokhtar (University of Malaya, Malaysia)
Co-Chair Heshalini Rajagopal (UCSI University, Malaysia)

OS1-1 Arduino Based Smart IoT Food Quality Monitoring System

Ashraf Ali Jamal Deen[1], Thivagar Chettiar Sarawanan[1], Heshalini Rajagopal[2], Devika Sethu[1], Neesha Jothi[2], Raenu Kolandaisamy[2]
([1]Manipal International University, Malaysia), ([2]UCSI University, Malaysia)

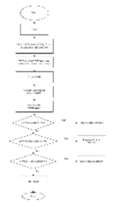
Food safety and hygienic as well as health are significant issues to stop food wastage. The high quality of the food requires to be kept track of and it should be also protected against deteriorating and decaying by the climatic variables like temperature level, humidity, and dark. In this paper, a comparable food quality monitoring tool will be created that will keep watch of ecological factors like temperature level, moisture, alcohol web content as well as exposure to light for fruits and vegetables. The system is built on Arduino UNO where it is interfaced with various sensors like DHT-22 to keep track of temperature level and humidity, MQ3 to identify alcohol material as well as LDR to gauge direct exposure to light. It sends the measured sensor data to an IoT system via ESP8266 Wi-Fi Module. The IoT system will certainly be made use of for logging and checking sensing unit data and this is beneficial in monitoring the food storage from anywhere and anytime.



OS1-2 Development of Image Quality Assessment (IQA) For Haze Prediction

Heshalini Rajagopal[1], Sayanth Sudheer[2], Neesha Jothi[1], Keoy Kay Hooi[1], Norrima Mokhtar[3]
([1]UCSI University, Malaysia), ([2]Manipal International University, Malaysia),
([3]University of Malaya, Malaysia)

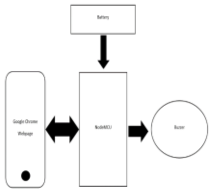
Haze is a term that is widely used in image processing to refer to natural and human activity-emitted aerosols. It causes light scattering and absorption, which reduce the visibility of captured images. This reduction hinders the proper operation of many photographic and computer vision applications, such as object recognition/localization. Therefore, an approach for haze density estimation is highly demanded. This paper proposes a model that is known as the haziness degree evaluator to predict haze density from a single image without reference to a corresponding haze-free image. The proposed model quantifies haze density by optimizing an objective function comprising haze-relevant features that result from correlation and computation analysis.



OS1-3 Development of IoT based Key Finder

Sayanth Sudheer[1], Heshalini Rajagopal[2], Azam Mohammed Al-Qussari[1], Norrima Mokhtar[3], Neesha Jothi[2], Raenu Kolandaisamy[2]
([1]Manipal International University, Malaysia), ([2]UCSI University, Malaysia),
([3]University of Malaya, Malaysia)

Typically, we misplace our keys and hunt for them throughout the home, eventually locating them with much difficulty after a long search. Therefore, we propose to develop a basic IoT-based Key finder utilizing NodeMCU, Buzzer, and Battery in this paper. In this paper, the development of key chain which can be attached to the keys will be explained. The paper also includes the creation of a website dedicated to the search for the missing keys. The missing keys can be located using a mobile phone's Google Chrome browser. When the webpage is enabled to find the missing keys, the designed IoT-based keychain is equipped with a buzzer that will make a beep sound. The developed IoT based key finder is important in terms of keeping track of the keys and could also save time.

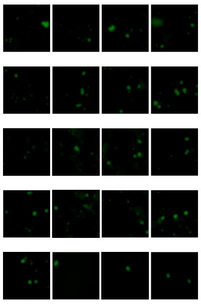


©ICAROB 2023 ALife Robotics Corp. Ltd.

OS1-4 Quality assessment for microscopic parasite images

Muhammad Amirul Aiman Bin Asri[1], Norrima Mokhtar[1], Heshalini Rajagopal[2], Wan Amirul Wan Mohd Mahiyiddin[1], Yvonne Ai Lian Lim[1], Masahiro Iwahashi[3], Anees ul Husnain[1]
([1]University of Malaya, Malaysia), ([2]UCSI University, Malaysia), ([3]Nagaoka University of Technology, Japan)

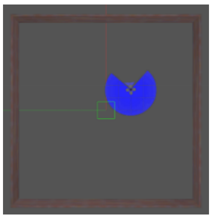
Water sample inspection is crucial for treated water monitoring, the quality of microscopic parasite images such as Giardia and Cryptosporidium need to be examined during treated water inspection. In this work, the subjective and objective evaluation of parasite images were performed. Parasite image database consisting of 20 reference images and 360 distorted images were used in the evaluation. The distorted images were generated from the reference images by applying distortion to the reference images with Gaussian White Noise and Motion Blur, at 9 levels of distortions. Twenty subjects scores obtained were transformed into Mean Opinion Score (MOS). In the objective evaluation, six Full Reference-IQA (FR-IQA) metrics, namely MSSIM, SSIM, FSIM, IWSSIM, GMSD and VIF were used to evaluate the distorted images. The subjective MOS scores were used as the benchmark to determine the most suitable objective IQA to assess parasite images. It was found that MSSIM is the most suitable IQA to assess parasite images distorted with Gaussian White Noise and Motion Blur.



OS1-5 A study on the Impact of Limitations in Multi-Rotor UAVs on Coverage Path Planning

Anees ul Husnain[1], Norrima Mokhtar[1], Noraisyah Mohamed Shah[1], Mahidzal Dahari[1], Muhammad Syazni[1] Ikmal Ramlee[1], Heshalini Rajagopal[2], Masahiro Iwahashi[3]
([1]University of Malaya, Malaysia), ([2]UCSI University, Malaysia), ([3]Nagaoka University of Technology, Japan)

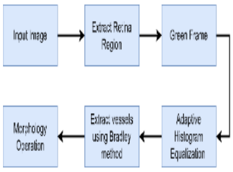
Search and explore missions through patrolling UAVs need effective strategies for area coverage. Various methodologies for coverage path planning were explored and analyzed through ROS-Gazebo simulation environment using Hector quadrotor model. Considering the impact of hardware limitations, simulations were conducted, for such missions where the UAV is needed to switch frequently between search and localize modes. This study investigated raster-scan exploration, expanding spiral search and zigzag pattern coverage to analyze the impact of limitations from Hector UAV on these models. The evaluation parameters were percentage of covered area, number of turns and time taken by the UAV.



OS1-6 Blood Vessels Segmentation in Eye Fundus Using Image Processing Algorithms

Obaid Al-quraan, Hiam Alquran, Mohammed Alsalatie, Wan Azani Mustafa, Wan Khairunizam
(University of Malaysia Perlis, Malaysia)

The retinal blood vessels have a huge impact on the diagnosis of eye diseases in addition to other systematic diseases in the human body. In this paper, we presented an automated segmentation method to extract retinal blood vessels, starting with preprocessing, then passing the image into segmentation stage using Bradley technique, and lastly, morphological operations. The proposed method was assessed and tested on STARE dataset, followed by comparing the auto-segmented images to the manually segmented ones. The comparison results Accuracy, Sensitivity, and Specificity were 94.63%, 95.02%, and 80.73% respectively.

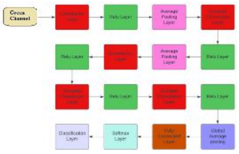

Input image
Extract Retina Region
Green Frame
Adaptive Histogram Equalization
Extract vessels using Bradley method
Morphology Operation

©ICAROB 2023 ALife Robotics Corp. Ltd.

OS1-7 Automated Diagnosis of Eye Fundus Images

Ala'a Zyout, Hiam Alquran, Wan Azani Mustafa, Mohammed Alsalatie, Aaa Al-Badarneh, Wan Khairunizam
(University of Malaysia Perlis, Malaysia)

Eye disease is a severe health problem. Advanced stages of the disease may lead to vision loss. Early detection may limit the development of the severity and enhance the chance of treatment. Computer-aided diagnosis (CAD) is the state-art-technology. This paper proposes a CAD system that combines image processing techniques and artificial intelligence. The proposed method used the green channel of fundus eye images to extract the most representative features by the trained convolutional neural network to classify five eye diseases of fundus images. The build CAD system exploits deep learning and support vector machine classifier to achieve a highly accurate model of 98% for five types of eye diseases.



OS2 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 Chuan-Kang Liu (National Chin-Yi University of Technology)

OS2-1 The Dam Gate Cybersecurity Testbed

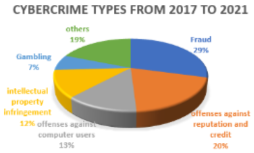
Chen-Yu Lee, I-Hsien Liu, Meng-Wei Chang, Jung-Shian Li (National Cheng Kung University, Taiwan)

The testbeds are very important for cybersecurity research on critical infrastructure. In today's drastic climate change, the dam gate control system is a very important part of the critical infrastructure for people's livelihood. In traditional research, because the real control system cannot be used directly, most of the research can only be carried out in a simulation way. The research based on simulation alone lacks practical value due to too many assumptions. This research was supported by the Water Resources Agency, Ministry of Economic Affairs and National Science and Technology Council in Taiwan. The gate control cybersecurity testbed was built with a blueprint of the real world.

OS2-2 Domain Name Infringement in Taiwan

Shih-Chin Lin (Ming Chuan University, Taiwan)

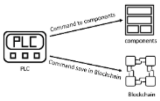
Registering a domain name can assist in attracting new customers and in cultivating a strong market presence. However, when the desired domain name has been registered by someone else or is similar to a trademark, proceeding with registration could infringe upon an existing domain name or a trademark. This is so-called domain name Infringement, which is often accompanied by trademark infringement or a criminal act. This study thoroughly examined the research reports published by the TWNIC and relevant authorities to gain an understanding of domain name Infringement in Taiwan, ultimately revealing that domain name ".tw" infringement has not yet gained its due attention in Taiwan.



OS2-3 Device's Operation Tracking using Blockchain in Industrial Control System

Chien-Hsin Wu, I-Hsien Liu, Jung-Shian Li (National Cheng Kung University, Taiwan)
Chu-Fen Li ((National Formosa University, Taiwan)

Many producing, monitoring and controlling needs are met by using programmable logic controllers. But there is no effective mechanism to audit PLC behavior. So this research designed a mechanism based on Blockchain for the purpose of effectively recording the commands and response actions received by the PLC. Due to the characteristics of the blockchain, the integrity of the data is also guaranteed.

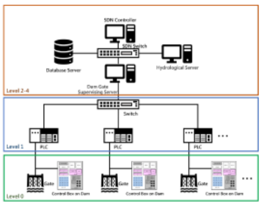


©ICAROB 2023 ALife Robotics Corp. Ltd.

OS2-4 Strengthen the security of the Industrial Control System using SDN technology

Min-Wei Huang, I-Hsien Liu, Hsin-Yu Lai, Meng-Huan Lee, Jung-Shian Li
(National Cheng Kung University, Taiwan)

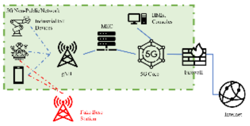
In the field of OT, most of the network architectures operated in the way of isolation from internal and external networks. Only firewalls are installed on the external network without any protection measures for the internal network. In this paper, we leverage a Software-defined network (SDN) with an industrial control system (ICS), so controllers can manage the equipment and keep track of each switch and its connection with the programmable logic controller (PLC) in the ICS. Additionally, by adding flow entries, only the critical flows can be allowed. So the transmission between the PLC and Human Machine Interface (HMI) can be protected. The transmission quality of the ICS and its availability can be improved.


SDN Controller
SDN Switch
Database Server
Hydrological Server
Dam Gate Supervising Server
Level 2-4
Switch
PLC
PLC
PLC
Level 1
Gate
Control Box on Dam
Gate
Control Box on Dam
Gate
Control Box on Dam
Level 0

OS2-5 Fake Base Stations threats in 5G Standalone Networks

Meng-Huan Lee, I-Hsien Liu, Jung-Shian Li (National Cheng Kung University, Taiwan)

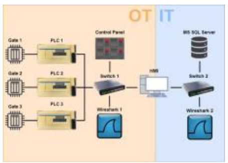
With 5G technology, traditional industrial and business equipment can now be connected wirelessly in a non-public network separated from public mobile services. Benefit from features such as high bandwidth, massive machine-type communications, and edge computing, while being able to control their own private 5G networks. But fake base stations or IMSI-Catchers used by law enforcement and hackers may collect private information and cause disruptions in cell services even if they're not public. In this research, we will analyze existing attack methods and detection mechanisms. And look at how those threats can affect the devices and operations in 5G non-public networks.


5G Non-Public Network
Industrial IoT Devices
gNB
MEC
HMIs, Consoles
5G Core
Firewall
Internet
Fake Base Station

OS2-6 Cyber-Physical Security Testbed for River Basin Gate Control System

Meng-Wei Chang, I-Hsien Liu, Jung-Shian Li (National Cheng Kung University, Taiwan)

Due to the flourishing development of critical infrastructures in recent years, increasing importance has been attached to the security of the Cyber-Physical System (CPS) of the infrastructures. Machine learning technology nowadays is evolving rapidly, and is widely implemented in detecting or preventing such attacks. As a result, This research constructs a Testbed to collect relevant data sets to support machine learning requirements, such as training models and analyzing attacks, etc.


OT
IT
Gate 1
PLC 1
Gate 2
PLC 2
Gate 3
PLC 3
Control Panel
Switch 1
Wireshark 1
HMI
MS SQL Server
Switch 2
Wireshark 2

©ICAROB 2023 ALife Robotics Corp. Ltd.

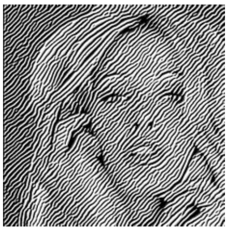
OS3 Advanced Information Processing Applications (4)

Chair Toru Hiraoka (University of Nagasaki, Japan)
Co-Chair Masaharu Hirota (Okayama University of Science, Japan

OS3-1 Generation of Arbitrarily-Oriented Ripple Images Using Smoothing Filter with Translated Window

Shogo Noma, Toru Hiraoka (University of Nagasaki, Japan)

A non-photorealistic rendering method for automatically generating ripple images from photographic images using region-division smoothing filter has been proposed. Ripple patterns are composed of continuous lines with fluctuations, and ripple images are expressed by superimposing ripple patterns on photographic images. To create ripple images that give different visual effects, this paper develops a method for generating ripple patterns with a texture different from the conventional method. The proposed method is executed by an iterative calculation using smoothing filter with the translated window. In the proposed method, the orientation of ripple patterns can be arbitrarily controlled by changing the amount of translation of the window used in smoothing filter. To verify the effectiveness of the proposed method, an experiment using various photographic images was conducted. Additionally, an experiment to visually examine how ripple patterns generated by changing the values of the parameters in the proposed method change.



OS3-2 Generation of Moire-Like Videos from RGB-D Videos

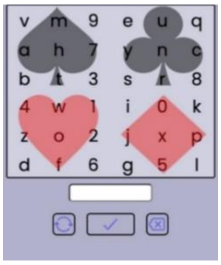
Sho Enomoto, Toru Hiraoka (University of Nagasaki, Japan)

A non-photorealistic rendering method has been proposed to generate moire-like images from photographic images using bilateral filter and unsharp mask. Extensions to the conventional method have also been proposed to generate moire-like videos from videos or to generate moire-like images from RGB-D images. In this paper, a method is proposed to generate moire-like videos from RGB-D videos. Flickering is a problem in NPR videos, but the proposed method can suppress flicker. Through experiments using an RGB-D video taken by the authors, the flicker of moire-like videos generated by the proposed method was evaluated visually and quantitatively.

OS3-3 A Proposal of Shoulder-surfing Attack Countermeasure Method with Improved Usability

Yoshihiro Kita, Shingo Nakamura (University of Nagasaki, Japan)

Shoulder-surfing attacks are one of the most familiar password exploitation attacks. It is vulnerable to be attacked while unlocking a smartphone screen. The smartphone users need to take countermeasures against that attack. The fingerprint-based screen unlock system has become the norm, but it is not safe, as there has been increased in the user's fingerprint theft. The existing methods to prevent the shoulder-surfing attacks are effective against such attacks, but many of them are complicated to operate, and difficult to use. In this paper, we propose the prevent method for surfing attacks that it is easy to use. The tool's user operates the lower buttons, moves characters to on the trump's marks in specified advance. The user can input as likely as the password. On the other hand, the attacker does not understand the input characters only shown these buttons has been pushed.



©ICAROB 2023 ALife Robotics Corp. Ltd.

OS3-4 User-movement Estimation in Social Media Sites Based on Seq2Seq Model

Masaharu Hirota (Okayama University of Science, Japan)

Many tourists upload content about tourist attractions to social media sites. The location information annotated on the content represents the user's movement. This movement information is an important source of information that can be used for recommendations and advertisements. However, users do not consistently post content about all the places they visit on social media. Therefore, this study aims to develop a method to estimate the location of users' movements. The proposed method uses a Seq2Seq model, which learns the reconstruction of users' movement trajectories.

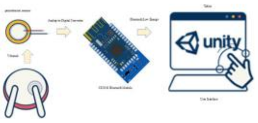
OS4 Artificial Life and Intelligent Systems (4)

Chair Chung-Wen Hung (National Yunlin University of Science and Technology, Taiwan)
Co-Chair Kuo-Hsien Hsia (National Yunlin University of Science and Technology, Taiwan)

OS4-1 Interactive Beating Drum Unity Game

Chung-Wen Hung, Cheng-Lung Ko, Wen-Huei Chou
(National Yunlin University of Science and Technology, Taiwan)

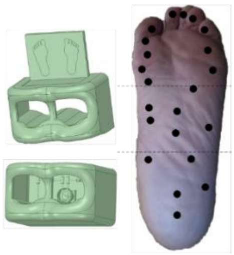
Interactive Beating Drum Unity Game is proposed in this paper. The pandemic forced individuals to maintain a prescribed social distance from others, and this way may ignore the individual's mental health and social needs, especially for the elderly. The Unity Engine is used to implement the system to provide a sport and social plane. And Bluetooth Low Energy (BLE) chip is adopted to transmit the beating signal to Unity Engine what following the Bluetooth (BT) protocol. Then the position beat by users will be showed on the user interface tablet. There are two play mode be implemented: single player mode and multiplayer mode, and the latter will include social function. The users are required to drum in short time interval accurately, and the game award points are evaluated for accuracy. The experimental results show the system workable, and the subjects unanimously stated that the operation of the game Easy-to-Use.



OS4-2 Error Backpropagation Neural Network Based Image Identification for a Foot Massage Machine and Its Mechanism Design

Chun-Chieh Wang (National Yunlin University of Science and Technology, Taiwan)

In the past ten years, many companies have developed different styles of foot massage machines. At present, the common massage products on the market include roller type and pressing type. However, it is very difficult to accurately stimulate all acupuncture points for different sizes of feet. Besides, the massage roller cannot be controlled independently. Therefore, a novel computer vision technology is proposed to identify the foot acupuncture points by error backpropagation neural network (EBNN) in this paper. First, we use cameras to capture the sole of users' soles and execute image preprocessing procedures to segment the region of interest (ROI) of soles. We map foot acupuncture points to foot images to obtain reference massage positions. Second, the YCbCr color space is used to separate the brightness to complete the segmentation of the foot image in the skin detection. Moreover, EBNN is used to train users' soles-image sets to improve the success rate of image segmentation. Finally, to improve the rate of image recognition and user convenience, a foot massage machine was redesigned. Experimental results validate the superiority and practicality of the proposed image identification method for foot massage machines.



©ICAROB 2023 ALife Robotics Corp. Ltd.

OS4-3 Cross-domain sharing of robots in the community caring and practice of university social responsibility

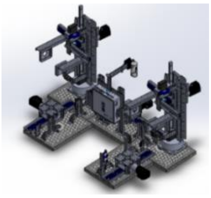
Jia-Ming Hsiao (National Yunlin University of Science and Technology, Taiwan)

This paper focuses on community caring and social practice that assisting robot education development in rural schools to echo the two goals of the United Nations' SDGs (Sustainable Development Goals)–"Goal 4: Quality education" and "Goal 10: Reduce inequality". At the same time, in response to the scientific and technological literacy-oriented needs of the Curriculum Guidelines of 12-Year Basic Education in 2019, it is shown how to assist rural schools in Tainan City of Taiwan to develop computational thinking and cross-domain applications of robots under the COVID-19. Four robot education suites including LEGO Mindstorms EV3, Makeblock mBot, mBot2, and VEX IQ are applied in the rural schools. The dilemma and solutions of robot education in rural schools are also discussed in this article. Universities, non-profit organizations and robot equipment vendors work together to practice social responsibility.

OS4-4 Research on Design of Implementation Mechanism for Similar Production Line

Chia-Nan Ko, Yi-Yu Li, Ting-Ru Ko, and Ting-Yi Chen (Nan Kai University of Technology, Taiwan)

Recently, enterprises are actively investing in the research and development of practical technologies for similar production line organizations. Faced with various technological innovations in smart automation. In this paper, the object is to research and integrate intelligence through practical functional solutions to design and establish a production line implementation mechanism. We use the myRIO controller as the main core to plan the factory automation production line structure through the practical function research for building similar production line implementation organizations. In this paper, combining theoretical methodology and practical experiments to construct institutional design, electromechanical system and intelligent automation procedure. Evaluate the efficiency of the design automated production lines to improve the capabilities of similar production lines.



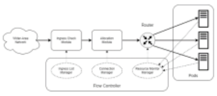
OS5 Intelligent Systems and Control (5)

Chair Kuo-Hsien Hsia (National Yunlin University of Science and Technology, Taiwan)
Co-Chair Chian C. Ho (National Yunlin University of Science and Technology, Taiwan)

OS5-1 Using Multithreaded Load Balancer to Improve Connection Performance in Container Environment

Pang-Wei Tsai, Hong-Yu Wei, Yu-Chi Hsu (National Central University, Taiwan)

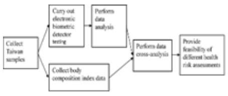
The virtualization technology has been widely used in computer system for a long period of time, it provides resource allocation, flexibility and efficiency by using hypervisor to manage virtual machine. Nevertheless, for lightweight workload, the container gains more advantages on transferability, rapid deployment, and robust ecosystem supports. Because of these characteristics, container is commonly used in edge computing and microservice to fulfill essential resource requirements. However, when service on the container receives too much requests without limitation, it may meet degrading effect which lowers down its performance and makes host network to be stuck. Hence, this research aims to focus on investigating this issue for finding better strategy to improve the connection performance of http-based service application with multithreaded load balancer.



©ICAROB 2023 ALife Robotics Corp. Ltd.

OS5-2 Electronic Biometric Detector and Body Composition Index in Predicting Disease Risk

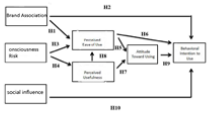
Wen-Fu Yang, Chung-Te Ting (Chang Jung Christian University, Taiwan)


Collect Taiwan samples
Carry out electronic biometric detector testing
Perform data analysis
Collect body composition index data
Perform data cross-analysis
Provide feasibility of different health risk assessments

According to the World Health Organization, COVID-19 has killed 14.9 million people worldwide by 2021. According to statistics from the World Obesity Alliance, if the domestic obese population exceeds 50% (three high diseases), the country's COVID-19 mortality rate will be 10 times higher. There seems to be a close relationship between obesity and the risk of hospitalization and treatment. In this study, the electronic biometric detector was used to collect the response values of the body cells of adults in Taiwan. Through the comparison and analysis of big data, the functional status of each organ system in the human body was calculated and compared with the body composition index. It was found that there is a high correlation between the two. The results of this study may provide feasibility of different health risk assessments.

OS5-3 Exploring Consumers' Intention to Use Mobile Payment APPs Based on Technology Acceptance Models - Taking Line Pay as an Example

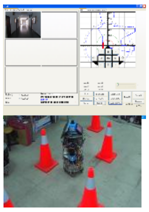
Chun -Nan Chen, Yu-Sheng Huang (Chang Jung Christian University, Taiwan)


Brand Association
onsciousness Risk
social influence
Perceived Ease of Use
Perceived Usefulness
Attitude Toward Using
Behavioral Intention to Use
H1
H2
H3
H4
H5
H6
H7
H8
H9
H10

According to the statistics of the Financial Supervisory Commission, as of the end of March 2021, the cumulative transaction amount of the latest five mobile payments has reached 503.2 billion, LINE, a communication software that is inseparable from our lives, has launched a service - "LINE Pay", which has become the most commonly used mobile payment by Taiwanese consumers. This research mainly takes people who have installed LINE PAY as the research object, and uses three factors that may affect consumers' consumption patterns, such as " Brand Association ", " Consciousness Risk " and "social impact", to explore consumers' usage intentions for mobile payment apps, hoping to gain a deeper understanding of their usage intentions.

OS5-4 Optimization of Robot Path and IoT Communication Path Based on Artificial Intelligence

Jr-Hung Guo, Kuo-Hsien Hsia (National Yunlin University of Science and Technology, Taiwan)



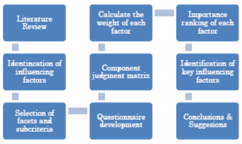
The Internet of Thin (IOT) and robotics are very popular research topics, and they have begun to enter people's daily life, and both robots and the Internet of Things have the problem of path optimization. For robots, Although the map can be established in advance and the robot can avoid obstacles, the robot's travel map is likely to change at any time, and a new path needs to be generated at this time. The Internet of Things will provide users or monitoring system information with the best transmission speed path. Based on the above requirements, this paper uses artificial intelligence to optimize the path between robots and the IoTs question. The method expected in this paper is to parameterize the length of each path, the number of obstacles, whether there will be collisions, etc., and then use the artificial intelligence algorithm of multi-tree and LSTM to find the best path.

©ICAROB 2023 ALife Robotics Corp. Ltd.

OS5-5 Key Success Factors Affecting Family Members' Intention to Withdraw from Life-sustaining Treatment for Long-term Ventilator-dependent patients: Nursing Professionals' Perspective

Hsiao-Fang Chen, Jin-Yuan Chern (Chang Jung Christian University, Taiwan)

Nursing professionals are major caregivers of VDPs. They have abundant experience in observing VDP family members' reflections when confronting LST withdrawal-or-not issues. Therefore, this study aimed to explore, from the perspective of nursing professionals, the key success factors which would significantly affect VDPs' family members' decision about withdrawal from LST. This is a quantitative survey research study. Nursing professionals from a government-affiliated region-level teaching hospital in southern Taiwan were potential participants. First, based on an action-research design, a structured questionnaire composed of three constructs (18 sub-constructs) were developed through a literature review. Second, senior nursing professionals with abundant experience in caring VDPs were recruited for data collection. Third, the analytic hierarchy process (AHP) was adopted to collect and analyze the participants' responses. Overall, "subjective norms" was considered relatively important than "behavior/attitude" and "behavior control". Further, "family consensus" was considered as the first priority, followed by "quality assurance" and "individual value". On the contrary, "grief counseling", "communication timing" and "ethics and legal" were with the least priorities. Family consensus and quality assurance are the two key factors when family members confront the decision whether to withdraw the LST for their beloved person.


Literature Review
Identincation of influencing factors
Selection of facets and subcriteria
Questionnaire development
Component judgment matrix
Calculate the weight of each factor
Importance ranking of each factor
Identification of key influencing factors
Conclusions & Suggesions

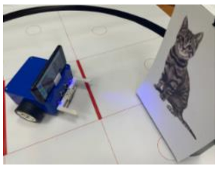
OS6 Modelling and Simulation in Robotics (8)

Chair Evgeni Magid (Kazan Federal University, Russia)
Co-Chair Kuo-Hsien Hsia (National Yunlin University of Science & Technology, Taiwan)

OS6-1 Android Based Educational Mobile Robot Design and Pilot Evaluation

Elvira Chebotareva, Maksim Mustafin (Kazan Federal University, Russia)

Educational robotics is a powerful tool for STEM and STEAM education. However, in practice, while trying to introduce educational robotics into a well-established educational process, a number of significant problems might arise. These include a high cost of an equipment and a lack of necessary competencies in the field of educational robotics among teachers. In this paper, we describe a concept of an inexpensive mobile robot equipped with a mobile device running Android operating system. We present sample projects that demonstrate how to extend capabilities of the educational mobile robot through the use of the mobile device. The proposed approach was preliminary evaluated in "Educational robotics" Bachelor level course and "Introduction to robotics" Master level course at Kazan Federal University.

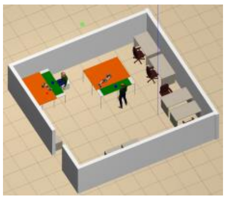


©ICAROB 2023 ALife Robotics Corp. Ltd.

OS6-2 Virtual Collaborative Cells Modelling for UR3 and UR5 Robots in Gazebo Simulator

Ramir Sultanov, Shifa Sulaiman, Tatyana Tsoy, Elvira Chebotareva
(Kazan Federal University, Russia)

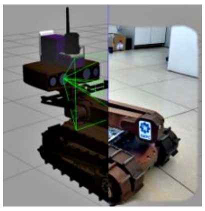
This paper presents virtual models of collaborative cells for two industrial collaborative robots UR3 and UR5 in the Gazebo simulator. Typically, the UR3 and UR5 robots are used by enterprises for packaging, assembly and sorting. Modeling and virtual experiments are an important stage in production processes planning, which involves joint human-robot work. Such models allow to plan safe human-robot interactions within a joint workspace and, if required, to rearrange the workspace. Our models of collaborative cells were adapted to several typical cases of joint human-robot operation scenarios and could be used in engineering design and testing for human-robot interaction in the field of production processes.



OS6-3 New Features Implementation for Servosila Engineer Model in Gazebo Simulator for ROS Noetic

Alexandra Dobrokvashina[1], Shifa Sulaiman[1], Timur Gamberov[1], Kuo-Hsien Hsia[2], Evgeni Magid[1]
([1]Kazan Federal University, Russia)
([2]National Yunlin University of Science & Technology, Taiwan)

Virtual experiments play an important part in robotics allowing to reproduce complex environments, perform complicated and risky tasks. Yet, a virtual model is not always a one-time build action and it requires revisions in a timely manner as operating systems and dependent software evolves. This article presents a number of technical updates of the Servosila Engineer crawler type robot virtual model. The model evolution necessity was caused by a migration from an outdated robot operating system (ROS) of Melodic version to the modern ROS Noetic version. In addition to migration issues, for the robot virtual model a new onboard torch control unit and a robot head aligning unit were developed.



OS6-4 Features of Interaction Between a Human and a Gestures-controlled Collaborative Robot in an Assembly Task: Pilot Experiments

Maksim Mustafin[1], Elvira Chebotareva[1], Hongbing Li[2], Martínez-García Edgar A[3], Evgeni Magid[1,4]
([1]Kazan Federal University, Russia) ([2]Shanghai Jiao Tong University, China)
([3]The Autonomous University of Ciudad Juarez, Mexico) ([4]HSE University, Russia)

This paper presents results of pilot experiments that were run to study a human interaction with the UR5e collaborative 6-axis robot manipulator in a cooperative assembly task. The participants controlled the equipped with a screwdriver UR5e robot using computer vision and gestures. The purpose of the experiments was to identify the features of user interaction with the UR5e robot controlled with gestures in a task of a complex object assembly. Ten people took part in the experiments. The results of the experiments allowed to conclude on practical efficiency of cobots in joint assembly tasks. In addition, we identified preferable by the users location areas during the assembly task.

©ICAROB 2023 ALife Robotics Corp. Ltd.

OS6-5 Modeling of Human Actions in a Collaborative Robotic Space Using AR601M Humanoid Robot: Pilot Experiments in the Gazebo Simulator

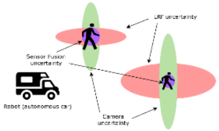
Aidar Zagirov, Artem Apurin , Elvira Chebotareva (Kazan Federal University, Russia)

To guarantee a safe human-robot collaboration, a collaborative system development requires a significant amount of real world experiments. Yet, it is critical to avoid injury risks for participants of such experiments. The risks could be reduced by introducing a virtual experiments' stage to detect mistakes in a robot behavior prior to the real world experiments. This paper presents a virtual model of a humanoid robot AR601M in the Gazebo simulator. Unlike the standard human models in the Gazebo, this model allows to simulate the gross and fine motor skills of a human and could be used when performing various human actions in collaborative robotic cells.

OS6-6 Modern Methods of Map Construction Using Optical Sensors Fusion

Ramil Safin[1], Tatyana Tsoy[1], Roman Lavrenov[1], Ilya Afanasyev[1], Evgeni Magid[1,2]
([1]Kazan Federal University, Russia) ([2]HSE University, Russia)

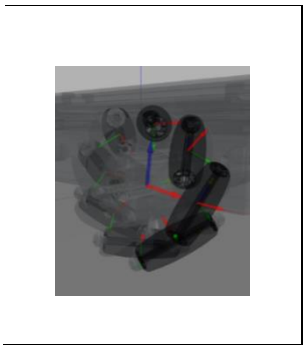
Map construction, or mapping, plays an important role in robotic applications. Mapping relies on inherently noisy sensor measurements to construct an accurate representation of a surrounding environment. Generally, individual sensors suffer from performance degradation issues under certain conditions in the environment. Sensor fusion allows to obtain statistically more accurate perception and to cope with performance degradation issues by combining data from multiple sensors of different modalities. This article reviews modern sensor fusion methods for map construction applications based on optical sensors, such as cameras and laser range finders. State-of-the-art mapping solutions built upon different mathematical theories and concepts, such as machine learning, are considered.



OS6-7 Omniwheel Chassis' Model and Plugin for Gazebo Simulator

Artem Apurin[1], Bulat Abbyasov[1], Alexandra Dobrokvashina[1], Yang Bai[2], Mikhail Svinin[2], Evgeni Magid[1]
([1]Kazan Federal University, Russia) ([2]Ritsumeikan University, Japan)

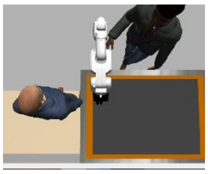
Increasing a mobility of a robot in a limited space or in a presence of a large number of people is an important task. Mecanum wheels could provide the required high flexibility locomotion in any direction. This article presents a virtual model of an omniwheel robot in the Gazebo simulator. Freely rotating rollers were implemented to simulate the robot motion. We developed a four-wheel mecanum mobile robot plugin controls the robot by publishing linear velocity data along X and Y axes, and angular velocity data along Z-axis. The plugin could optionally publish a standard ground-truth odometry of the Gazebo or a calculated in real time wheel odometry. The open source code is extendable for similarly structured platforms and is available for a download via GitLab.



OS6-8 Using Optical Sensors for Industrial Human-Robot Interaction in Gazebo Environment

Elaman Kidiraliev, Roman Lavrenov (Kazan Federal University, Russia)

This paper presents an overview of a robot operating system based architecture for human-industrial robot interactions using peripheral optical sensors for real-time object detection and collision avoidance with an industrial robot in the virtual world of the Gazebo simulator. Machine vision plays a huge role in production automation, and develop a system based on Kuka KR3 industrial robot for detecting and tracking a human and other objects in a working area using optical sensors. The ability to work in low light and crowded conditions, as well as the ability to reconstruct a method of execution of a task, while maintaining control of a robot. This work considers several optical sensors and a comparative analysis.



©ICAROB 2023 ALife Robotics Corp. Ltd.

OS7 Advanced Control Systems (7)
Chair Takuya Kinoshita (Hiroshima University, Japan)
Co-Chair Takao Sato (University of Hyogo, Japan)

OS7-1 Predictive Functional Controller Design with Disturbance Observer and Its Application

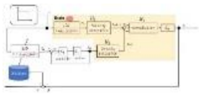
Syota Yoshida, Shin Wakitani (Hiroshima University)

This paper discusses predictive functional controller (PFC) design using a disturbance observer for a system with a dead time. The PFC is a kind of model predictive controller that is an effective method for systems with a dead time. However, in PFC, the control performance is strongly affected by the accuracy of the designed model. Therefore, this research proposes a method that the PFC maintains good control performance under the modeling error by using the disturbance observer that can estimate the disturbance even if a system has dead time. The effectiveness of the proposed control scheme is evaluated by experimental results.

OS7-2 Study on a Construction of Velocity Perception Model and Kansei Feedback Control System in Active Behavior

Sota Takemura, Takuya Kinioshita, Toru Yamamoto (Hiroshima University, Japan)

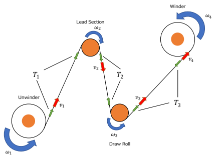
In Japanese, there is a word "Kansei" which means "feelings, impulses, and desires stimulated by the senses. In addition, there is a field called Kansei Engineering which is recognized worldwide. In the case of human-operated machines, it is necessary to consider Kansei such as whether the operability is comfortable. Therefore, this paper describes a model that focuses on Kansei and control based on the model. It has been proposed that a human Kansei model based on the Weber-Fechner law. However, the Weber-Fechner law is applied to stimulus such as sound, smell, and light. The model needs to be improved to reduce the error between the velocity assumed in the brain and the actual velocity of the machines. Therefore, this paper proposes a new Kansei model that focuses on the relationship between the actual velocity and the perceived velocity when a human actively operates the machines.



OS7-3 Design of a Database-Driven Control for a Web Conveyor

Atsushi Takatani, Takuya Kinioshita, Toru Yamamoto (Hiroshima University, Japan),
Tomohiro Hirakawa, Hiroki Hamamoto, Takashi Ochiiwa, Hideki Tomiyama (Japan Steel Works, Ltd.)

Web conveying equipment used in the processing of plastic films and other materials is a large-scale system consisting of multiple drive rolls, and it is difficult to understand the system characteristics during driving. This study proposes a database-driven control method based on the FRIT method to implement motor control of this equipment without relying on system identification. In order to reduce the computational load, the concept of database design based on similarity is incorporated. The database, which holds the operation results, enables the web conveyor's operation output to quickly reach the reference signal.



©ICAROB 2023 ALife Robotics Corp. Ltd.

OS7-4 Development of IoT self-tuning control device using Wi-Fi

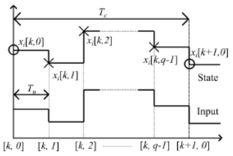
Shinichi Imai (Tokyo Gakugei University, Japan)

In this paper, development of IoT control device using Wi-Fi. In recent years, IoT has been attracting attention, and there are growing expectations in the industrial world for the utilization of data obtained from many sensors. These data are stored in databases in real time through communication between sensors and the cloud and communicating with the cloud. Meanwhile, digital controllers are widely used in the process industry as general-purpose controllers. However, it is difficult to incorporate AI, machine learning, and databases into general-purpose controllers due to data memory limitations. Therefore, in this paper, we develop an IoT self-tuning controller using Wi-Fi. As a result of experiments, the controller and computer were connected via Wi-Fi, self-tuning was performed on the computer side, and the calculated PID gains could be sent to the controller to achieve control.

OS7-5 Consensus Control for Dual-rate Multi-agent Systems

Takaya Tanaka, Natsuki Kawaguchi, Takao Sato (University of Hyogo, Japan)

The present study discusses the consensus control multi-agent systems. In such systems, the consensus is achieved through the exchange of information between neighboring agents. In order for mobile systems to consensus with each other, they must be designed with power consumption in mind. Therefore, since communication consumes power, it is important that the communication interval be as long as possible. The present study proposes a design methodology for the consensus achievement of dual-rate multi-agent systems, where the update period of agents is shorter than the communication period.



OS7-6 Data-driven Control Experiments of a Quadrotor Drone

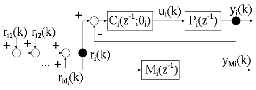
Tomonori Yao, Natsuki Kawaguchi, Takao Sato (University of Hyogo, Japan)

Multi-rotor unmanned aerial vehicle (UAV)have various advantages due to their high expected performance. Most conventional model-based design methods require the dynamic characteristics of UAV, whereas data-driven design methods allow the controller to be designed directly from flight data without a model. This study reports on the creation and flight experiments of a quadrotor drone with the aim of preparing an environment for implementing data-driven design for UAV. In addition, the usefulness of the controller directly designed from flight data is also reported.

OS7-7 Design of Data-driven Multi-agent Systems

Kenta Nagao, Natsuki Kawaguchi, Takao Sato (University of Hyogo, Japan)

This study discusses the consensus control of multi-agent systems. The consensus can be achieved when the closed-loop system of multi-agent systems is stable. In conventional model-based methods, since the controller is designed based on the dynamic characteristics of the agents, models of the agents must be used. On the other hand, this study examines data-driven design of multi-agent systems. In the proposed method, the controller of a multi-agent system is designed directly from the control data, where the controller structure is fixed. The usefulness of the proposed method is shown through numerical examples.



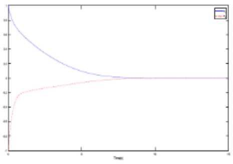
©ICAROB 2023 ALife Robotics Corp. Ltd.

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

OS8-1 A Self-triggering Control Based on Adaptive Dynamic Programming for Nonzero-sum Game Systems

Yibo Shi, Chaoli Wang (University of Shanghai for Science and Technology, China)

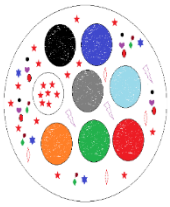
Recently, for the optimal control problem of nonzero-sum game systems, although it is discussed that these methods are event-triggered, it is still necessary to continuously monitor measurement errors during execution, which is difficult to achieve by hardware. In order to avoid continuous detection measurement errors, a selftriggered control based on adaptive dynamic programming is proposed to solve the optimal control problem for continuous-time nonlinear nonzero-sum game systems with unknown drift dynamics. Firstly, the principle of IRL method is used to avoid the requirement of system drift dynamics in the controller design. Then, to approximate the Nash equilibrium solution, a critic neural network is used to estimate the value function. Furthermore, a selftriggered adaptive control scheme is proposed according to Lyapunov theory to ensure the uniform ultimate boundedness (UUB) of the closed-loop system state. The self-triggered control obtained in this paper can calculate the next trigger point by the information of the current trigger moment.



OS8-2 Harmony of Agent System with Heterogeneity

Yunzhong Song[1], Weicun Zhang[2], Fengzhi Dai[3], Huimin Xiao[4], Shumin Fei[5], Jichao Zhao[6]
([1]Henan Polytechnic University, P.R.China) ([2]University of Science and Technology Beijing, P.R.China)
([3]Tianjin University of Science and Technology, P.R.China)
([4]Henan University of Economics and Law, P.R.China) ([5]South East University, P.R.China)
([6]Tianjin University of Science and Technology, P.R.China)

This paper focuses on the integration of heterogeneous agents, also known as harmony of them, where heterogeneity emphasized so often. To advance the idea of heterogeneous agents, harmony of the agents, where homogeneity of the agent system turns into heterogeneity one at start, and then the heterogeneity turns back to homogeneity again after. Initiated from hybrid order agents, heterogeneity like social status, encapsulated agents will be introduced step by step. Finally, Chinese philosophy, which always inspires new ideas of thought, can be a good source of research topic. Conclusion that encapsulated agent is the capital Tao of agent systems was drawn from the paper.

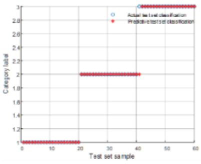


©ICAROB 2023 ALife Robotics Corp. Ltd.

OS8-3 Apple Grading based on IGWO Optimized Support Vector Machine

Yi Zhao[1], Qunpo Liu[1,2], Yuxi Zhao[1], Yueqin Sheng[1]
([1]Henan Polytechnic University, P.R.China)
([2]Henan International Joint Laboratory of Direct Drive and Control of Intelligent Equipment, P.R.China)

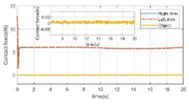
In order to improve the accuracy of apple external quality classification based on support vector machine, an improved grey wolf optimization algorithm IGWO was proposed by adding Logistic chaos mapping, nonlinear convergence factor and Cauchy variation to the grey wolf optimization algorithm. Firstly, different benchmark functions are used to test the improved IGWO algorithm. The test results show that the IGWO algorithm has improved the convergence speed and accuracy. Secondly, the image processing method is used to extract apple's external features as the data set. The improved grey wolf algorithm was used to optimize the penalty parameters and kernel parameters in support vector machine, and the optimal IGWO-SVM classification model was obtained. Finally, compared with the classification results of SVM and GMO-SVM, the results show that IGWO-SVM has the highest classification accuracy.



OS8-4 Cartesian Space Coordinated Impedance Control of Redundant Dual-Arm Robots

Yang Zhang, Yingmin Jia (Beihang University (BUAA), China)

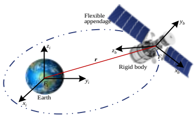
This paper presents a cartesian space coordinated impedance control method to achieve coordination when a dual-arm robot operates an object. First, the relative positional and force errors when the two arms operate the object are defined. Then, these relative errors are introduced into the general impedance controller to achieve coordinated impedance control. Compared to the conventional impedance control, this scheme ensures the coordination between the two arms and reduces the contact force error between the end-effectors and the object.



OS8-5 Disturbance Observer-based Anti-unwinding Control for Flexible Spacecrafts

Qian Sun[1], Yingmin Jia[1], Weicun Zhang[2]
([1]Beihang University (BUAA), China) ([2]University of Science and Technology Beijing, P.R.China)

The anti-unwinding control problem for the six-degrees-of-freedom (6-DOF) motion of the flexible spacecraft is studied in this paper. Firstly, the translation-rotation-vibration coupling motion of the flexible spacecraft is described by dual quaternion. Then, a nonlinear disturbance observer (NDO) is applied to estimate and compensate the lumped disturbances including the flexible vibration and unknown external disturbances. An anti-unwinding controller is designed based on the sliding mode technology. The stability of the closed-loop system is verified via Lyapunov method. Finally, numerical simulations indicate the effectiveness of the designed controller.

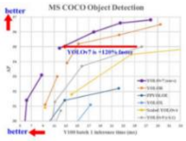


©ICAROB 2023 ALife Robotics Corp. Ltd.

OS9 Intelligent Systems and Robotics – 1 (4)
Chair Hucheng Wang (Tianjin University of Science and Technology, China)
Co-Chair Fengzhi Dai (Tianjin University of Science and Technology, China)

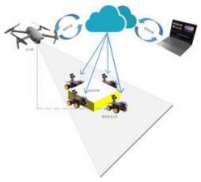
OS9-1 A Survey of Target Detection Based on Deep Learning
Hucheng Wang, Fengzhi Dai, Min Zhao (Tianjin University of Science and Technology, China)

Object detection is a hot topic in the field of visual detection. Deep learning can greatly compensate for the defect that traditional methods sacrifice real-time for improving accuracy. This paper mainly introduces the main networks and methods of two-stage deep learning algorithm and single-stage deep learning algorithm in the field of target detection. The advantages and disadvantages, usage scenarios and development of each network are described in detail. Finally, the follow-up development in this field is prospected.



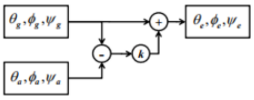
OS9-2 A Design of New Air Ground Cooperative Unmanned Transportation System
Hucheng Wang, Min Zhao (Tianjin University of Science and Technology, China)

With the development of science and technology, the shortage of human resources in the labor market and the increase of human costs, intelligence and unmanned logistics have become the mainstream development trend. This paper introduces a new type of space ground coordinated unmanned transportation system, which aims to solve the problems of liberating labor, reducing the operating costs of logistics enterprises, reducing the contact between people, and impeding the spread of the COVID-19 epidemic. This paper introduces the research background and advantages of the system, and describes the selection and working principle of the system in detail.



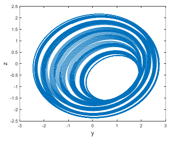
OS9-3 Attitude Solution of Quadrotor UAV
Siyuan Liu, Zhihao Zhao, Haoran Gong (Tianjin University of Science and Technology, China)

In this paper, the attitude solution of a quadcopter UAV is introduced, and the integrated circuit designed by the UAV is verified. The angular velocity data and acceleration data required for this solution are provided by two modules: gyroscope and accelerometer. In this attitude control, three values are used: quaternions, rotation matrices, and Euler angles. Through data algorithm analysis, integrated circuit design and program design, the feasibility of the attitude algorithm is proved. On this basis, the problem of combining two systems: attitude flight and fixed height and fixed point is solved.


θ_g, ϕ_g, ψ_g
θ_a, ϕ_a, ψ_a
-
k
+
θ_e, ϕ_e, ψ_e

OS9-4 A four-dimensional conservative chaotic system and its application in image encryption
Hongyan Jia, Wei Li, Jingwen Liu (Tianjin University of Science and Technology, Tianjin, China)

In this paper, based on numerical analysis, NIST test and FPGA implementation, a four-dimensional conservative chaotic system is firstly analyzed to investigate and find an applicable pseudo-random signal generator for image encryption algorithm. It is found that the four-dimensional conservative chaotic system shows some complex dynamics, such as multi-stability, strong pseudo-randomness, and physical characteristics. Finally, based on pseudo-random sequences and two-dimensional discrete wavelet transform, an image encryption algorithm is realized. Both the experiment results and the security analysis show that the algorithm show good encryption characteristics, which further prove the image encryption algorithm.



©ICAROB 2023 ALife Robotics Corp. Ltd.

OS10 Intelligent Systems and Robotics – 2 (4)

Chair Fangyan Li (Tianjin University of Science and Technology, China)
Co-Chair Yande Xiang (Tianjin University of Science and Technology, China)

OS10-1 A Research on Image Defogging Algorithm Based on Enhancement

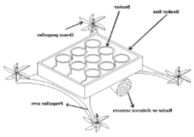
Fangyan Li, Haokang Wen, Chang Sheng, Min Zhao (Tianjin University of Science and Technology, China)

In order to solve the problem of low contrast image and loss of image details in the foggy weather, the image defogging technique is used to remove the noise in the image and improve the image contrast, so as to recover a clear and fog-free image. In this paper, we mainly introduce three image defogging algorithms: global histogram equalization, local histogram equalisation and the Retinex algorithm. The advantages and shortcomings of each algorithm are summarised through the study of the principles of each algorithm and the comparative analysis of the experimental result.

OS10-2 Autonomous Microcontroller-Based Aerial Water Sampling Device

Weifang Wang[1], Mingxia Kang[2], Ruming Kang[3]
([1]North Minzu University, China), ([2]Tianjin University of Commerce, China), ([3]Chongqing University, China)

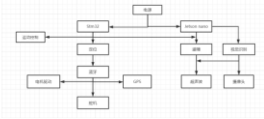
Water quality testing starts with sampling. Traditional methods may cause water quality properties to change due to the limitation of sampling methods. In addition, it is difficult and risky to take samples in parts of special terrain, such as ravines and swamps, where manpower is difficult to reach. To this end, we combine a single-chip machine STM32-based sampling device with a UAV that can span complex terrain to solve the above problems. The characteristics of water quality without human intervention and equipment automation are of great significance.



OS10-3 Intelligent Electronic Guide Dog

Mengyu Liu, Yande Xiang, Zhi Qiao, Tao Zhu (Tianjin University of Science and Technology, China)

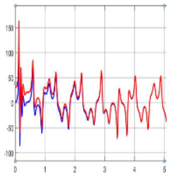
An intelligent mobile blind guide robot is designed in this paper. The robot is equipped with a variety of sensors and combined with the powerful computing power of the main control board to form a complete obstacle avoidance system to achieve autonomous navigation and obstacle avoidance in a complex environment. It controls GPS positioning through Jetson-Nano, plans the route, detects obstacles with the help of cameras and ultrasonic sensors, and gives users feedback through the alarm sound of the buzzer.



OS10-4 Research on Chaos Synchronization of Qi System and Lü System with Different Structures

Hanyuan Wang, Yiting Gao (Tianjin University of Science and Technology, China)

his paper introduces a three-dimensional chaotic synchronization method is introduced, and the advantages and disadvantages of the synchronization controller designed by this method are analyzed. Firstly, the characteristics of two chaotic systems with different structures are studied. Secondly, the mathematical model is established, and the synchronization controller is designed by direct method, so that the two chaotic systems with different structures can be synchronized at different initial values. With the help of MATLAB, the error curve of the synchronous system is drawn when the synchronous controller acts on the response system.



©ICAROB 2023 ALife Robotics Corp. Ltd.

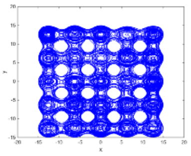
OS11 Intelligent Systems and Robotics – 3 (4)

Chair Jiaxin Li (Tianjin University of Science and Technology, China)
Co-Chair Fengzhi Dai (Tianjin University of Science and Technology, China)

OS11-1 Generation and Analysis of a Multi-scroll Conservative Chaotic System

Jiaxin Li, Yong Liu, Min Zhao (Tianjin University of Science and Technology, China)

Chaos is one of the hot subjects in recent years. With the further study of chaotic systems, more and more chaotic systems have been found. Multi volume conservative chaotic system is a new kind of chaotic system, which has attracted extensive attention due to its complex dynamic characteristics. In this paper, we study a conservative chaotic system and introduce a sine function without multiple angles to make the conservative chaotic system generate multiple volumes, so as to construct a multiple volume conservative chaotic system. The system generates one-dimensional linear multi roll and two-dimensional grid like multi roll distributions by adjusting nonlinear functions.



OS11-2 A Design of Fire Detection Device Based on YOLOv5

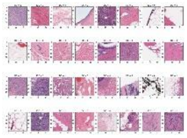
Zhiyang Li, Yande Xiang, Haoyu Guo, Yu Chen, Wenxuan Pan, Fengzhi Dai
(Tianjin University of Science and Technology, China)

In recent years, fires have become more and more frequent, which has a great impact on people's production and life and even their lives. This paper designs a fire detection device based on Yolov5, which is mainly composed of Raspberry Pi, OpenCV and buzzer. It can be widely used in narrow corridors, parking lots, shopping malls, forests and other scenarios. The device has the characteristics of high recognition rate, fast recognition speed and strong sensitivity, and has excellent recognition effect in fire detection.

OS11-3 Application of Convolutional Neural Network in Accurate Breast Cancer Identification

Zhiyang Li, Haoyu Guo, Yande Xiang, Wentao Kuang, Lu Chen
(Tianjin University of Science and Technology, China)

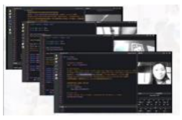
As one of the most common cancers in women, breast cancer has the highest incidence in the world. Nearly 600,000 people die from breast cancer each year, and early detection is essential for breast cancer treatment. In recent years, the rapid development of artificial intelligence has provided unprecedented ideas for the precise diagnosis and treatment of breast cancer. In this paper, the practical application of artificial intelligence convolutional neural network in breast cancer recognition is studied, which greatly improves the detection speed and saves a lot of time for doctors to further judge the condition.



OS11-4 Intelligent Infusion Service Based on Open MV

Haoran Gong, Zongyi Li, Qi Chu, Siyuan Liu, Feiyang Qu, Lu Wang
(Tianjin University of Science and Technology, China)

This paper presents an intelligent infusion service system designed to reduce the workload of doctors and nurses and for the health of patients. The system is applied to the drip stand to create a "new type of drip stand" that will effectively solve a number of problems. It is dedicated to: "intelligent infusion process, improving the efficiency of health care workers and reducing their work stress and psychological burden". It embodies the concept of economy and structural rationality.



©ICAROB 2023 ALife Robotics Corp. Ltd.

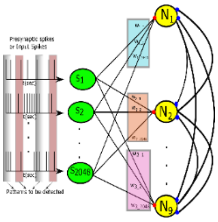
OS12 Theory and Implementation of Neuromimetic Systems (2)

Chair Takashi Kohno (University of Tokyo, Japan)
Co-Chair Takuya Nanami (University of Tokyo, Japan)

OS12-1 Adaptive STDP Learning with Lateral Inhibition for Neuromorphic Systems

Ashish Gautam, Takashi Kohno (University of Tokyo)

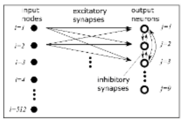
Implementing biologically plausible learning rules on neuromorphic chips is essential to explore the learning mechanisms in the brain. Spike-timing dependent plasticity (STDP) is one such rule but its multi-bit circuit implementation requires too much area. In our previous study, we proposed a bioinspired hardware-friendly learning rule named adaptive STDP and experimentally showed that its performance was similar to STDP learning in a very basic biologically plausible spike pattern detection task using a single neuron. In this study, we extend the adaptive STDP learning rule with lateral inhibition, a common motif observed in the brain, and solve a competitive spike pattern detection task with multiple neurons that compete to detect multiple patterns. Our results show that the performance is similar to STDP learning.



OS12-2 Spike pattern detection with close-to-biology spiking neuronal network

Takuya Nanami, Takashi Kohno (University of Tokyo)

The nervous system contains a variety of different types of neurons, each with different electrophysiological properties. However, their roles in information processing are poorly understood. Using the piecewise quadratic neuron (PQN) model that can reproduce a variety of electrophysiological properties, we demonstrate that performance on a biologically plausible task of spike pattern detection varies depending on the electrophysiological properties.



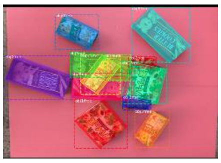
OS13 Industrial Artificial Intelligence Robotics (9)

Chair Eiji Hayashi (Kyushu Institute of Technology, Japan)
Co-Chair Sakmongkon Chumkamon (Kyushu Institute of Technology, Japan)

OS13-1 Object Status Detection in clutter environment for robot grasping using Mask R-CNN

Kasman, Eiji Hayashi (Kyushu Institute of Technology, Indonesia, Japan)

Detecting object status in cluttered manipulator's robot environment before grasping is quite challenging to recognize the target because of unstructured and uncertainty scenes. Using Mask R-CNN for detecting the status of the object i.e. free for picking, close, overlapping and piling up to the other objects is very useful as computer vision before the manipulator doing next procedures to complete its task. This paper provides a systematic summary and analysis target detecting and recognizing object status using Mask R-CNN. Unlike related solution methods that use machine vision and deep learning directly and combine together for doing robot controlling, pushing and grasping, we are doing image processing separately and simply for detecting the object's status and location before performing like pushing the object for making free and easy grasping. Experiment with this method shows that it has good accuracy and can be implemented like as input control reinforcement learning in advance.

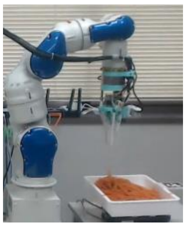


©ICAROB 2023 ALife Robotics Corp. Ltd.

OS13-2 Deep-Learning-Based Designed Weight Picking Noodle-like Object

Nattapat Koomklang, Sakmongkon Chumkamon, Prem Gamolped, Tomofumi Tsuji, Eiji Hayashi
(Kyushu Institute of Technology, Japan)

For food packaging line, manual picking up of the noodle-like objects according to specific weight requires worker's experience for picking quickly and accurately. This article presents a robot arm with 6-fingurs gripper picking up the noodle-like objects in specific weight using deep-learning-based to find the best possible action. For measuring the action, we use direct variation to probability of picking action at specific weight in this research use value for the likelihood of weight probability given an action. To find likelihood of weight probably deep-learning-based and use normal distribution for model distribution of the systems. For evaluation we passed any possible action to the network and find action that get maximum likelihood.



OS13-3 Research on grasping of string foods in the home meal replacement industry

Akihiro Ooya, Sakmongkon Chumkamon, Prem Gamolped, Tomofumi Tsuji, Eiji Hayashi
(Kyushu Institute of Technology, Japan)
Abbe Mowshowitz (The City College of New York, USA)

In recent years, automation by industrial robots has been desired in Japanese food manufacturing plants. This paper describes the development of an autonomous robot for automating the preparation of home meal replacement. The serving of lunchtime meals includes not only solid foods such as rice balls, but also string foods such as spaghetti. Unlike solidified foods, string foods require quantitative grasping. However, in the grasping experiments of string foods, the spaghetti deteriorates with time, and thus, a problem arises where an accurate quantitative grasping experiment cannot be performed. Therefore, in this study, we perform a quantitative grasping experiment by deep reinforcement learning using a material like string foods to verify the grasping and serving system for string foods.

OS13-4 Development of Drifting Debris Detection System using Deep Learning on Coastal Cleanup

Shintaro Ogawa, Sackmongkon Chumkamon, Eiji Hayashi (Kyushu Institute of Technology, Japan),
Ayumu Tominaga (National Institute of Technology, Kitakyushu College, Japan)

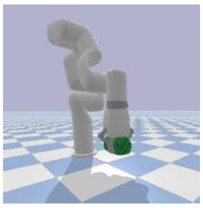
To solve the problem of litter drifting ashore, this study developed a litter detection system using deep learning. I used HTC (Hybrid Task Cascade) as a deep learning network. I also used Mask R-CNN for comparison. HTC is a model for instance segmentation that combines Cascade R-CNN and Mask R-CNN. Mask R-CNN is a model that adds segmentation capabilities to Faster R-CNN, a model for object detection. The dataset is a combination of a public dataset of general garbage called TACO and images taken at actual cleanup sites such as Hokuto Mizukumi Park. The discriminator was trained 2000 times with a learning rate of 0.002, and multi-class cross-entropy was used as the loss function.

©ICAROB 2023 ALife Robotics Corp. Ltd.

OS13-5 Soft Object Dexterous Manipulation Using Deep Reinforcement Learning

Sornsiri Promma, Sakmongkon Chumkamon, Eiji Hayashi (Kyushu Institute of Technology, Japan)

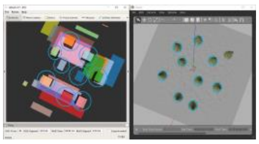
Manipulation of objects is one of the basic tasks that has been studied for a long time in the robotic field. Many experiments were setting the environment and using Deep Reinforcement Learning to train robot arm to grasp various objects. Still, most of those objects are solid objects, whereas nowadays, robot arms are used for grasping soft objects as well. In this study, we develop object manipulation tasks in pybullet simulation, focusing on soft objects by using Deep Reinforcement Learning (DRL) based on Soft Actor-Critic and Proximal Policy Optimization algorithm which aims to make the robot able to grasp soft objects in the exact position with proper force that does not damage them.



OS13-6 Practical Implementation of FastSLAM for Forestry Robot

Sylvain Geiser[1], Sakmongkon Chumkamon[1], Ayumu Tominaga[2], Takumi Tomokawa[1], Eiji Hayashi[1]
([1]Kyushu Institute of Technology, Japan), ([2]National Institute of Technology, Kitakyushu College, Japan)

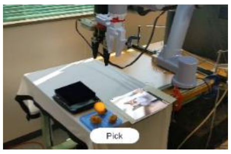
As the Japanese forestry workforce is shrinking, field robots are gaining interest in performing dangerous tasks in this environment. This paper presents research conducted on the SOMA robot designed at Hayashi Laboratory for this purpose. It focuses on theoretical and practical issues encountered through the implementation of the particle filter based FastSLAM algorithm on this mobile robot. In particular, the determination of the positions of trees from the raw pointcloud of the lidar, the side effects occurring at the boundary of the lidar visibility scope, and the modelling of motion and observation noises are discussed.



OS13-7 Research on AR system for industrial robot introduction

Takuya Matsumoto[1], Eiji Hayashi[1], Sakmongkon Chumkamon[1], Tomofumi Tsuji[1], Ayumu Tominaga[2], Abbe Mowshowitz[3]
([1]Kyushu Institute of Technology, Japan), ([2]National Institute of Technology Kitakyushu, Japan), ([3]The City College of New York, USA)

In recent years, labor shortages in small and medium-sized enterprises (SMEs) have become a serious issue, and the demand for automation by robots is increasing. Another challenge is the high cost of introducing industrial robots. In order to reduce the introduction cost, we are developing an AR system with the aim of providing robot introduction support to SMEs. The AR system developed enables the display of the robot's movements on the AR screen of a smartphone by communicating with ROS while developing Google's ARCore in Unity.In this paper, we describe the system configuration and evaluate the application of an AR application that enables AR simulation and intuitive GUI operation to check the safety range at the time of robot introduction.



©ICAROB 2023 ALife Robotics Corp. Ltd.

OS13-8 The BCRobo dataset for Robotic Vision and Autonomous Path Planning in Outdoor Beach Environment

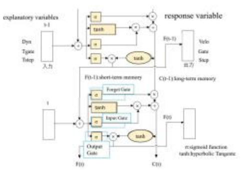
Tan Chi Jie[1], Takumi Tomokawa[1], Sylvain Geiser[1], Shintaro Ogawa[1], Ayumu Tominaga[2],
Sakmongkon Chumkamon[1], Eiji Hayashi[1]
([1]Kyusyu Institute of Technology, [2]National Institute of Technology (Kitakyushu College), Japan)

Along with the universalization of autonomous driving and image segmentation, various datasets are available freely for anyone to use to train their own neural network which speeds up the growth of deep learning technology. However, most of the datasets target only urban environments and other offroad environments are still lacking in datasets. This paper presents a beach environment dataset, BCRobo with the aim to contribute to closing the gap of robotic visual perception in offroad environment especially in beach. The dataset is also evaluated with two state-of-art image segmentation techniques to demonstrate the practical usage of the dataset.

OS13-9 The research about editing system of performance information for player piano. -Develop inference methods using machine learning -

Takaaki Ueno, Sakmongkon Chumkamon, Eiji Hayashi (Kyushu Institute of Technology, Japan)

In order for an automatic piano to perform like a human, it is necessary to have data with intonation for each note. However, existing automatic pianos do not have the ability to infer inflected data. For this reason, data inference has been performed manually until now. Therefore, the data We therefore attempted to construct a system that can infer data automatically by using machine learning. In this paper, we describe the machine learning system we have developed to infer performance information. This paper describes the actual inference system we developed. The data to be inferred was a performance of "Prelude Op.28 No.15" by Fryderyk Franciszek Chopin, composed by the world-famous pianist Vladimir Davidovich Ashkenazy. The performance is by the world famous pianist "Vladimir Davidovich Ashkenazy



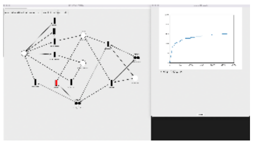
OS14 Software Development Support Method (4)

Chair Tetsuro Katayama (University of Miyazaki, Japan)
Co-Chair Tomohiko Takagi (Kagawa University, Japan)

OS14-1 Automated Random Simulation for Checking a Behavioral Model of Systems Based on Extended Place/Transition Net with Attributed Tokens

Sho Matsumoto[1], Tetsuro Katayama[2], Tomohiko Takagi[1]
([1] Kagawa University, Japan), ([2] University of Miyazaki, Japan)

Extended Place/transition Net with Attributed Tokens (EPNAT) is one of formal modeling languages, and it enables system engineers to construct an executable and abstracted behavioral model of multiple software systems. In this paper, we propose an automated random simulation technique of an EPNAT model in order to detect failures in the model. In the simulation, input is randomly selected for model execution. When a constraint given for each system or multiple systems is violated through the model execution, a failure is revealed. The simulation is terminated by the detection of a failure or the satisfaction of a criterion focusing on the combination of marking, data writing and reading between different systems. A prototype tool of the simulation technique was developed and applied to a trial model to discuss its effectiveness.

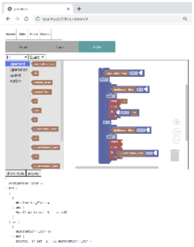


©ICAROB 2023 ALife Robotics Corp. Ltd.

OS14-2 Training of Software Formal Modeling Using Visual Blocks for Actions and Guards of Extended Place/Transition Net

Akio Usuda, Ryoichi Ishigami, Tomohiko Takagi (Kagawa University, Japan)

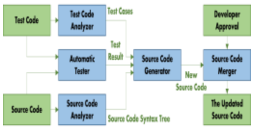
Extended Place/transition Net (EPN) that is one of software formal modeling languages consists of the parts of PN and VDM++. The former part represents abstracted state transitions of software. On the other hand, the latter part represents actions and guards on the transitions, and requires skills of system engineers who construct EPN models. We propose an extended training technique of EPN-based modeling using visual blocks for the VDM++ part. The visual blocks implemented by Blockly will be useful to accelerate the trainees' understanding of syntactical aspects of VDM++, and thus they are introduced into each step of the existing training technique for EPN. The effectiveness of the proposed technique is discussed through a preliminary experiment using our prototype tool.



OS14-3 Proposal of a Framework to Improve the Efficiency of the Implementation Step in Test Driven Development (TDD)

Takeaki Miyashita[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)

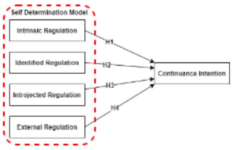
Test Driven Development (TDD) has three steps: test design, minimal implementation that passes the test, and refactoring. This research proposes a framework to support the implementation step in TDD to improve the efficiency of the implementation. The proposed framework firstly receives from the developer the test code and the source code that the test code does not fully pass. Next, It automatically generates new source code by modifying the given source code that the given test code can fully pass. And then, it shows the developer the differences between the given source code and the generated source code. Finally, it updates the source code if the developer approves the modification. By using this framework, developers can reduce the time required for the implementation step in TDD.



OS14-4 Continuance Intention Factor of Online Learning Management System in Case on Faculty of Computer Science at Brawijaya University in Indonesia

Mochamad Chandra Saputra[1], Tetsuro Katayama[2], Yoshihiro Kita[3], Hisaaki Yamaba[2], Kentaro Aburada[2], Naonobu Okazaki[2]
([1]Brawijaya University, Indonesia), ([2]University of Miyazaki, Japan), [3]University of Nagasaki, Japan)

The outbreak of the COVID-19 pandemic changed the model of the learning process. Online learning became one of the best solutions for many educational institutions, including the University, especially the Faculty of Computer Science, Brawijaya University. The continuance intention factor for using an online learning management system is important to ensure learning sustainability. To understand students' continuance intention this study proposes self determination model as a research model to find the factors affecting students' continuance intention toward online learning management systems. The proposed model is useful for investigating continuance intention factors. The study finds that intrinsic regulation, external regulation, identified regulation, and introjected regulation significantly positive impact on continuance intention in learning management systems. Finally, this study provides suggestions for the Faculty to improve the continuance intention of the student in using an online learning management system.



©ICAROB 2023 ALife Robotics Corp. Ltd.

OS15 Robot Control (5)

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

OS15 -1 Design of Intelligent Crutch System Based on STM32 and Raspberry Pie

Zongxuan Zhang, Jianhao Hu, Yizhe Sun, Yizhun Peng (Tianjin University of Science and Technology, China)

In order to strengthen the development of China's elderly care industry, the project has designed an intelligent crutch based on stm32 and raspberry pie. This product adds many functions on the basis of general intelligent crutches. It is equipped with real-time GPS monitoring and uploading APP. When the elderly fall, they can immediately and automatically alarm. When they fall, they can simultaneously emit an alarm sound of about 100 decibels to the surrounding. It provides a touch screen, supports voice recognition, and can achieve multiple functions. It is equipped with a step counting function, temperature and humidity display and other functions. The product uses 4G modules to communicate with the APP.

OS15-2 Design of Intelligent Fish Box Based on Machine Vision and Internet of Things Technology

Suqing Duan, Jiangyu Wu, Shuai Chen, Yizhun Peng (Tianjin University of Science and Technology, China)

The intelligent fish box based on machine vision and Internet of Things includes many fields such as machine vision, Internet of Things technology, single-chip microcomputer control and so on. It is an intelligent system that receives user instructions or voice control through the Internet to achieve a series of operations. At the same time, there is a temperature sensor in the fish box, which can transmit the temperature to the mobile phone APP in real time, and intelligent control the water temperature in the fish box. In addition, the fish box granaries send information to the user when the fish food is insufficient, and the camera in the fish box can monitor the situation of the fish box in real time, when a fish belly notifies the user to deal with.

OS 15-3 Design and Implementation of Internet of Things Planting System Based on esp32 MCU

Cuiying Ji, Yizhun Peng (Tianjin University of Science and Technology, China)

Agriculture is the country's fundamental industry, but there is still a lot of hard work because of climate pests. It is the development direction of agriculture to integrate traditional agriculture into modern Internet of Things technology and realize visualization and digital management of agricultural production.

This study combined modern information technology and used esp32 as the master to design and produce a plant planting system model for the Internet of Things.This research product uses a variety of sensors, Monitor the essential elements required for crop growth, including temperature, humidity, nutrient solution, soil moisture, etc.In the process of crop growth, automatic irrigation, nutrient solution supplement, ventilation cooling, light and other operations, greatly reduce the intensity of manual labor.

©ICAROB 2023 ALife Robotics Corp. Ltd.

OS15 -4 A Customized Dispensing Robot Based on OpenMV Visual Recognition

Hongze Liu, Yizhun Peng (Tianjin University of Science and Technology, China)

Dispensing medicine in hospital wards is a mechanized process, but it is difficult to develop on a large scale due to the uncertainty of the environment. In this study, a customized dispenser robot was designed to automatically deliver drugs to the designated ward according to the given instructions. The Visual identity module (OpenMV) is designed to collect images from the progress and then compare them with pre-stored images to achieve automatic pathfinding. Among them, the tracking part uses PID open-loop operation to increase the accuracy of movement. The ward number enables the camera to recognize the number in different scenes for many times, which increases the accuracy of recognition. The unique structure of this study provides an innovative and effective way to customize ward dispensing.

OS15-5 An Intelligent Guide Hat Based on The Internet of Things

Suqing Duan, Yizhun Peng (Tianjin University of Science and Technology, China)

With STM32 as the control core, the new guide cap has ranging obstacle avoidance, ranging obstacle avoidance, intelligent recognition, voice interaction, GPS positioning and other functions. The system uses OpenMV4 Cam H7 Plus intelligent camera for continuous real-time monitoring and effective feedback of traffic lights, zebra crossings and other important road information; Assisted by ultrasonic sensor, the electrical signal is converted into ultrasonic output through the transmitter to effectively measure the distance of obstacles; At the same time, the user's location information is sent to the surrounding vehicles and passers-by in real time through the Wi-Fi module and GPS positioning, ensuring the user's safety to the greatest extent.

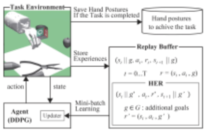
OS16 Robotic Manipulation (3)

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

OS16-1 Acquisition of Synergy for Low-dimensional Control of Multi-fingered Hands by Reinforcement Learning

Kazuki Higashi, Tomohiro Motoda, Akiyoshi Hara, Kensuke Harada (Osaka University, Japan)

We propose a reinforcement learning platform to learn to perform various tasks with a robotic hand to acquire a synergy. The model of deep reinforcement learning is trained to grasp an object with a multi-fingered hand. The synergy space is calculated by principal component analysis of hand postures when the task is successfully executed. The reward system is designed to minimize the distance of orthogonal projection between the posture and the synergy space, and the synergy space is acquired simultaneously with reinforcement learning.


Task Environment
Save Hand Postures If the Task is completed
Hand postures to achive the task
Store Experiences
Replay Buffer
$(s_t \| g, a_t, r_t, s_{t+1} \| g)$
$t = 0...T$
$r = (s_t, a_t, g)$
HER
$(s_t \| g', a_t, r', s_{t+1} \| g')$
$g \in G$: additional goals
$r' = (s_t, a_t, g')$
action
state
Mini-batch Learning
Agent (DDPG)
Updater

©ICAROB 2023 ALife Robotics Corp. Ltd.

OS16-2 Error Recovery Techniques Focused on Revival Process from Failures in Robotic Manufacturing Plants

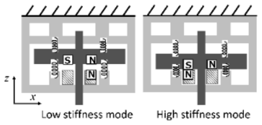
Akira Nakamura[1], Kensuke Harada[2] ([1]Saitama Institute of Technology, [2]Osaka University, Japan)

In recent years, working robots have been used in various fields from manufacturing industries to human living spaces. Therefore, more and more difficult tasks are performed by robots, necessitating the introduction of error recovery techniques. Our proposed error recovery technique is based on a method of going back to the process before the step in which the failure occurred and starting over from there. Of course, in practice, not only such a backward recovery but also a forward recovery that moves forward even after a failure occurs is used. This paper considers the various paths from failure occurrence to recovery execution and further discusses which path should be selected.

OS16-3 Flexible assembly system with stiffness switching joint

He Maike, Tokuo Tsuji, Naoki Ichikawa, Takuro Sawada, Tatsuhiro Hiramitsu, Hiroaki Seki
(Kanazawa University, Japan)

In this research, we aim to construct a flexible assembly system and realize precise assembly work by using a flexible-rigidity switchable joint for the wrist part of an industrial robot. We install magnets in the joints and change the structure of the joints by changing the state of attraction and repulsion of the magnets. The joint has two states, which are low stiffness and high stiffness.



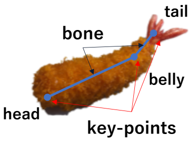
OS17 Artificial Intelligence for Embedded Systems and Robotics (6)

Chair Hakaru Tamukoh (Kyushu Institute of Technology, Japan)
Co-Chair Yuma Yoshimoto (National Institute of Technology, Kitakyushu College, Japan)

OS17-1 Pose Detection for Flexible-Indefinite Objects using Pseudo-Bone Data

Yuma Yoshimoto[1], Hakaru Tamukoh[2]
([1]National Institute of Technology, Kitakyushu College, Japan), ([2]Kyushu Institute of Technology, Japan)

This paper proposes a method for recognizing the poses of a flexible-indefinite object. Some flexible-indefinite objects, such as fried shrimp, differ between individuals. Therefore, it is difficult to estimate the pose of these foods by point cloud fitting or other methods. We propose "pseudo-skeletal" data for these objects. Pseudo-skeletal data consists of "key-points," which are joints, and "bones," which connect between key-points. For example, fried shrimp are given 3 key-points; "head," "belly," and "tail." In addition, the bones that connect them are given. In the experiment, a objects pose recognition model based on the human pose recognition model trains pseudo-skeletal data of fried shrimp. We confirmed that the model estimates the poses in the images

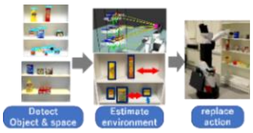


©ICAROB 2023 ALife Robotics Corp. Ltd.

OS17-2 Object Search and Empty Space Detection System for Home Service Robot

Tomoya Shiba[1], Tomohiro Ono[1,2], Hakaru Tamukoh[1]
([1]Kyushu Institute of Technology, Japan), ([2]JSPS Research Fellow, Japan)

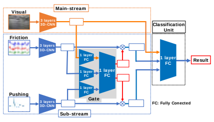
Home service robot has a pickup task to grasp and transport objects. When performing a pickup task, robots need to process to search for objects in the case of missing target objects. We propose approaches for the robot to search for objects from a shelf and a method for selecting empty spaces to move off-target objects. Proposed method is based on object recognition of graspable objects and shelf recognition model to select replacement locations and planned motions. The object to be moved and replace location are selected based on the size and position of empty spaces. In the experiments, the robot planned actions to search for the target object from a shelf and to replace the object near a group of similar objects. We used these in RoboCup@Home competition to evaluate the effectiveness of the proposed method.


Detect Object & space
Estimate environment
replace action

OS17-3 Robust Classification Model with Multimodal Learning for Home Service Robots

Ikuya Matsumoto, Daiju Kanaoka, Hakaru Tamukoh (Kyushu Institute of Technology, Japan)

We propose an auxiliary data stream structure as a robust classification model as shown in the figure. The model treats one modal as a main input and other modals as support. We experimented with two and three modal inputs. Moreover, we added pseudo shadows to visual information for the experiment of three modal inputs. In all experiments, our proposed model improves accuracy and robustness to environmental disturbances by using multiple modals. In future works, we will attempt to implement and evaluate the proposed method on a home service robot.


Main-stream
Visual
3 layers 3D-CNN
Classification Unit
Friction
3 layers 3D-CNN
1 layer FC
1 layer FC
1 layer FC
1 layer FC
1 layer FC
Result
Pushing
3 layers 3D-CNN
Gate
FC: Fully Conected
Sub-stream

OS17-4 Flexible Human-Robot Interaction in Domestic Environment Using Semantic Map

Yuga Yano[1], Yukiya Fukuda[1], Tomohiro Ono[1,2], Hakaru Tamukoh[1]
([1]Kyushu Institute of Technology, Japan), ([2]JSPS Research Fellow, Japan)

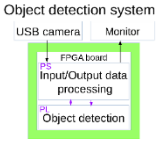
We propose an efficient semantic map to realize flexible human-robot interaction (HRI) in domestic environments. Our proposed map is created from an output of Simultaneous Localization and Mapping and already-known environmental information such as furniture and room. In this study, we evaluated the effectiveness of our proposed method on two benchmark tests for HRI in RoboCup@Home held in Bangkok in 2022. In the RoboCup@Home, we employ 3D human recognition to apply our proposed map to HRI, such as "find and offer an empty seat." We had the best score of all teams on both tests. The results of our experiments are available at https://youtube.com/playlist?list=PLfbN50Mwh2DG3OPDeCHo4TNuyrU4qYCrJ

3D human recognition
Left chair
Add furniture information

OS17-5 Impact of PS Load on FPGA Object Detection System Performance

Yusuke Watanabe[1,2], Hakaru Tamukoh[2]
([1]CRAFTWORK Co. Ltd, Japan,), ([2]Kyushu Institute of Technology, Japan)

A field-programmable gate array (FPGA) device which has Zynq architecture recently became popular. It is featured by inclusion of both processing system (PS) and programmable logic (PL) to a single chip. In a system using FPGA, while we tend to focus on the performance of PL, we can not ignore PS load completely. In this paper, using our object detection system which works on a Zynq FPGA board, we explore how our FPGA object detection system performance changes depending on PS load and report experiment results.

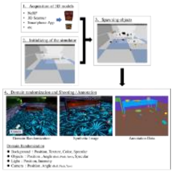

Object detection system
USB camera
Monitor
FPGA board
PS
Input/Output data processing
PL
Object detection

©ICAROB 2023 ALife Robotics Corp. Ltd.

OS17-6 An Effective Method for Minimizing Domain Gap in Sim2Real Object Recognition Using Domain Randomization

Tomohiro Ono[1,2], Akihiro Suzuki[1], Hakaru Tamukoh[1]
([1]Kyushu Institute of Technology, Japan), ([2]JSPS Research Fellow, Japan)

Data-centric deep learning has attracted considerable research attention in an era where data is at the core of every decision-making process. Manual annotation is a common practice, but it is very expensive and problems such as oversight and mislabeling caused by human error occur. These problems are known to affect the quality of the datasets significantly. To resolve these problems, in this study, we propose a method to automatically generate high-quality and large datasets in a short time using a simulator. Our proposed method aims to minimize the domain gap using domain randomization without faithfully reproducing real scenes in the simulator. The generated data were trained on You Only Look Once v7 and achieves an accuracy of more than 80% against the real data, demonstrating the feasibility of Sim2Real.



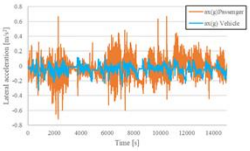
OS18 Mathematical Informatics (7)

Chair Takao Ito (Hiroshima University, Japan)
Co-Chair Makoto Sakamoto (University of Miyazaki, Japan)

OS18-1 A Fundamental Study on Car Sickness Using Data Science

Tsutomu Ito[1], Seigo Matsuno[1], Makoto Sakamoto[2], Satoshi Ikeda[2], Takao Ito[3]
([1]NIT, Ube College, Japan), ([2]University of Miyazaki, Japan), ([3]Hiroshima University, Japan)

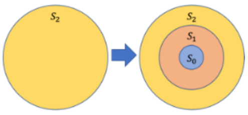
Car sickness often occurs with dizziness and discomfort accompanied by vomiting and headaches lasting several days in severely affected person. Car sickness has been studied from various standpoints on countermeasures and onset mechanisms, but a general measure has not been established yet. In this study, all dataset of the motion that occurs in the car and the head of the person sitting and/or driving in the car are collected based on data science. By validating the data, the characteristics of person who has experience of car sickness are tested. This study proposes a new measure aiming at development of motion sickness countermeasures that do not depend on car performance and find that the quick rotational motion of the head generated in the car could be considered as one of the factors that cause car sickness.



OS18-2 An Analysis of Quoridor by reusing the results of reduced version

Satoshi Ikeda[1], Takuro Iwanaga[1], Makoto Sakamoto[1], Takao Ito[2]
([1]University of Miyazaki, Japan), ([2]Hiroshima University, Japan)

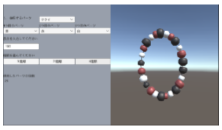
This paper presents a study of retrograde analysis using the board game " Quoridor " released by Gigamic Games. The retrograde analysis is performed by enumerating all the possible game phases. This leads to a huge search area and a problem of spatial computational complexity. In this study, we attempted to reduce the number of stations to be enumerated by reusing the results of the reduced version; when enumerating the S_2 stations, only the S_0 and S_1 stations can be omitted. This solves the Quoridor under the condition of having two fences each other.



©ICAROB 2023 ALife Robotics Corp. Ltd.

OS18-3 Prototype Software for Designing Hula Accessories

Takumi Nakahara[1], Satoshi Ikeda[1], Amane Takei[1], Kenji Aoki[1], Makoto Sakamoto[1]*, Tsutomu Ito[2], Takao Ito[3]
([1]University of Miyazaki, Japan), ([2]NIT, Ube College, Japan), ([3]Hiroshima University, Japan)



Today, there are two main types of hula: the classical hula "kahiko" and the modern hula "auana". Music, costumes, and accessories are essential to these hula. Auana, in particular, is danced to music played on Western instruments and expresses the mythology, history, and various aspects of Hawaiian culture. The costumes and accessories worn by the dancers are designed to match the music. Therefore, we thought that it would be possible to design costumes and accessories more in line with the image of the dancers if we could freely design them using a personal computer and simulate how they would look when finished. In this study, we focused on accessories and conducted basic research to develop a CAD system for the design of "leis," one of the accessories.

OS18-4 A Basic Study on Museum Exhibition Support Using AR Technology

Kakeru Takemura[1], Satoshi Ikeda[1], Amane Takei[1], Masahiro Yokomichi[1], Makoto Sakamoto[1],
Shuichi Kurogi[2], Tsutomu Ito[3], Takao Ito[4]
([1]University of Miyazaki, Japan), ([2]Miyazaki Prefectural Museum of Nature and History, Japan), ([3]NIT, Ube College, Japan), ([4]Hiroshima University, Japan)

Enasirassitake are mushrooms that can be observed a lot in Aoshima, Miyazaki. Garnet Ochibatake was discovered in 2016 and confirmed in Miyazaki. Both are glow-in-the-dark mushrooms and are attracting attention as a new tourist resource for Miyazaki. When I visited the exhibition at the Miyazaki Prefectural Museum, these mushrooms were very small and difficult to observe. 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.

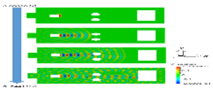
OS18-5 Tourism Support for Bioluminescent Fungi Using Video Technology

Bidesh Biswas Biki[1], Kodai Hasebe[1], Fumito Hamakawa[1], Satoshi Ikeda[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 about many bioluminescent fungi all over the world, and in recent years, 11 types have been confirmed in Miyazaki Prefecture. It is a considerable number from the perspective of Japan as a whole. The purpose of this study is to make the general public aware of bioluminescent fungi, and to widely publicize mushrooms of Miyazaki prefecture. Therefore, we conducted basic study on three-dimensional computer graphics models and virtual reality, and on educational promotional video for the general public cocerning bioluminescent fungi such as the *Favolaschia peziziformis* found in Miyazaki Prefecture. We also confirmed the usefulness of educational promotional video produced by conducting a questionnaire survey.

OS18-6 Parallel wave sound analysis based on hierarchical domain decomposition method

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



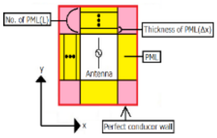
We are investigating a large-scale non-steady wave sound analysis method based on the parallel finite element method. The iterative domain decomposition method is employed in the analysis method as a parallel technique. We have confirmed that the non-steady wave sound analysis code is very high-accuracy with errors within the allowable range in a numerical analysis.

©ICAROB 2023 ALife Robotics Corp. Ltd.

OS18-7 Parallel full-wave electromagnetic field analysis based on hierarchical domain decomposition method

Amane Takei, Nanako Mizoguchi, Kento Ohnaka, Makoto Sakamoto (University of Miyazaki, Japan)

In this presentation, a parallel full-wave electromagnetic field analysis code based on an iterative domain decomposition method is explained that is named ADVENTURE_Fullwave. A stationary vector wave equation for the high-frequency electromagnetic field analyses is solved taking an electric field as an unknown function. Then, to solve subdomain problems by the direct method, the direct method based on the LDL^T decomposition method is introduced in subdomains. The simplified Berenger's PML is introduced which these eight corners are given the average value of all PML's layers. And, we show a numerical example of a microwave. More detail will be shown in the conference.



S19 Natural Computing (2)

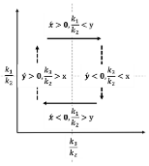
Chair Marion Oswald (Technische Universität Wien, Austria)

Co-Chair Yasuhiro Suzuki (Nagoya University, Japan)

OS19-1 Changes in the Behavior of a Small Number of Molecular Systems

Yasuhiro Suzuki (Nagoya University, Japan)

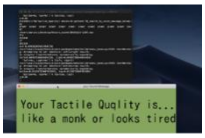
Molecular systems in chemical reaction systems have been considered continuous systems. However, chemical reactions in living organisms involve small molecules and cannot be considered a continuous system in some cases. In this study, we examine the behaviors of the two-party Lotka-Volterra model with a small number of molecules. We then show that there are cases in which intrinsic fewness is dominant.


$\dot{x} > 0, \frac{k_1}{k_2} < y$
$\frac{k_1}{k_2}$
$\dot{y} > 0, \frac{k_3}{k_2} > x$
$\dot{y} < 0, \frac{k_3}{k_2} < x$
$\dot{x} < 0, \frac{k_1}{k_2} > y$
$\frac{k_3}{k_2}$

OS19-2 Retrieval by Sensory Information

Yasuhiro Suzuki (Nagoya University, Japan)

Until now, information retrieval has been conducted by language. In recent years, information retrieval using smart microphones has also developed. However, voice-based retrieval is conducted by converting voice into linguistic information. In this study, we propose a search based not on linguistic information but on language sensitivity information. Language sensitivity information is a way of saying things. We perceive differences in the way we say the same word. In other words, the difference in how we say a word is the sensory information of the language. This study proposes a method for extracting and retrieving sensitive information from language.



```
Your Tactile Quqlity is...
like a monk or looks tired
```

©ICAROB 2023 ALife Robotics Corp. Ltd.

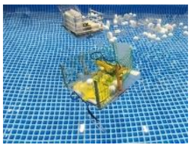
OS20 Robot Competitions and Education (7)

Chair Kazuo Ishii (Kyushu Institute of Technology, Japan)
Co-Chair Yasunori Takemura (Nishinippon Institute of Technology, Japan)

OS20-1 Report on Underwater-Robot-Festival Junior Division Aiming at Marine Debris Clean up

[1]Takayuki Matsuo, [2]Masanori Sato, [3]Masayoshi Ozawa, [4]Seiichiro Miura, [5]Masakazu Arima, [6]Kazuo Ishii
([1]National Institute of Technology, Kitakyushu College, Japan),
([2]Nagasaki Institute of Applied Science, Japan),
([3]Kobe City College of Technology, Japan), ([4]National Institute of Technology, Tokuyama College, Japan)
([5]Osaka Metropolitan University), ([6]Kyushu Institute of Technology)

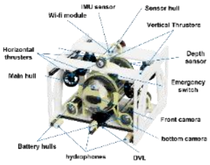
This paper reports the results of the Junior Division of the 8th Underwater Robot Festival held in August 2022. The junior division is a robotics competition for junior high and high schools, and 18 teams participated. The subject of the competition is Marine Debris Clean-up using surface ships with the theme of "Life Below Water" which is the 14th goal of SDGs. The competition consists of a poster session on the problem of garbage in the sea and a robot competition to compete the ability of their handmade ships to clean up floating objects. This paper gives an overview of the Junior Division and discusses its learning effects



OS20-2 Image-based navigation of Small-size Autonomous Underwater Vehicle “Kyubic” in International Underwater Robot Competition

Yusuke Mizoguchi, Daiki Hamada, Riku Fukuda, Irmiya R. Inniyaka, Kaito Kuwata, Keisuke Nishimuta,
Akihiro Sugino, Rikuto Tanaka, Yoshiki, Tanaka, Yuya Nishida, Kazuo Ishii
(Kyushu Institute of Technology, Japan)

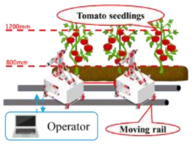
International underwater robot competition “RoboSub” is held in USA to demonstrate robot’s autonomy by completing underwater tasks, with a new theme each year. Student project team “Kyutech Underwater robotics” developed autonomous underwater vehicle (AUV) “KYUBIC” with built-in image processing board for the RoboSub. Sensors on the AUV are connected via ethernet and its automatic navigation program is built using ROS. The AUV moves through gate to the target panel based on self-localization using doppler velocity log and IMU. When approaching the panel, the AUV knows the direction to the panel based on the red line detection by image processing, and the type, the position and the attitude of the target panel are detected by deep learning on the board.



OS20-3 Development of Harvesting Robot for Tomato Robot Competition 2022 and Its Evaluation

Takeru Oshige[1], Moeko Tominaga[2], Takuya Fujinaga[3], Yasunori Takemura[2], Jonghyun Ahn
([1]Hiroshima Institute of Technology, [2]Nishinippon Institute of Technology, [3]Fukuoka University, Japan)

Agriculture is one of the most important industries for human food production. Recently, the number of farmers in Japan is decreasing, and the age of farmers is increasing. Therefore, automation of agriculture using robot systems is highly required. The “Tomato Robot Competition” is held every year in Fukuoka, Japan, to arouse student’s interest in the field of agriculture, and to promote the development of agricultural automation technology by robot systems. In this paper, we introduce the developed tomato harvesting robot to participate the “Tomato Robot Competition 2022”, and its harvesting performance evaluation. Developed tomato harvesting robot is composed 3 linear arms, end effector, rail movement system and electronic system of communication and control. In the evaluation, developed tomato harvesting robot took approximately 1.3 [min] to harvest a single tomato.

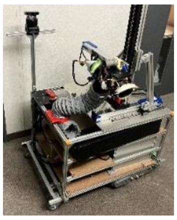


©ICAROB 2023 ALife Robotics Corp. Ltd.

OS20-4 Development of a Tomato Harvesting Robot for Farm Field

Shunsuke Oda, Ryuma Fukumoto, Kensuke Hirata, Shu Tahara, Keisuke Yoshida, Shinsuke Yasukawa,
Kazuo Ishii (Kyushu Institute of Technology, Japan,)

The 9th Tomato Harvesting Robot Competition is held at the green house in Kitakyushu Science and Research Park. The competition consists of two leagues of different field aresa; rail-style for greenhouse and free-style area aiming for farm field, and our team jointed the free-style. In order to participate in free-style, tomato harvesting mechanism, mobility mechanism to move on the rough terrain, camera to photograph tomato, and self-location system are required. We development of 3-axis cartesian coordinates manipulator tomato harvesting robot that using crawler move on soil and tomatoes are harvested by suction and cutting. In this paper, we describe the system architecture of tomato harvesting robot and the results of 9th Tomato Harvesting Robot Competition in 2022.



OS20-5 Tomato-Harvesting Robot Competition: Developed Robots and Results of 9th Competition

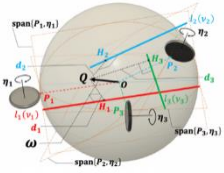
[1]Kazuo Ishii, [2]Takayuki Matsuo, [3]Yasunori Takemura, [3]Takashi Sonoda, [3]Atsushi Sanada, [1]Yuya Nishida,
[1]Shinsuke Yasukawa, [4]Takuya Fujinaga, [3]Moeko Tominaga, [5]Daisaku Arita, [6]Kazushi Kawajiri,
[7]Kenich Ohshima, [7]Masayuki Okada, [1]Kanako Shirahashi
([1]Kyushu Institute of Technology, Japan), ([2]National Institute of Technology, Kitakyushu College, Japan),
([3]Nishinippon Institute of Technology, Japan), ([4]Fukuoka University, Japan), ([5]University of Nagasaki, Japan),
([6]Hibikinada Greenhouse, Japan), ([7]Kyushu Polytechnic College, Japan)

Tomato is one of the important fruit vegetables and most tomatoes are produced in the greenhouses, or large-scale farms, where the high temperature and humidity, and long harvest age force the farmer heavy works. To develop the tomato harvesting robot, many research issues exist such as manipulator design, end-effector design, collaborative behavior, artificial intelligence, motor control, image processing, target recognition and so on. With an aim to promote the automation of tomato harvesting, we have organized the tomato harvesting robot competition since 2014, and currently changed the competition field to the greenhouse in 2020. In this paper, we discuss the results of 9th tomato harvesting robot competition in 2022.

OS20-6 A Modeling of Sphere Considering Slipping Adapted Three-Rollers

[1]Kenji Kimura, [2]Kouki Ogata,[3]Hiroyasu Hirai, [3]Takumi Ueda, [3]Kazuo Ishi
([1]National Institute of Technology, Matsue College, Japan), ([2]Saga University, Japan),
([3]Kyushu Institute of Technology, Japan)

Many types of spherical robots use friction-drive systems for locomotion because such systems enable omnidirectional movement and are more capable of climbing steps than mobile robots equipped with multiple omni-wheels. Slipping between spheres and rollers is a remarkable issue with friction-driven mechanisms. However, the previously established sphere kinematic models do not consider slipping, and kinematic models consider slipping in only two constraint rollers. In this study, we propose a mathematical model that allows for slipping on three constraint rollers and simulate the angular velocity vector of the sphere and slip vectors in each contact points.

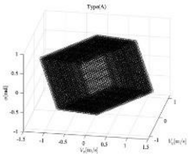


©ICAROB 2023 ALife Robotics Corp. Ltd.

OS20-7 Roller Arrangement Problem of Omnidirectional Mobil Robot Adapted Three Omni Rollers

[1]Kenji Kimura, [2]Yuki Shigyo,[3]Kazuo Ishi
([1]National Institute of Technology, Matsue College, Japan), ([2]Fujitsu, Japan),
[3]Kyushu Institute of Technology, Japan)

Mobile robots adapted to omni rollers are required to have efficient mobility in areas like logistics. Since such systems are easily controlled and provide omnidirectional locomotion. But theoretical research about motion efficiency has not been conducted. In this study, we evaluate rollers arrangement from speed efficiency point of view assumed a mechanism where the roller placement position can be changed arbitrarily on a round shape mechanism and we consider roller arrangement using the theory of linear transformation.



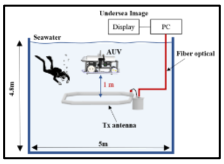
OS21 Advances in Field Robotics and Their Applications (7)

Chair Keisuke Watanabe (Tokai University, Japan)
Co-Chair Kazuo Ishii (Kyushu Institute of Technology, Japan)

OS21-1 Underwater Live Video Streaming Experiment Using Radio Frequency Communication for AUVs

Raji Alahmad, Yuya Nishida, Kazuo Ishii, Yukihiro Fukumoto (Kyushu Institute of Technology, Japan)

Many underwater applications need high data rate transmission to transmit the data such as images in real-time. Autonomous Underwater Vehicles (AUVs) require long-distance communication, especially in the deep sea. The technology of acoustic communication is the commonly used way for AUVs. The major drawback in underwater acoustic communication is the low transmission baud rate which is up to several kilobits per second (kbps). However, the reliable distance can be in tens of kilometers. On the other hand, the radio frequency (RF) communication provide much higher data rate. The electromagnetic wave is seriously limited by high attenuation in a water medium. In this paper, we investigate the Radio frequency communication in seawater, the experiment results show the effects of the distance between the transmitter and receiver, and the stability of the antennas. We could achieve HD video transmission with 25fps.



OS21-2 Control strategy to change the locomotion mode of a reconfigurable wheel/track robot based on the soil conditions

Supaphon Kamon, Enrico di Maria, Kazuo Ishii (Kyushu Institute of Technology, Japan)

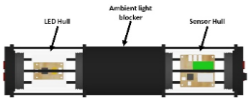
The use of agricultural machinery damages the soil by compaction and distortion, where the compaction is caused more by small contact areas like in wheeled vehicles, and the distortion is caused especially by tracks. In this work, we discuss a wheel/track reconfigurable robot which changes locomotion based on the soil conditions, to minimize the soil damage, energy consumption, and adapt the traversability. After giving an overview of the system, we propose a control strategy for switching between the locomotion modes.

©ICAROB 2023 ALife Robotics Corp. Ltd.

OS21-3 Sea-floor Image Restoration with Variable Absorbance Coefficient.

Irmiya R. Inniyaka, Yuya Nishida, Kazuo Ishii (Kyushu Institute of Technology, Japan)

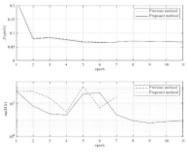
Scattering from suspended particles and visible light attenuation with water depth are constant phenomena that reduce underwater image quality due to low contrast and color distortion. Common image restoration techniques such as the image formation model, assumes a constant attenuation coefficient across color channels. This results in a restoration solution with limited application. We propose a method of image restoration that takes into account the wavelength-dependent attenuation of underwater images by taking in situ measurements of absorbance per image color channel. In this paper, a description of the design of a turbidity meter is made. It used to extract absorbance of light in the RGB channel. To evaluate the proposed method, image and absorbance data are collected concurrently from different water types. An analysis of image processing method presented in comparison to other methods.



OS21-4 Design of A Parameter Update Method of the Database-Driven PID Controller Considering Norm of the System

Takumi Ueda, Kazuo Ishii (Kyushu Institute of Technology, Japan)

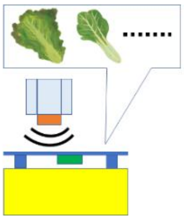
PID controllers are used in many applications, and they considered as typical examples of fixed-structure controllers. For each specific system, the PID parameters of the controller must be adjusted appropriately. Therefore, database-driven control, which automatically adjusts the parameters of the PID controller using a large amount of experimental data, has attracted attention. Database-driven control is one of the methods that does not require model information. Therefore, in the Database-driven PID (DD-PID), the Jacobian of the system used for the steepest descent method is unknown. In this paper, we present an optimization method using the Gradient descent method based on the local model of the system. And we proposed method that can suppress the increase of norm more than the conventional DD-PID method.



OS21-5 Occluded Objects Detection by Ultrasonic Sensors

Ryuugo Mochizuki, Yuya Nishida and Kazuo Ishii (Kyushu Institute of Technology, Japan)

Crops are required to be processed in sanitary manner for lunchbox production in industries. Particularly, foreign object detection is crucial to guarantee consumer’s health. Actually, foreign objects are mainly detected by visual check manually. Image processing method is an example to automate the foreign object detection process, however, is not valid to detect hidden objects, thus, non-destructive detection is essential. We propose a method of foreign object detection with ultrasonic reflection analysis. In our experiment, we selected four crops (lettuce, spinach, Perilla, komatsuna) for masking material, and a coin for foreign object. The coin was attached on the backside of masking. Reflected wave was recorded with the masking materials linearly moving. As the result, reflection intensity tripled by the existence of the foreign object.

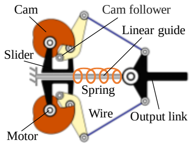


©ICAROB 2023 ALife Robotics Corp. Ltd.

OS21-6 Development of a Variable Stiffness Function for a New Multifunctional Wire Driven Joint Mechanism

Katsuaki Suzuki, Yuya Nishida, Kazuo Ishii (Kyushu Institute of Technology)

A mechanical variable stiffness can adapt to external forces beyond the control cycle, such as overturning of walking robots, which can help solve problems such as actuator and joint destruction. In this research, we will clarify the structure of a two-input, one-output joint mechanism that can realize three functions: normal motion, instantaneous motion, and variable stiffness function. As part of the development of the variable stiffness function, a mathematical model is derived and simulated for the relationship between the joint angle of the output link and the disturbance torque when a disturbance torque is applied to the output link of the proposed mechanism.



OS21-7 Analyzing an OFDM using Cyclic Prefix to Improve the Underwater Communication System

Alraie Hussam, Kazuo Ishii (Kyushu Institute of Technology, Japan)

The Underwater Communication is one of the most difficult challenges facing the researchers. OFDM technique has been used widely in 4G communication system, and recently it was approved to successfully implemented in 5G. Many researchers have improved the Underwater wireless communication system including the Acoustic communication using OFDM technique. In this study, we applied the Cyclic Prefix to improve the underwater communication system. We started by study the effects of Rician fading based on OFDM system on the wireless channel. Then we analyzed the Bit Error Rate (BER) of the system in several scenarios, by applying the AWGN, Rician fading, and absorption factor to simulate the underwater channel. The results showed that using Cyclic Prefix could improve the BER in underwater communication environment.

N inputs
MUX
1 Link for N channels
DEMUX
N outputs

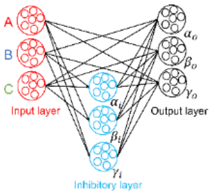
OS22 Bio-inspired Artificial Vision -Algorithms and Systems-(5)

Chair Shinsuke Yasukawa (Kyushu Institute of Technology, Japan)
Co-Chair Yuki Hayashida (Mie University, Japan)

OS22-1 Increasing selectivity to a feature combination using inhibitory synaptic plasticity in a spiking neural network

Mahiro Ikeda, Hirotsugu Okuno (Osaka Institute of Technology, Japan)

Applying spiking neural networks to artificial intelligence requires various techniques, such as a learning algorithm and a topology for extracting a feature combination, which have already been established for conventional neural networks. In this study, we designed a spiking neural network that uses synaptic plasticity to increase selectivity to a particular combination of features. The model of neurons used is a leaky integrate-and-fire model, and the learning rule used is the long-term potentiation of inhibitory inputs (LTPi). We investigated how the time constant of inhibitory presynaptic neurons whose weights were updated by LTPi affects to selectivity of the postsynaptic neurons. The results showed that the selectivity was increased effectively when the time constant of inhibitory neurons was slightly longer than that of postsynaptic neurons.

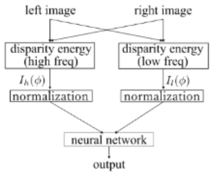


©ICAROB 2023 ALife Robotics Corp. Ltd.

OS22-2 A binocular disparity estimation algorithm using multiple spatial frequency information and a neural network

Ryoka Sato, Hirotsugu Okuno (Osaka Institute of Technology, Japan)

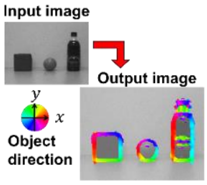
In this study, we developed a disparity estimation algorithm using multiple spatial frequency information and neural-network-based regression. First, the algorithm computes a value called disparity energy based on a model of neurons in the visual cortex that respond selectively to a particular disparity. A Gabor filter with a particular spatial frequency is used to compute the disparity energy, and therefore the disparity energy depends on the spatial frequency of the input image. To reduce the spatial frequency dependency, our algorithm uses disparity energy values computed from two different spatial frequencies for neural-network-based regression.



OS22-3 A figure-ground discrimination algorithm inspired by border-ownership selective cells

Tomoya Kobayashi, Hirotsugu Okuno (Osaka Institute of Technology, Japan)

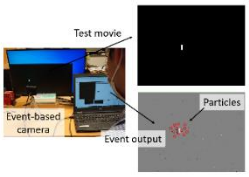
Figure-ground discrimination is one of the most important functions for image classification. In this study, we developed a figure-ground discrimination algorithm that is computationally inexpensive and is suitable for robot vision. The algorithm is inspired by the response properties of border-ownership (BO)-selective cells, which possibly support figure-ground discrimination in the visual nervous system. The output of our algorithm is affected by spatial parameters such as the characteristics of the spatial filter used for edge enhancement and the interval of model BO-selective cells. Therefore, we investigated relationship between the algorithm output with various spatial parameters and object sizes.



OS22-4 Event-Driven Particle Filter for Tracking Irregularly Moving Objects

Yuki Kawasaki[1], Masahiro Ohtani[2], Shinsuke Yasukawa[1]
([1]Kyushu Institute of Technology, [2]Mie University, Japan)

Conventional object tracking techniques that use general-purpose cameras and particle filters find it difficult to track irregularly and rapidly moving objects. To track an irregularly moving object without losing sight, quickly measuring the position of the object is necessary. In this study, we used a fast-response event-based camera, which is a bioinspired camera that produces a spiking output. We propose an event-driven particle filter that performs processing in response to the input from an event-based camera. Our proposed method was evaluated by presenting an event-based camera with a rectangular motion pattern that combines periodic and constant-velocity motions at various speeds. The experimental results demonstrated that our proposed method could track objects in a test video.

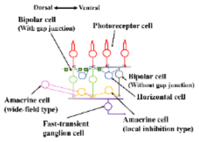


©ICAROB 2023 ALife Robotics Corp. Ltd.

OS22-5 A simulation model for analyzing the spatiotemporal receptive field of retinal ganglion cells in the presence of fixational eye movements

Hiroyuki Yokota[1], Yuki Hayashida[2], Shinsuke Yasukawa[1]
([1]Kyushu Institute of Technology, [2]Mie University, Japan)

Understanding the receptive field dynamics of vertebrate retinal ganglion cells in the presence of fixational eye movements is considered to provide insights into information processing/encoding optimized for behaviors in animals and autonomous robots in the future. Previous studies have proposed computational models that account for fixational eye movements, suggesting that the long-range spatial inhibitions provided by wide field amacrine cells to some subtypes of bipolar cells and ganglion cells dynamically shape the responses of ganglion cells. In this study, we constructed a simulation model of a retinal circuit based on recent physiological findings. The model validity was tested through computer simulation experiments and by analyzing the time series of spike outputs from the ganglion cell unit. As a preliminary result, we were able to quantify the spatiotemporal receptive field by applying a simple white-noise movie stimulus.



OS23 STEM Education for fostering Innovators (4)

Chair Hiroyuki Y. Suzuki (Hiroshima University, Japan)
Co-Chair Kazuo Kawada (Hiroshima University, Japan)

OS23-1 Junior High School Rescue Robot Challenge for Fostering Problem-Solving Skills

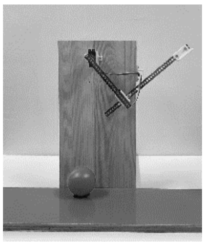
Kazuo Kawada, Keita Murai, Yuta Susawa, Hiroyuki Y. Suzuki (Hiroshima University, Japan)

Junior High School Student Rescue Robot Challenge" is an annual activity organized by Hiroshima University in cooperation with construction equipment manufacturers. The Challenge has been continued for 18 years with a given theme in each year. Nonetheless, the theme was essentially changed from the last year, according to revised version of the Courses of Study for junior high schools announced by the Japanese Ministry of Education, Culture, Sports, Science and Technology, in which the both problem-finding and problem-solving skills are emphasized. The theme for 2022 includes proposal of the problem to be solved in disaster-stricken areas, and it will be solved by themselves by making prototype rescue robot by modifying remote-controlled excavator model of 1/14 scale.

OS23-2 A Study of Experiential Learning Activities using Model Materials for the Kicking Motion

Teruyuki Tamai, Shoki Takeuchi, Yoshihiro Ohnishi (Ehime University, Japan)
Kazuo Kawada (Hiroshima University, Japan)

In this research, model teaching material to be used for cross-curricular learning, which is expected to be promoted in elementary and junior high school education, is explained. This material produced using materials such as servo motors and frames were proposed as teaching materials that would allow learners to consider the kicking motion. First, the relationship between cross-curricular learning and the proposed teaching material is explained. In particular, the relationship with the content of technology education in junior high schools is explained. Next, the structure and mechanism of the teaching materials and the programs used to make them are described. Finally, data and results obtained from the use of the materials are presented.



©ICAROB 2023 ALife Robotics Corp. Ltd.

OS23-3 Making High Precision Single Balance in Active Learning Seminar for Hiroshima Univ. Monozukuri Junior Doctor Special Educational Program

Hiroyuki Y. Suzuki, Kazuo Kawada, Masayasu Nagamatsu (Hiroshima University, Japan)

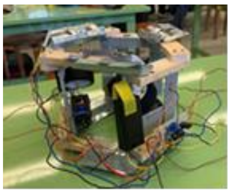
Hiroshima Univ. Monozukuri Junior Doctor" is a special educational program for young (11 – 15 in age) people. In the program we give several seminars for nurturing innovative minds and capacities, including an active learning seminar on making single balance. We prepared a number of pre-fabricated parts of deferent sizes, giving them to young participants teams (about 5 people each) to assemble single balance, putting their own ideas in the design by selecting favorite parts by themselves. Majority of the students could understand key concept of the balance (rule of moment) and assemble single balance with accuracy of milligrams.

.

OS23-4 Capstone Class of "Mechatronics Innovation Project" as STEM Educational Curriculum for Teacher Training Course

Hiroyuki Y. Suzuki, Masayasu Nagamatsu, Kazuo Kawada (Hiroshima University, Japan)

We have a set of STEM oriented special curriculum with name of "Mechatronics" in our teacher training course of "Program in Technology and Information Education, in School of Education, Hiroshima Univ.". The mechatronics classes start from first grade, and they are eventually integrated in a capstone class of "Mechatronics Innovation Project" in third grade. We set a different theme in each year, and students are going to solve it by making automatic "mechatronics" machines. The theme in 2022 is guardrail cleaning robots. We are going to report the struggle of students to the theme and their outcomes, achievements and problems.



OS24 Applications of Unmanned Aerial System (3)

Chair Hazry Desa (Universiti Malaysia Perlis, Malaysia)
Co-Chair Nurfadzillah Ishak (Universiti Malaysia Perlis, Malaysia)

OS24-1 Effect of Spraying Dispersion Using UAV Spraying System with Different Height at Paddy Field

Hazry Desa[1], Muhammad Azizi Azizan[1], Nik Noriman Zulkepli[1], Nurfadzillah Ishak[1], Teh Xi Hang[1], Siti Syuhaidah Yahya[1], Aisyah Arina Mohammad Shahrazel[1], Fakhrul Mukmin Mansor[1], Siti Zaleha Abdul Aziz[2], Abadal-Salam T Hussain[3]
([1]universiti Malaysia Perlis, Malaysia), ([2]Mara Japan Industrial Institute (MJII), Malaysia)
([3]Al-Kitab University, Iraq)

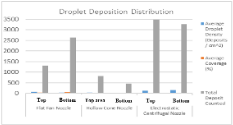
This study investigated the UAV spraying system height in relation to spraying uniformity and dispersion. The operating heights of the UAV spraying system at height of 1 m, 1.5 m, and 2 m from the hollow cone nozzles were investigated within a wind speed of 2.8 m/s. The tests were to determine the spray uniformity and dispersion on the water sensitive paper that was placed on the paddy plant. The results of water droplet samples were evaluated using ImageJ software. The results show the droplet distribution at 1.5 m height has high values for average droplet density, which is 162.7 deposits/cm² at the top area and 161.8 deposits/cm² at the bottom area. The percentage of coverage was also high, at 55.21% at the top area and 51.4% at the bottom area.

©ICAROB 2023 ALife Robotics Corp. Ltd.

OS24-2 Spraying Dispersion Analysis with Different Nozzle Types Using UAV Spraying System in a Paddy Field

Hazry Desa[1], Muhammad Azizi Azizan[1], Nik Noriman Zulkepli[1], Nurfadzillah Ishak[1], Tan Yew Tian[1], Siti Syuhaidah Yahya[1], Aisyah Arina Mohammad Shahrazel[1], Fakhrul Mukmin Mansor[1], Siti Zaleha Abdul Aziz[2], Abadal-Salam T Hussain[3]
([1]universiti Malaysia Perlis, Malaysia), ([2]Mara Japan Industrial Institute (MJII), Malaysia)
([3]Al-Kitab University, Iraq)

This study investigates the ability of Unmanned Aerial Vehicle (UAV) spraying systems to be used as an agriculture spraying method in Malaysia. The operating height of the UAV was 1.5 m with three different nozzles were investigated within a wind speed of 1.15 m/s to determine spray uniformity and dispersion in the paddy field conditions. The results from these samples were evaluated by using ImageJ software. The results show that the droplet distribution by using an electrostatic centrifugal nozzle has a high average droplet density, which is 134.03 deposits/cm2 for the top area and 153.93 deposits/cm2 for the bottom area. The electrostatic centrifugal nozzle also testified to the high value of total droplet deposit at 3478 for the top area and 3255 for the bottom area.



OS24-3 The capabilities and Readiness of Unmanned Aerial System (UAS) implementation in Construction Work Progression

Nurfadzillah Ishak, Muhammad Azizi Azizan, Hazry Desa (universiti Malaysia Perlis, Malaysia)

Unmanned Aerial Systems (UAS) have evolved over the past decade as both advanced military technology and off-the-shelf consumer devices. Nowadays, the implementation of Unmanned Aerial Systems technology is becoming increasingly widespread in construction industry. The uses of drones are increasingly moving towards successful remote procedures which can take place in a range of in building industry. Unmanned Aerial Systems remote sensing equipment is a way in which current building construction work progress can be analyzed and inspected. Therefore, this study would be beneficial as eye opener for building construction practitioner that this remote sensing equipment can replaced the manual paper-based supervision and unsystematic database of current conventional construction work progression process.

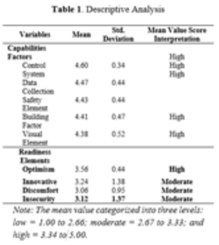


Table 1. Descriptive Analysis

Variables	Mean	Std. Deviation	Mean Value Score Interpretation
Capabilities Factors			High
Control System	4.60	0.34	High High
Data Collection	4.47	0.44	
Safety Element	4.43	0.44	
Building Factor	4.41	0.47	High
Visual Element	4.38	0.52	High
Readiness Elements			
Optimism	3.56	0.44	**High**
Innovative	3.24	1.38	**Moderate**
Discomfort	3.06	0.95	**Moderate**
Insecurity	**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.

©ICAROB 2023 ALife Robotics Corp. Ltd.

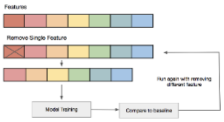
OS25 Emerging Technologies in Industrial Revolution 4.0 (IR 4.0) Era (8)

Chair Takao Ito (Hiroshima University, Japan)
Co-Chair Wei Hong Lim (UCSI University, Malaysia)

OS25-1 Chaotic African Vultures Optimization Algorithm for Feature Selection

Wy-Liang Cheng[1], Li Pan[1], Mohd Rizon Bin Mohamed Juhari[1], Chin Hong Wong[2,3], Abhishek Sharma[4], Tiong Hoo Lim[5], Sew Sun Tiang[1], Wei Hong Lim[1]
([1]UCSI University, Malaysia), ([2]Maynooth University, Ireland), ([3]Fuzou University, China), ([4]Graphic Era University, India), ([5]Universiti Teknologi Brunei, Brunei Darussalam)

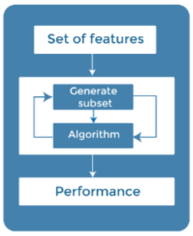
Feature selection is widely used to decrease the number of features by removing undesirable, noisy and inaccurate data while maintaining the classification accuracy. One way to improve the performance of feature selection is through metaheuristic search algorithms. Traditional optimization algorithms have poor initialization scheme that tend to trap into local optima when processing datasets with complex feature. A new variant known as Chaotic African Vultures Optimization Algorithm (CAVOA) is proposed to solve feature selection problems with better performances by leveraging the benefits of chaotic map in population initialization. Twelve datasets obtained from UCI Machine Learning Repository are used to investigate the capability of CAVOA in feature selection and compare with six competing algorithms. The proposed CAVOA has shown leading performance when against other competing algorithms by producing ten out of twelve best mean accuracies and four out of twelve datasets that have the lowest number of features.



OS25-2 Multi Chaotic Flow Directional Algorithm for Feature Selection

Wy-Liang Cheng[1], Li Pan[1], Mohd Rizon Bin Mohamed Juhari[1], Abhishek Sharma[2], Hameedur Rahman[3], Chun Kit Ang[1], Sew Sun Tiang[1], Wei Hong Lim[1]
([1]UCSI University, Malaysia), ([2]Graphic Era University, India), ([3]Air University, Pakistan)

Feature selection is crucial in optimizing performance as it can reduce processing time and preserve classification accuracy. One way to further improve the capability is to implement a metaheuristic search algorithm as it is widely used by many researchers. This method has some disadvantages due to random population initialization scheme as it can easily trap in the local optima especially handling large quantities of features in the dataset. A modified algorithm known as Multi Chaotic Flow Directional Algorithm (MCFDA) is presented by incorporating multiple chaotic maps for population initialization. For validating the performance of the proposed algorithm, six competing algorithms are used to compare MCFDA on fourteen datasets received from UCI Machine Learning Repository. The results show that the proposed algorithm can obtain eight out of fourteen best mean accuracies and six out of fourteen datasets that have the least number of features.

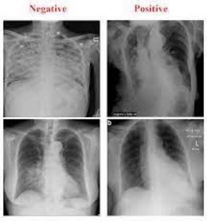


©ICAROB 2023 ALife Robotics Corp. Ltd.

OS25-3 Performance Comparison of Convolutional Neural Networks for COVID-19 Diagnosis

Suhaim Parvez Wadekar[1], Koon Meng Ang[1], Chin Hong Wong[2,3], Abhishek Sharma[4], Tiong Hoo Lim[5], Chun Kit Ang[1], Sew Sun Tiang[1], Wei Hong Lim[1]
([1]UCSI University, Malaysia), ([2]Maynooth University, Ireland), ([3]Fuzou University, China), ([4]Graphic Era University, India), ([5]Universiti Teknologi Brunei, Brunei Darussalam)

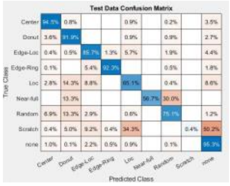
COVID-19 has devastated the global healthcare system as well as the economy with more than 600 million confirmed cases and 6 million deaths globally. A timely and accurate diagnosis of the disease plays a vital role in the treatment and preventative spread of disease. Recently, deep learning such as Convolutional Neural Networks (CNNs) have achieved extraordinary results in many applications such as medical classifications. This work focuses on investigating the comparison of nine state-of-the-art architectures: Alexnet, Googlenet, Inception-v3, Mobilenet-v2, Resnet-18, Resnet-50, Shufflenet, Squeezenet and Resnet-50 RCNN for COVID-19 classification by comparing with performance metrics such as accuracy, precision, sensitivity, specificity and F-score. The datasets considered in current study are divided into three different classes namely Normal Chest X-Rays (CXRs), Pneumonia patient CXR and COVID-19 patient CXR. The results achieved shows that Resnet-50 RCNN achieved an accuracy, precision, sensitivity, specificity and F-score of 95.67%, 95.71%, 95.67%, 97.84% and 95.67% respectively.



OS25-4 Classification of Wafer Defects with Optimized Deep Learning Model

Koon Hian Ang[1], Koon Meng Ang[1], Mohd Rizon Bin Mohamed Juhari[1], Chin Hong Wong[2,3], Abhishek Sharma[4], Chun Kit Ang[1], Sew Sun Tiang[1], Wei Hong Lim[1]
([1]UCSI University, Malaysia), ([2]Maynooth University, Ireland), ([3]Fuzou University, China), ([4]Graphic Era University, India)

Wafer defect inspection is one of the crucial semiconductor processing technologies because it can help to identify the surface defects in the process and eventually improve the yield. Manual inspection using human eye is subjective and long-term fatigue can lead to erroneous classification. Deep learning technology such as convolutional neural network (CNN) is a promising way to achieve automated wafer defect classification. The training of CNN is time consuming and it is nontrivial to fine tune its hyperparameters to achieve good classification performance. In this study, Arithmetic Optimization Algorithm (AOA) is proposed to optimize the CNN hyperparameters, such as momentum, initial learn rate, maximum epochs, L2 regularization, to reduce the burden brought by trial-and-error methods. The hyperparameters of a well-known pretrained model, i.e., Google Net, are optimized using AOA to perform wafer defects classification task. Simulation studies report that the AOA-optimized Google Net achieves promising accuracy of 91.32% in classifying wafer defects.

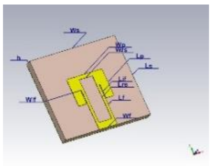


.

OS25-5 Compact Wearable Antenna for Millimeter-Wave (mm-Wave) Fifth Generation (5G)

Wai Kiat Wong[1], Sew Sun Tiang[1, *], Wei Hong Lim[1], Mastaneh Mokayef[1], Chin Hong Wong[2,3]
([1]UCSI University, Malaysia), ([2]Fuzou University, China), ([3]Maynooth University, Ireland)

The need for networking, communication, and data sharing capabilities among users of wearable terminal devices has increased, and this has made the new wearable antenna one of the most active research areas. This work presents a wearable antenna for 5G applications based on a microstrip patch antenna type operating at 28GHz millimeter-wave (mm-wave). The operating frequency of 28GHz is expected to be appropriate for 5G mm-wave wearable antenna design. The design makes use of the semi-flexible Rogers Duroid RO3003 substrate, which has a thickness of 0.75mm, a loss tangent of 0.001 and a relative permittivity of 3. CST Microwave Studio software is used to analyze and evaluate the proposed antenna's performance to other existing designs in terms of return loss, bandwidth, gain, directivity, and point SAR value.



©ICAROB 2023 ALife Robotics Corp. Ltd.

OS25-6 Driver's Fatigue Recognition using Convolutional Neural Network Approach

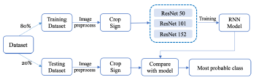
Samer Abbas[1], Sew Sun Tiang[1], Wei Hong Lim[1], Li Sze Chow[1], Chin Hong Wong[2,3]
([1]UCSI University, Malaysia), ([2]Fuzou University, China), ([3]Maynooth University, Ireland)

Drowsy driving is a serious issue that has been leaking in our communities since long time, the definition of drowsy driving is when the driver is not aware enough to proceed with driving the vehicle causing catastrophic accidents. Multiple methods were found to approach this complication across the years. Convolution Neural Network has approved to be a reliable approach to treat this issue by using face feature detection. In this paper, the effect of key parameters of the trained framework based on the driver's fatigue recognition model are analyzed, and the accuracy of the driver's fatigue recognition model is investigated, as well as a driver's fatigue recognition is studied under different conditions using CNN. Transfer learning is used to develop a reliable method for detection, Mediapip Face Mesh model is used to extract the features from the face. MAR (Mouth Aspect Ratio) as well as EAR (Eyes Aspect Ratio) are obtained through the detection, these terms are responsible for detecting the eye and mouth closure ratio, the model has proved to work with accuracy of 98.3% and in different light conditions with accuracy of 94.7% outperforming several past models.

OS25-7 Deep Residual Neural Network for Efficient Traffic Sign Detection System

Hanlin Cai[1,2], Jiaqi Hu[1,2], Zheng Li[1,2], Wei Hong Lim[3], Sew Sun Tiang[3], Chin Hong Wong[1,2]
([1]Fuzou University, China), ([2]Maynooth University, Ireland), ([3]UCSI University, Malaysia)

This paper has proposed a deep residual neural network (RNN) model for traffic signs detection system (TSDS) research. Experiments are conducted to verify the feasibility of implementing the RNN model for traffic sign detection and recognition. Moreover, a new systematic analytic hierarchy process (AHP) method for model performance evaluation has been suggested, which is sufficient for deployment in the practical performance measurement of the deep learning model.



OS25-8 Wall Crack Detection based on Adaptive Double Threshold Grayscale Transform

Mingrui Lin[1,2], Xin Xu[1,2], Tengxiang Li[1,2], Yuhang Hong[1,2], Weiqin Wang[1,2], Shilin Chen[1,2], Wei Hong Lim[3], Chin Hong Wong[1,2]
([1]Fuzou University, China), ([2]Maynooth University, Ireland), ([3]UCSI University, Malaysia)

The construction industry is an important supporting social-economic development, it is of great significance to detect cracks in walls to realize the analysis and provide early warning of building. A wall crack detection algorithm is proposed based on an adaptive double-threshold grayscale transform. The MATLAB built-in function graythresh is modified to have a flag bit to preliminarily select the grayscale transformation threshold. The grey transform and binarization threshold can be automatically adjusted based on the image processing effect using a preliminary-selected threshold. The MATLAB simulation results show that the algorithm has an accuracy of 96.16%.

©ICAROB 2023 ALife Robotics Corp. Ltd.

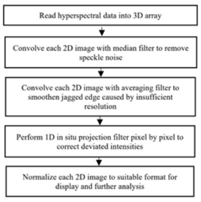
OS26 Signals Processing and Automation I (6)

Chair Norrima Mokhtar (University of Malaya, Malaysia)
Co-Chair Heshalini Rajagopal (UCSI University, Malaysia)

OS26-1 Noise Filtering of Hyperspectral Data of Oil Palms by Median-Mean Projection Filtering

Imanurfatiehah Ibrahim[1], Hamzah Arof[1], Mohd Izzuddin Anuar[2], Mohamad Sofian Abu Talip[1]
([1]Universiti Malaya, Malaysia), ([2]Malaysian Palm Oil Board, Malaysia)

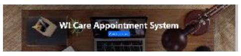
One of the many applications of hyperspectral imaging is in agriculture. However, hyperspectral data captured from airborne UAV sometimes contain noise that make their spectral signatures different from those of field spectroscopy using similar wavelengths. Therefore, there exists a need to filter noisy hyperspectral data to improve their quality so that a strong correlation can be established between the airborne and field hyperspectral data for effective analysis. For oil palm hyperspectral data, an efficient and effective method is introduced to filter noise using median-mean projection filtering. This novel approach generates superior results compared to those produced by the conventional method of convolving the data with 2-D filters, in terms of output quality and signal to noise ratio.


Read hyperspectral data into 3D array
Convolve each 2D image with median filter to remove speckle noise
Convolve each 2D image with averaging filter to smoothen jagged edge caused by insufficient resolution
Perform 1D in situ projection filter pixel by pixel to correct deviated intensities
Normalize each 2D image to suitable format for display and further analysis

OS26-2 Smart Telehealth Appointment System – WI Care

Siah Cheong Lin, Heshalini Rajagopal, Chloe Thong Chee Ling (UCSI University, Malaysia)

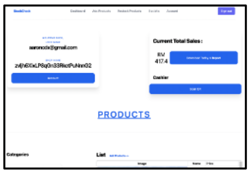
This study aims to create a Smart Telehealth Appointment System -WI Care that enables user to make an appointment in ease. Due to COVID-19 pandemics, constant lockdown had resulting social distancing measures, movement control order, and people avoid outdoor activities. The proposed system, WI Care will have an appointment system that can be managed by doctors meanwhile appointer can keep submitting form. WI Care ensure both parties can be conducting the same procedure as how every GPs and doctors able to work anywhere anytime. With dashboard and time scheduling, these are able to create chance for doctors and GPs having manageable workload as individual. This study suggests the healthy environment whereas every doctor can join WI Care to share skills and tackles problem together.


WI Care Appointment System

OS26-3 Web-based Stocktaking application in Businesses

Cheah Dei Xuan, Heshalini Rajagopal, Shayla Islam (UCSI University, Malaysia)

This study aims to develop a web-based application on business stock checking where workers able to track their quantity of their stock from running out of stock as well as to track the business statistic. This application is developed using Vue 2 framework, Java script language, html 5, and tailwind CSS. In this study, a database was developed for shopkeeper to store all the data related to the shop on cloud-based that able to prevent from disaster or human errors. Furthermore, a web-based system was developed to perform stock checking and monitoring which allow shopkeeper to check on statistic of the product and also the notify shopkeeper which stock are running low. In addition, a softcopy receipt was implemented as a default for customer instead of hardcopy.

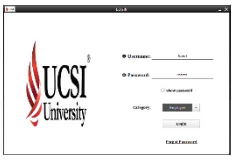

Current Total Sales :
RM 417.4
Cashier
PRODUCTS
Categories
List

©ICAROB 2023 ALife Robotics Corp. Ltd.

OS26-4 Pharmacy Warehouse Management System

Gan Jhui Ken, Heshalini Rajagopal, Shaik Shabana Anjum (UCSI University, Malaysia)

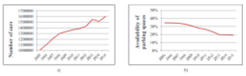
Managing of the stock and supply of medication plays an essential part for the provisioning of health care, therefore this study aims to develop the web-based warehouse system. With the use of modern technology, the warehouse management system can be made better to help improve the quality of work for both the employees and admins. Additionally, this system will include the basic functions of a warehouse management system that will allow users to utilize the system with ease. The proposed system aims to improve the current pharmacy warehouse management systems so that it will be viable for users to understand and utilize the system efficiently.



OS26-5 A Development of an Automatic Allocation Parking System

Ng Wai Lam, Neesha Jothi, Shayla Islam (UCSI University, Malaysia)

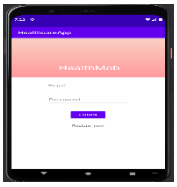
As the number of cars increases, the demand for parking spaces in the parking lot also increases. Due to the lack of parking spaces in the parking lot, people spend more time looking for parking spaces after entering the parking lot, and also consume a lot of car oil. Therefore, this project is to develop a new parking system that will automatically allocate parking spaces for user. In this study, literature review was done to determine how this type of system functions and how it is made. Then to make a questionnaire cantered on the users' previous experience on existing systems. Lastly, to make a comparison of existing systems and to analyze the systems strengths and weaknesses to determine what functions to add to this new system.



OS26-6 Healthcare Mobile Application

Yee Chee Hong, Neesha Jothi, Javid Iqbal (UCSI University, Malaysia)

This project aims to assist those people who live in a busy city to take care their health. They suffered a lot from working and other stress that caused themselves always to keep in unhealthy status. Although they know they are lives such unhealthy, still they have no time to manage it. Also, the unfriendly user interface of current healthcare mobile application that caused inconvenience to use them. This project will improve those weaknesses that allocated from current existing system. The healthcare mobile will open use to public users. It will come with online appointment, online pharmacy purchase, chat support and check information or history. New function added will be e wallet which can make payment or top up using serial number or smart wallet such as TNG pay and Shopee Pay.



©ICAROB 2023 ALife Robotics Corp. Ltd.

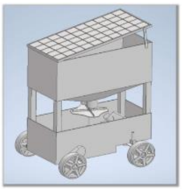
OS27 Research Towards the Sustainable Development Goals (SDG's) (6)

Chair Ammar A.M. Al Talib (UCSI University, Malaysia)
Co-Chair Takao Ito (Hiroshima University, Japan)

OS27-1 Solar Powered Seed Sprayer Machine

Ammar A.M. Al Talib, Yap Chee Xian, Ain Atiqa, Nor Fazilah Abdullah (UCSI University, Malaysia)

The main objective of this project is to fabricate a complete functional seed sprayer machine which is purely powered by solar energy. The performance is tested through the seed amount sprayed over time and the area that can be covered. In this project, a broadcast spreader has been used. It is solar powered with a wireless communication to remotely control the machine. The fabrication of the solar seed sprayer machine in this study has been divided into four main parts which are remote driving system, solar charging system, seed storage dispenser system, and impeller spreader system. Several experiments are conducted to analyze the performance of solar seed sprayer machine. The performance of the tested machine has been indicated through the capability of the machine on spreading different types of seeds with various sizes and shape.



OS27-2 Solar Powered Outdoor Air Purifier With Air Quality Monitoring

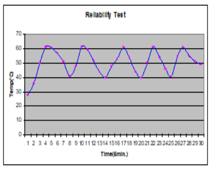
Ammar A.M.Al Talib, IK Chu Aung, Noor Idayu M. Tahir (UCSI University, Malaysia)

This paper is discussing the design and fabrication of a solar powered air purifier and its performance. Its aiming for the production of an outdoor air purifier powered by solar energy and with a High Efficiency Particulate Air Filter (HEPA) and Carbon Filters which can achieve air purification with self-sustainable ability. Several tests have been conducted to proof the good performance of air purifier. In the first indoor test, 15.79% of error and 67.37 % of efficiency was achieved during the solar panel efficiency test. Second test of air purifier test has shown the efficiency of cleaning ammonia pollutant in the air as 43.55% for burning cigarette and 35.33% for floor detergent using the equipped two MQ135 sensor. The findings has shown that the floor detergent might have higher rate of diffusion than ammonia molecule found in cigarette smoke. Comparisons have been conducted of theoretical and actual test results.

OS27-3 Design and Fabrication of a Mutual Control Electronic Circuit for Solar and Electrical Water Heating

Ammar A.M.Al Talib[1], Sarah Atifah Saruchi[2] ([1]UCSI University, Malaysia), ([2]UTP, Malaysia)

This research is a temperature controller that will be implemented to ensure that the water temperature of the solar water heating unit is maintained at the desirable level all the time. This control circuit is designed to control the action of electrical heater according to specific temperature range. Temperature sensor will sense the water temperature constantly and send signal to a micro-controller unit. The micro-controller will process the data according to a written program and control the actions of electrical heater. At the same time, temperature reading will display through LCD and real-time data can be viewed from a computer via serial port. During times of sufficient sunlight, solar will be the main energy source used for heating water; otherwise, there will be an automatic switching to the electrical operated immersion heater. This controller will give reliability to users of solar water heating systems.



©ICAROB 2023 ALife Robotics Corp. Ltd.

OS27-4 Optimization of the Major Factors Affecting The Home Recycling of Disposed (LDPE) Plastics

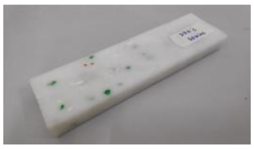
Ammar A.M.Al Talib, Chua Ray Sern Grayson, Ain Atiqa (UCSI University, Malaysia)

Low-density polyethylene (LDPE), which is widely adopted in many daily life products, took more than hundreds of years to decompose in the landfill. Although various efforts were carried out to recycle LDPE, the operations were proved to be uneconomical as it needs sophisticated equipment. Hence, this study is aimed to encourage simple domestic recycling procedures and to investigate the optimum recycling conditions for disposed LDPE plastics and explore potential green applications for the recycled material. As proposed in this study, disposed LDPE bags are cut, washed, dried and then inserted into oven for melting at specified heating temperatures and different time duration's inside the oven. Mechanical tests are conducted which include hardness test, compression test and tensile test. The recycled LDPE achieved 90.2 Shore-A hardness point, 9.2599 MPa of compressive strength and tensile strength of 9.0705 MPa. The recycled LDPE plastics are found useful in many life appliances.

OS27-5 Investigation of the Mechanical Properties and Applicability of Recycled Plastic Bags

Ammar A.M.Al Talib, Wan Wai Kit, Nor Fazilah Binti Abdullah (UCSI University, Malaysia)

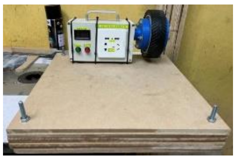
Low Density Polyethylene (LDPE) is a thermo-plastic made from the monomer ethylene. It is estimated that only a mere 5.7% of LDPE is recycled worldwide. This study is to reveal some of the mechanical properties of the recycled disposed plastic bags and finding the possible applications for them.. Mechanical properties such as hardness testing, compression testing and water absorption has been carried out on the LDPE recycled samples according to ASTM Standards. The mechanical tests conducted on the recycled specimens have indicated that the most suitable applications based on safety and satisfying needs would be the spring post that can be used on roads, field track or any situation where separation of section would be needed.



OS27-6 Design and Fabrication of Power Generating Treadmill

Ammar A. M. Al-Talib1, They Kai Yang[1], Sarah Atifah[2], Noor Idayu M.Tahir[1]
([1]UCSI University, Malaysia), ([2]UTP, Malaysia)

This paper aims to take advantage of treadmill’s wasted electrical energy during a person’s workout and utilize the energy for charging electrical appliances. In order to take advantage of this wasted energy, it could be harnessed by a power generator and stored in a battery bank. The prototype machine did not interrupt a person’s workout flow and it can be attached to any treadmill due to its design features. Wasted energy was harnessed in this project by a non-traditional manner of using shaft and wheel method. A multi-meter was used to measure the voltage, current and power generated. The prototype machine is found able to fully charge a 3096mAh smartphone in 135 minutes and the phone could be fully charged for 2 charge cycles. This power generating treadmill machine ia proved to be a very good application of the Sustainable development Goals (SDG’s).



©ICAROB 2023 ALife Robotics Corp. Ltd.

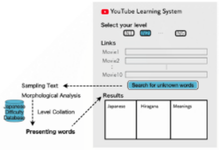
OS28 Is That Narratology ‘Post’? -Computational and Cognitive Approaches and Beyond (5)

Chair Jumpei Ono (Aomori University, Japan)
Co-Chair Hiroki Fxyma (Tainan University of Technology, Taiwan (ROC))
Co-Chair Yuki Hayashi (Chiba University / NINJAL, Japan)
Co-Chair Takashi Ogata (Iwate Prefectural University, Japan)

OS28-1 Proposing YouTube-based vocabulary learning system

Hiroki Fxyma (Tainan University of Technology, Taiwan (ROC))

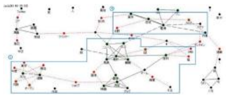
This study proposes a language learning support system using YouTube. Motivations for language learning are becoming more diverse: exposure to multi-language information is becoming more common through SNS. Thus, it is expected that the number of potential "sub-learners," rather than explicit learners in the classroom, is increasing. This study proposes a system that allows these potential learners to focus on the words they need to watch videos that interest them. As a manual prototype, we will use an existing subtitle extraction system to extract text and create a corpus.



OS28-2 A Study on the Impact of “Trending” and “Flaming” on Content from “A Crocodile Who Will Die in 100 Days”

Taiki Sugimoto, Jun Nakamura (Chuo University, Japan)

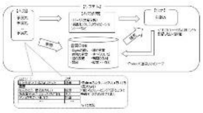
The subject of this research is “A Crocodile Who Will Die in 100 Days” which has experienced getting flamed. This study uses Sentiment Analysis and Key Graphics Analysis to capture a bird's eye view of the impact of trending and flaming generated through social media on content. The analysis and discussion clarify the following two points. First, there is a significant gap between the tendency of comments directly sent from recipients to senders and the tendency of comments shared by recipients during the period of being flamed on social media. Second, content that has gone up in flames will not be extinguished for a long period, and the fact that it has gone up in flames will itself be consumed as content.



OS28-3 Protocol analysis for constructing Verbalizing Support System

Yuki Hayashi (Chiba University/NINJAL, Japan)

In this paper, I will show a result of an experiment. The experiment was conducted using voice recorder and the participants was required to report what they think orally. I have been conducting experiments to know how people express their feelings and thinking in their daily lives. Previously, I suggested the basical concept of Verbalizing Support System (in Japanese). However, in this experiment, I tried to show the process of verbalizing people's feelings and thinking. In addition, I'll show the revised concept of Verbalizing Support System compared with the previous paper.



©ICAROB 2023 ALife Robotics Corp. Ltd.

OS28-4 Story generation during appreciating an artwork based on an actual tale (Ugetsu-monogatari)

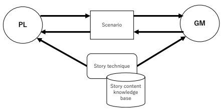
Akinori Abe (Chiba University, Japan)

In this paper, I will show a result of an experiment. The experiment was conducted using an artwork based on an actual tale (Ugetsu-monogatari). Actually the tale is rather old one. I have been conducting experiments to know how people can generate stories from the artworks. Previously I did not use artworks based on actual tales. Only the exception was artworks based on "Alice in Wonderland." These artworks were used to see how Visual Thinking Strategies (VTS) functions during art appreciation. Of course in the experiment, we asked participants to generate stories. However, in this experiment, I tried to see the process of the story generation during art appreciation.

OS28-5 Designing a Narrative Generation Game Based on the Russian Invasion of Ukraine

Jumpei Ono[1], Takashi Ogata[2] ([1]Aomori University, Japan), ([2]Iwate Prefectural University, Japan)

We have previously proposed an automatic narrative generation game. The purpose of this paper is to propose a game concept based on the framework of the automatic narrative generation game. The proposed game uses the ongoing Russian invasion of Ukraine since February 2022. The game based on an analog game in which participants advance a shared storyline, mainly through dialogue. The proposed game is a game with story generation, and as a method of application, we intend to support the creation of stories through the game.



OS29 Signals Processing and Automation II (6)

Chair Norrima Mokhtar (University of Malaya, Malaysia)
Co-Chair Heshalini Rajagopal (UCSI University, Malaysia)

OS29-1 Weather Forecast System for Mobile Devices

Hu Jiahao, Neesha Jothi, Chloe Thong (UCSI University, Malaysia)

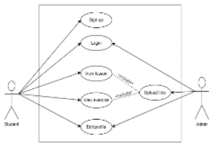
In the 21st century, a large number of personal computers appeared in daily life, and people began to be familiar with obtaining the required information through the Internet. Subsequently, as mobile development technology became increasingly popular and rapidly occupied the Internet market, the development of mobile terminals also made a breakthrough. At present, Android and iOS platforms dominate mobile platforms. Most applications are developed based on a single platform, and the development efficiency is not ideal. Therefore, cross platform technology has gradually entered the vision of developers, easing the problem of cross platform development. It provides opportunities to meet people's needs for real-time, accurate and diverse weather information.

©ICAROB 2023 ALife Robotics Corp. Ltd.

OS29-2 Development of a Novel E-Learning System for Improved Usability

Soo Yang Yew, Shaik Shabana Anjum, Shayla Islam (UCSI University, Malaysia)

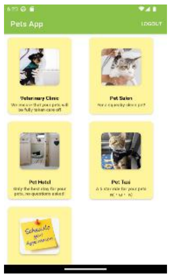
E-learning is a learning approach that combines organized instruction with the use of technological resources such as laptops, computers, and tablets. It allows students to study at any time and from any location with the help of Information Communication Technologies (ICT) that can connect instructors and pupils who are separated by thousands of miles. The criticality of this study is to fulfill students' and instructors' expectations of using virtual learning systems while having thrilling and interesting learning and teaching experience. This proposed system will help them to know more about the technology tools functioning and teach them a better understanding of using the system's features. Most lecturers and students faced challenges in utilizing e-learning systems with such factors being reviewed by some researchers.


Sign up
Login
View lesson
View exercise
Edit profile
Upload files
«include»
«include»
Student
Admin

OS29-3 A Development of a Prototype-based Mobile Pet Care Application

Gan Ai Leen, Shaik Shabana Anjum, Chloe Thong Chee Ling (UCSI University, Malaysia)

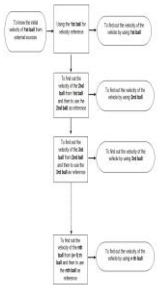
The keeping of domesticated pets has become a norm in today's society, and taking care of them has become a part of the humane culture. Therefore, this project aims to create prototype-based mobile pet care application that is able to make the user's life easier. The pet care application is based on the android application. Its objective is to provide the users a way to book the necessary appointment through the use of mobile phone without having to do it manually. Additionally, the application will include a chat function which can connect with the users and the admin. The application will enable the pet owners to get timely updates about their pets when they are away and under pet care. It will also provide options for grooming and healthcare facilities for the pet.


Pets App
LOGOUT
Veterinary Clinic
Pet Salon
Pet Hotel
Only the best stay for your pets, no questions asked!
Pet Taxi
Schedule your Appointment

OS29-4 Modeling of an Environmentally Independent and Contactless Speed Sensor for Measuring the Speed of Ships, Submarines, and Aircraft in Relation to the Ground Development of Image

Jakaria Mahdi Imam, Mohammad Aminul Islam, Norrima Binti Mokhtar, S. F. W. Muhamad Hatta
(Universiti Malaya, Malaysia)

We are presenting a theoretical and mathematical model for an environment-nondependent contactless speed sensor that can measure directly the horizontal speed of ships, submarines, and/or aircraft with respect to the ground. Currently, available standalone onboard speed sensors used in ships, submarines, and aircraft measure the speed of the vehicle with respect to the water or air; not the ground. In this paper, we have shown that a novel speed sensor can be designed by using small size objects dropped inside a vacuum chamber. Unlike the conventional methods, our proposed method can measure the speed of ships, submarines, and aircraft with respect to the ground directly. Another issue is that current speed sensors are environment-dependent, meaning they or their probes require some type of touch with the operational environment. The complete sensor assembly can be placed inside the vehicle.

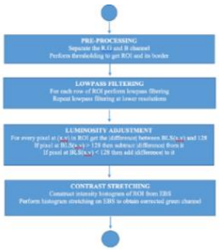

To know the initial velocity of 1st ball from external sources
Using the 1st ball for velocity reference
To find out the velocity of the vehicle by using 1st ball
To find out the velocity of the 2nd ball from 1st ball and then to use the 2nd ball as reference
To find out the velocity of the vehicle by using 2nd ball
To find out the velocity of the 3rd ball from 2nd ball and then to use the 3rd ball as reference
To find out the velocity of the vehicle by using 3rd ball
To find out the velocity of the nth ball from (n-1) th ball and then to use the nth ball as reference
To find out the velocity of the vehicle by using nth ball

©ICAROB 2023 ALife Robotics Corp. Ltd.

OS29-5 On Correcting Luminosity and Contrast of Retinal Images with Reflectance

Mofleh Hannuf AlRowaily[1], Hamzah Arof[1], Imanurfatiehah Ibrahim[1], Wan Amirul Wan Mohd Mahiyiddin[1], Norrima Mokhtar[1], Haniza Yazid[2]
([1]Universiti Malaya, Malaysia), ([2]Universiti Malaysia Perlis, Malaysia)

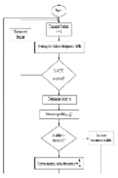
In this paper, we perform an automatic correction of luminosity and contrast of retina images. Retina images with varying level of reflectance taken from online databases are used to test the effectiveness of the method. The approach is implemented in 4 stages. In pre-processing, the three components of a color retina image are separated and only the green channel is processed further. Then the region of interest (eye region) and its border are marked. Next, the ROI of the green channel is filtered using a lowpass filter row by row to create a smooth background brightness surface. Three types of lowpass filters are used and their performances are compared. Using the background surface, the luminosity of the ROI is adjusted so that every pixel has the same brightness. Finally, the contrast of the ROI is improved by histogram stretching.


PRE-PROCESSING
Separate the R,G and B channel
Perform thresholding to get ROI and its border
LOWPASS FILTERING
For each row of ROI perform lowpass filtering
Repeat lowpass filtering at lower resolutions
LUMINOSITY ADJUSTMENT
For every pixel at (x,y) in ROI get the (difference) between BLS(x,y) and 128
If pixel at BLS(x,y) > 128 then subtract (difference) from it
If pixel at BLS(x,y) < 128 then add (difference) to it
CONTRAST STRETCHING
Construct intensity histogram of ROI from EBS
Perform histogram stretching on EBS to obtain corrected green channel

OS29-6 Rate Adaptation for Quality of Service Improvements in IEEE 802.11ax

Hazwani Zawawi[1], Wan Norsyafizan W Muhamad[1], Nani Fadzlina Naim[1] ([1]Universiti Teknologi, MARA)

This paper presents rate adaptation for Quality of Service (QoS) improvement in IEEE802.11ax. This study implemented a link adaptation technique which adapts the transmission data rate in IEEE 802.11ax WLAN. The main objectives of this paper are to design a link adaptation which adapts the transmission data rate based on radio channel conditions and to verify the effectiveness of the proposed algorithm. Simulation results show that the data rate adaptation technique offers better performance in terms of throughput and delay.



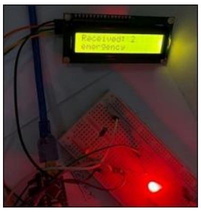
OS30 Embedded systems projects (3)

Chair Mastaneh Mokayef (UCSI University, Malaysia)
Co-Chair Takao Ito (Hiroshima University, Japan)

OS30-1 Data Transmission by Li-Fi in Coal Mining

Myadadha Goutham Reddy[1], Tan Jia Wei[1], Mandigiri Golla Jagan Kumar[1], Vytla Poojitha[1], Kaki Harshavardhan Goud[1], Wang Ji[1], MHD Amen Summakieh[1], Atefeh Mohammadpour[2], Mastaneh Mokayef[1*]
([1]UCSI University, Malaysia) ([2]California State University, USA)

In this paper, we proposed and designed a method to utilize the Li-Fi technology for data transmission which are equipped with sensors to provide real time situation stat to the coal miners to ensure their safety. One of the most dangerous gases in coal mining is the carbon monoxide (CO). Hence, specific type of sensor is used and connected to a core component which is the microcontroller and placed at different parts of the coal mine. When the concentration level exceeds a certain threshold value, the buzzer will be triggered, and the notification is sent to the central database to inform the workers. Li-Fi provides the solution for slow data transmission or data loss which enhances the safety and improve the working condition of labors in coal mining sectors.


Received: 2
emergency

©ICAROB 2023 ALife Robotics Corp. Ltd.

OS30-2 Embedded Table Tennis Ball Launcher with a Trajectory Path Analyzer for Junior Players

Mastaneh Mokayef[1], Lee Qi Jian[1], Bushra Naeem[2], Miad Mokayef[1], M.K.A Ahamed Khan[1], MHD Amen Summakieh[1]
([1]UCSI University, Malaysia) ([2]Balochistan University of Information Technology, Pakistan)

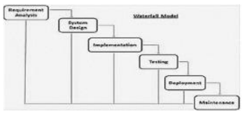
In this project, a pitcher machine as a personal Table Tennis trainer has been designed and the trajectory path analysis of the practical performance of player is analysed accordingly. In the proposed system, a camera is presented to track the table tennis ball. The obtained images from the camera are processed by OpenCV software and the flying trajectory is predicted based on the X and Y coordinates of a ball position to analyse the performance of player in each direction. The pitcher machine is then set via smart phone to the direction with weak performance to shoot more balls in that direction. However, the spinning serve is yet to be added in the proposed system, this project offers an acceptable platform for the early stage of the table tennis training. Experimental results show the acceptable performance analysis in complex mess environment and background.

Frame
tennis

OS30-3 Table Tennis Tournament Scores and Statistics Web Application

Belal Khaled[1], Mastaneh Mokayef[1*], Chin Hong Wong[2], Sew Sun Tiang[1], Wei Hong Lim[1], MHD Amen Summakieh[1], Miad Mokayef[1]
([1]UCSI University, Malaysia) ([2] Fuzou University, China)

The objective of this project is to develop a functional web application that will retrieve or store table tennis tournament statistical data and visualize them using tables, bar charts, pie charts, histograms as a medium of demonstration for the user. Moreover, the application will perform specific analysis on the scores and display insightful data about the tournament for league structures investigation. The developed system will include a database to store and retrieve data for display on the user interface. The development of the project is fully stacked (front-end and back-end), so it is built with the appropriate web technologies to function in the background (PHP, MySQL, Apache) while displaying results (HTML, CSS, JavaScript) on the page for the user. The project is managed using effective project management methods to plan, design, implement, develop, and maintain the application.



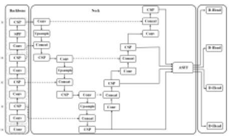
OS31 Image Processing and Reconstruction (5)

Chair Chair Xiaoyan Chen (Tianjin University of Science and Technology, China)
Co-Chair Shiming Wang (Tianjin University of Science and Technology, China)

OS31-1 Small-Target Detection Based on YOLOX

Keying Ren[1], Xiaoyan Chen[1*] (Tianjin University of Science and Technology, China)

With the development of drone, Small-Target detection has become a hotspot of current research. In this paper, the network of Small-Target detection is based on YOLOX is studied. There are a lot of small targets in the images taken by UAV, which brings great difficulty to the detection task. A four-scale adaptive spatial feature fusion pyramid is used to filter the conflicting information between different scales and improve the expressive accuracy of the small target features. Experiments show that the proposed method has a mAP average accuracy of 32.83% in the self-built dataset, which is 2.53% higher than that of YOLOX-s, 1.53% higher than that of Yolov6-s, and 5.73% higher than that of YOLOv5-s.

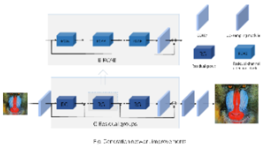


©ICAROB 2023 ALife Robotics Corp. Ltd.

OS31-2 Super Resolution Reconstruction Model Based on Attention Mechanism and Generative Adversarial Network

Xia Miao, Xiaoyan Chen (Tianjin University of Science and Technology, China)

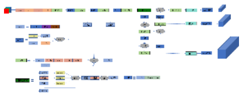
With the continuous improvement of the screen quality of various electronic devices, the public's pursuit of fine visual effects is also increasing. However, due to the degradation phenomenon in the actual imaging and transmission process, there is a contradiction between the obtained low resolution image and the high quality screen. The purpose of image super resolution reconstruction is to recover high frequency details from low resolution images containing little information, so as to improve the visual effect of images. Based on the problems existing in the current Adversarial network generation methods, this paper constructs Channel Attention (CA) and Generative Adversarial Networks. CA-SRGAN is applied to image super resolution reconstruction, and the image super resolution reconstruction module is implemented.



OS31-3 Pedestrian-vehicle detection based on YOLOv7

Xiaoyan Chen, Zhihui Chen (Tianjin University of Science and Technology, China)

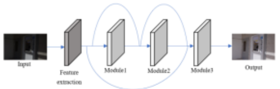
With the continuous development of information technology, it is constantly improving. The security industry is constantly expanding. Therefore, object detection has become more important. In this paper, a large number of experiments have been made to improve the detection accuracy of pedestrians, vehicles and license plates in cities. The improved YOLOv7 algorithm was used to conduct a large number of experiments on the urban pedestrian vehicle dataset. Experiments show that new improvements have improved detection accuracy.



OS31-4 A lightweight, faster, more generalized low light image enhancement network

Yutao Jin, Xiaoyan Chen (Tianjin University of Science and Technology, China)

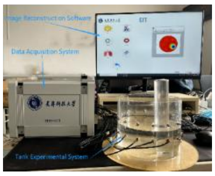
Existing methods based on deep learning have achieved great success in low light image enhancement. However, these methods generally have a very large amount of computation and poor generalization ability for low light images in different scenes. In order to solve this problem, this paper explores a wonderful way to assist training by introducing auxiliary blocks in the training process. These blocks can connect the feature map input to each stage with the input to the first stage of the model to explore whether the model converges. In the prediction phase, due to the superior capabilities of the network, we use only one basic block for inference, so as to greatly reducing the computational cost. Thanks to the flexibility of the network, the model can easily handle low light images in a variety of complex environments.



OS31-5 An Improved Landweber Method for Electrical Impedance Tomography

Qian Wang, Xiaoyan Chen (Tianjin University of Science and Technology, China)

Electrical impedance tomography (EIT) is a non-destructive monitoring technique. Its image reconstruction problems have serious nonlinear and ill-posed nature, which leads to the low spatial resolution of the reconstructed images. The iterative algorithm is an effective method to deal with imaging inverse problem. However, the existing iterative imaging methods have many problems, such as many iterations and unstable convergence. To solve this problem, this paper proposes an improved Landweber iterative image reconstruction method. By adding a regularization term to the objective function of the inverse problem of EIT, the method improves the convergence speed. The physical experiment results show that the improved Landweber method is superior to the Landweber method in convergence speed.



©ICAROB 2023 ALife Robotics Corp. Ltd.

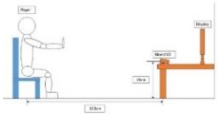
OS32 AR, VR and Digital Contents

Chair R.P.C. Janaka Rajapakse (Tainan National University of the Arts, Taiwan)
Co-Chair Yoshimasa Tokuyama (Transworld University, Taiwan)

OS32-1 A Kinect-based Augmented Reality Game for Arm Exercise

Toshimasa Tokuyama (Transworld University, Taiwan)
R.P.C. Janaka Rajapakse (Tainan National University of the Arts, Taiwan)

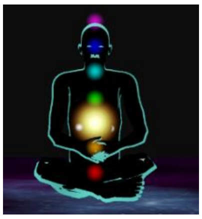
Augmented reality (AR) is where 3D virtual objects are integrated into a 3D physical environment in real-time. Augmented reality applications such as medical visualization, maintenance and repair, robot path planning, entertainment, military aircraft navigation, and targeting applications have been proposed. This paper introduces the development of an augmented reality game that allows the user to carry out arm exercises using a natural user interface based on Microsoft Kinect. The system has been designed as an augmented game where users' hands are in a world augmented with virtual objects generated by computer graphics. The player is sitting in a chair, grasping the yellow stars displayed on the stage. It encourages the activities of a large number of arm muscles which will prevent decay. It is also suitable for rehabilitation.



OS32-2 Development of EEG-based VR Application for Chakras Guided Meditation

Chien-Tung Lin, R.P.C. Janaka Rajapakse (Tainan National University of the Arts, Taiwan)

In a rapidly changing world, everyone is under a lot of pressure, both visible and invisible. Therefore, stress causes several diseases. Meditation through mindfulness is one of the ways to relieve stress. This paper studies how to properly meditate in a VR environment and increase the effect of chakra meditation by real-time EEG data. The theory of chakras has existed for thousands of years and still exists in modern society. The balance of their energy is essential to everyone. Meditation on the chakras is beneficial to the health of the body. But for ordinary people, chakra is invisible and unimaginable. It is easy to associate it with religion. Therefore, we developed the application in a VR environment. Users can sit inside a non-distracting environment with chakra visualization. It's easier for ordinary people to understand chakras meditation. The application can understand the user's condition by the EEG data and change to the next chakra in sequence.



OS32-3 A Study on Flower Patterns of Temple Cut-and-Paste Decorations based on L-system

Meng-Fan Huang, Tzu-Hsien Yuan, R.P.C. Janaka Rajapakse (Tainan National University of the Arts, Taiwan)

“Cut-and-paste' is an architectural skill used in traditional Chinese temples to make various decorations by pasting tiles and glass. This traditional craft has a history of more than 400 years, and each piece of work has noble artistry. This kind of craftsmanship is usually shown on roofs and walls of temples, so it is also called "art on the roof." However, due to the time-consuming or labor-intensive production and the aging of many craftsmen, this traditional handicraft is slowly disappearing. Although the paste decoration of the Temple has been registered and preserved by the Taiwan government many years ago, there are very few related studies, and many works are also in a state of being damaged and difficult to repair. Therefore, this study will discuss the decorative flower patterns in the paste decoration of the temple, analyze the growth of plants with the L-System, and then use TouchDesigner to create interactive works, hoping to transform traditional handicrafts into public art so that more people realize the importance of craft preservation.

©ICAROB 2023 ALife Robotics Corp. Ltd.

OS32-4 Pass-By: Development of Pedestrian Counts-based Art Installation for Passive Interaction.

Jianlin Zhong, R.P.C. Janaka Rajapakse (Tainan National University of the Arts, Taiwan)

The trees in the streetscape grow as pedestrians pass by; we are always taking away and bringing in our surroundings without realizing it, perhaps some of the changes are very small, and we only notice them when they have grown to colossal size. This work presents a pedestrian counts-based art installation for multi-user passive interaction. This work focuses on the impact of people on their surroundings which concerned a contemporary street scene projected in an open space and captured pedestrians' pass-by counts using an Arduino ultrasonic sensor. Based on the measured counts of pedestrians passing by, the developed application controls the color of the street scene and gradually decreases as the pedestrians cross. Using the L-system application, the trees in the street scene grow progressively as the pedestrians cross.

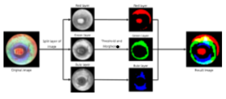
©ICAROB 2023 ALife Robotics Corp. Ltd.

GS abstracts
GS1 Machine Learning & Neural Network & Artificial Life (3)
Chair Masato Nagayoshi (Niigata College of Nursing, Japan)

GS1-1 Defect Solder Classification in Print Circuit Boards using Machine Learning

Watcharin Tungsuksan, Hattapat Silangren, Phattharaporn Iamcharoen, Wisanu Jitviriya
(King Mongkut's University of Technology North Bangkok, Thailand)

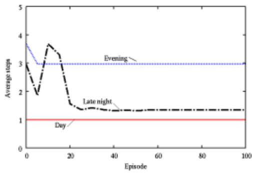
This research proposes the solder inspection using the digital image processing technique and machine learning base with our machine vision prototype. There are five classes of classifying solder, including acceptable, short circuit, insufficient, blow hole and too much of solder. Automatic Optical Inspection (AOI) is used for the light source in the designed prototype and industrial camera which are installed on the mini-CNC. For the algorithm, this research applies the scanning line of binary image for detecting short circuit defection and the Random Forrest model for classifying other defects. According to the experiments, the system can classify the defect types for two classes (acceptable and unacceptable types) and five classes as 89% and 71% of accuracy, respectively



GS1-2 A dynamic nurse scheduling using reinforcement learning: Dealing with various sudden absences of a nurse

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

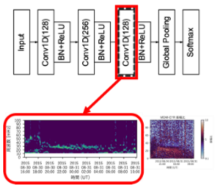
In nurse scheduling, whereby work schedules for nurses are created, it is very difficult to create a work schedule that satisfies all the various requirements. Hence, various studies have been conducted on the nurse scheduling problem. However, for practical use, adjustments including various constraints and evaluation values are required, as the created shift schedule is often not practical as it is. Therefore, we have proposed a work revision method using reinforcement learning on a constructive nurse scheduling system. Furthermore, we have proposed an extension of the proposed method to dynamic nurse scheduling, in which the work schedule is revised or rescheduled when an absence occurs. In this paper, we confirm whether or not the extended method can be used to create a work schedule that is feasible in various sudden absences of a nurse.



GS1-3 Classification of Time Series Data Obtained by the Satellite by Using Rule-Based and Machine-Learning Methods

Satoko Saita[1], Mariko Teramoto[2], Kentarou Kitamura[2]
([1]National Institute of Technology, Kitakyushu College, Japan), ([2]Kyushu Institute of Technology, Japan)

Scientific observations by nano-satellites need to reduce downlink data by onboard data preprocessing because of its communication volume constraints. We tried to classify time series data of geomagnetic field obtained by the SWARM satellite in order to integrate machine-learning techniques, and to determine the most appropriate method for onboard classification of a phenomenon in the geomagnetic field. The classifications have been executed by using rule-based, K-means, and combined CNN methods. The experimental results demonstrated the effectiveness of machine-learning model with LSTM networks. This study will contribute to development of machine-learning capable nano-satellites.



©ICAROB 2023 ALife Robotics Corp. Ltd.

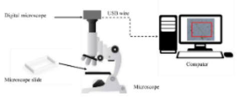
GS2 Image Processing I (5)

Chair Watcharin Tangsuksant (King Mongkut's University of Technology North Bangkok, Thailand)

GS2-1 Microalgae Detection by Digital Image Processing and Artificial Intelligence

Watcharin Tangsuksant, Pornthep Sarakon
(King Mongkut's University of Technology North Bangkok, Thailand)

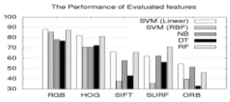
This article presents a technical approach to the video computer analysis, to automatically identifying the two most frequently identified microalgae in water supplies. To handle some difficulties encountered in image segmentation problem such as unclear algae boundary and noisy background, we proposed a deep learning-based method for classifiers or localizers to perform microalgae detection and counting process. The system achieves approximately 91% accuracy on Melosira and Oscillatoria detection, which around 4.82 seconds per grid. (Intel Xeon(R) CPU E5-2667 12 CPU at 2.66GHz and 32.0GB RAM, NVIDIA Quadro K5200 with 2304 CUDA cores). The system can significantly reduce 33.33 - 55.56% of the counting time when compared with the visual inspection of manual methods, and eliminate the error due to the human fatigue.



GS2-2 SmartCropPlanting: IoT-Based Mobile Application for Hydroponic System

Sung Jun Kyu[1], Chit Su Mon[2], Kasthuri Subaramaniam[3]
([1,3]UCSI University, Malaysia) ([2]Heriot-Watt University Malaysia, Malaysia)

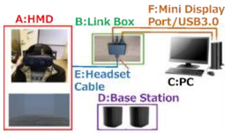
Hydroponic crop production is challenging for farmers and gardeners; they must continuously monitor and control the crop environment to achieve or maintain the best plant growth. Climate changes and disease/pests might occur anytime at anywhere which will consequence in crop-damaging. Studies show that slight rises in temperature from 1°C to 4°C can resulting in a decrease in the production of 10 to 41%. Another study shows that depending on the severity of leaf spot disease it will damage the crop by 10 to 50%. Therefore, this study proposed to improve the current hydroponic system by implementing the Internet of Things and Image Processing technique to optimize and reduce the infeasible tasks with mobile applications. This study has conducted a mixed-mode method with home gardeners, farmers, and related experts. The results show that 80% of respondents agree that SmartCropPlanting will improve the productivity and efficiency of crop planting with IoT and Image Processing techniques.



GS2-3 Detection of Eye Misalignment Using an HMD with an Eye-tracking Capability

Yoki Nagatomo, Noriyuki Uchida, Takuya Ikeda, Kayoko Takatsuka, Masayuki Mukunoki, Naonobu Okazaki
(University of Miyazaki, Japan)

In this study, we implemented the Cover-Test, a test method for diagnosing eye misalignment using a head-mounted display with an eye-tracking capability. Specifically, we created a virtual examination environment in a VR space. The eye-tracking technique collected eye movements immediately after the covering or uncovering of the eyes. Thus, we calculated the amount of eye deviation and developed a system to determine the presence and magnitude of strabismus and heterophoria. We then assessed the system in the verification experiment by examining the consistency between the judgment results provided by this system and the clinical evaluation approach with the Maddox rod. The result was that we could verify the horizontal eye movements more accurately.

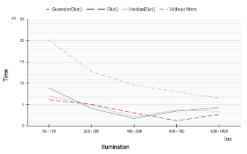


©ICAROB 2023 ALife Robotics Corp. Ltd.

GS2-4 Using OpenCV for real-time image recognition through augmented reality devices

Gabdullina Dinara, Zykov Evgeniy, Kugurakova Vlada (Kazan Federal University, Russia)

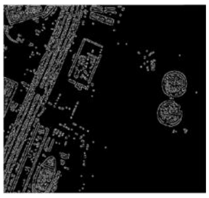
This article describes the peculiarities of writing libraries and the rapid development of augmented reality applications on the Unity platform and the OpenCV open-source library. The application of OpenCV functions to work with augmented reality objects is discussed in detail. On their basis, algorithms and methods were developed and an open library was created that implements the stable recognition of objects in various conditions by means of binding to the marker and their visualization on the augmented reality devices. On the basis of the developed library, a complete application was created to illustrate the possibility of applying virtual reality technologies for optical control of the assembly of radio electronic boards and for personnel training.



GS2-5 A Structure Pattern Extraction by Using Morphological Component Analysis in the Aerial Image Edge Detection

Wataru Oshiumi, Ankur Dixit, Hiroaki Wagatsuma (Kyushu Institute of Technology, Japan)

T An automated extraction of areas of interest from aerial images is a central issue of map makers. The road skeleton is one of the most frequent targets and then the prefiltering of road structures is highly important; however, the isolation of man-made structures from a natural landscape is technically difficult, because image parts to represent natural geographical features are not uniform patterns. In the present study, we focused on the function of Morphological Component Analysis (MCA) method to extract structural patterns when dictionaries were appropriately given and demonstrated an effectiveness of the edge detection if the prefiltering was done by using MCA. In our computer experiments, MCA decomposes the image into two patterns with Curvelet transform and Local Discrete Cosine Transform (LDCT) dictionaries. This approach will explore extensive possibilities of structural data extraction from complex images as an automated method.



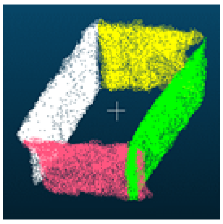
GS3 Image Processing II (4)

Chair Joo Kooi Tan (Kyushu Institute of Technology, Japan)

GS3-1 3D Point Cloud Registration and Segmentation of Reflective Metal Objects Using Go-ICP and Improved RANSAC

Kairi Morita, Ziyue Liu, Jing Cao, Seiji Ishikawa, Masuhiro Nitta, Joo Kooi Tan
(Kyushu Institute of Technology, Japan)

Registration and segmentation of 3D data are necessary in many fields, such as factory automation, automated driving, or even in the medical field. However, the technique is generally applied to non-metal objects. One of the problems of registration of a metal object is that the point clouds representing a metal object contain many outliers and missing points because of its reflective nature. This makes the accuracy of the registration and segmentation degrade. In this paper, we propose registration and segmentation techniques that are robust to outliers. For registration, we use globally optimal Go-ICP (Global optimal - Iterative Closest Point) algorithm considering goodness of a combination of point cloud sets to escape from convergence to a local solution. In segmentation, we address the problem of RANSAC generating false segments consisting of nearly identical multiple planar points by improving RANSAC. We use three kinds of metal tray to show the effectiveness of the proposed technique.



©ICAROB 2023 ALife Robotics Corp. Ltd.

GS3-2 Development of a Safe Walking Assistance System for Visually Impaired Person Using MY VISION — Estimation of a Safe Passage from Sidewalk Information Based on Transfer Learning of VGG-16 Network

Takumi Yokote, Joo Kooi Tan (Kyushu Institute of Technology, Japan)

In recent years, the number of visually impaired person has been increasing year by year, and outdoor accidents have also been increasing when they go out. It is difficult to detect hazards on sidewalks even with the currently popular technique, such as a semantic segmentation technique or YOLO, because sidewalk situations are complicated and change frequently. For this reason, we propose a method of recognizing sidewalk situations from a self-viewpoint video called MY VISION. Conventional methods detect objects surrounding the sidewalk by learning the objects' features beforehand and guiding visually impaired person according to the position/direction of the detected object. The proposed method neither learns objects nor detects objects. We focus on sidewalk situations and use a multi-class classification based on transfer learning of VGG-16 to guide visually impaired person' walk according to three kinds of sidewalk information to ensure more safety. The effectiveness of the proposed method was confirmed by experiments.

GS3-3 Detection of Fallen Persons and Person Shadows from Drone Images

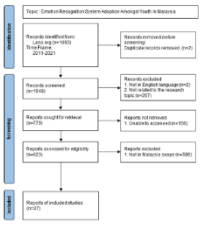
Taisei Ono, Haruka Egawa, Seiji Ishikawa, Joo Kooi Tan (Kyushu Institute of Technology, Japan)

In recent years, the development of automatic search methods based on aerial images taken by drones has been attracting attention in order to prevent secondary disasters and to perform rescue operations quickly in the search for victims of natural disasters. Although various methods exist for automatic human detection for search, they are based on the assumption that the background area of a person captured by a drone camera is a uniform ground in which only those persons who require rescue exist without any shadows or trees. In this paper, we propose a method of automatic detection of both fallen persons and person shadows, or trees on the ground. The method is a combination of Ri-HOG and Ri-LBP features to search for fallen persons. These features are robust to rotation. We then employ Grabcut and brightness values to detect shadows. The effectiveness of the proposed method was verified by experiment.

GS3-4 A Systematic Literature Review on Emotion Recognition System in Malaysia

Muhammad Nadzree Mohd Yamin, Kamarulzaman Ab. Aziz, Tan Gek Siang, Nor Azlina Ab. Aziz
(Multimedia University, Malaysia)

Artificial intelligence (AI) is an important technology that evolved from theories into tangibility with significant impacts and applications across sectors. AI is one of the key technologies that gave rise to the fourth industrial revolution (IR 4.0). One key subcategory of AI is the automated emotion recognition system (ERS); the application of AI to recognize human emotional states. ERS can be embedded in our daily lives. The importance of ERS will become more significant as we move towards the fifth industrial revolution (IR 5.0). ERS is a bridging technology from IR 4.0 into IR 5.0. Crucial for this, is good adoption or diffusion levels of ERS amongst society. Therefore, there is a need to understand the factors that affect the adoption of ERS. This paper seeks to establish and discuss the current ERS research landscape in Malaysia through systematic literature review covering works over a decade; from 2011 to 2022.


Identification
Screening
Included
Topic : Emotion Recognition System Adoption Amongst Youth in Malaysia
Records identified from: Lens.org (n=1050) Time Frame: 2011-2021
Records removed before screening: Duplicate records removed (n=2)
Records screened (n=1048)
Records excluded 1. Not in English language (n=2) 2. Not related to the research topic (n=267)
Reports sought for retrieval (n=779)
Reports not retrieved 1. Unable to accessed (n=156)
Reports assessed for eligibility (n=623)
Reports excluded 1. Not in Malaysia scope (n=586)
Reports of included studies (n=37)

©ICAROB 2023 ALife Robotics Corp. Ltd.

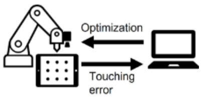
GS4 Robotics I (4)

Chair Amornphun Phunopas (King Mongkut's University of Technology North Bangkok, Thailand)

GS4-1 A Basic Study of Hand Eye Calibration using a Tablet Computer

Junya Sato, Takayoshi Yamada, Kazuaki Ito (Gifu University, Japan)

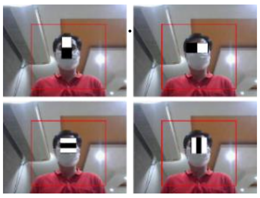
In this study, we describe a hand eye calibration method for calibrating an attached camera and a handmade end-effector to a robot by utilizing a tablet computer. For accurate manipulation, calibrating the camera, robot hand, and end-effector is important. However, general hand eye calibration methods do not consider dimensions of an end-effector and attached position to the robot in the calibration process. Thus, additional calibration of the attached end-effector for accurate manipulation is required when there is no detailed information of the end-effector such as a handmade one. To avoid this, we propose a novel method using a tablet computer. By iterating the touching black dots displayed in the computer and optimizing robot parameters to minimize the touching error, the hand eye calibration is achieved without the end-effector information such as the dimensions and attached position.



GS4-2 Human Detection with Uprisen Angle of a Camera for the Service Robot

Watcharin Tangsuksant, Amornphun Phunopas, Pornthep Sarakon, Aran Blattler
(King Mongkut's University of Technology North Bangkok, Thailand)

In order to improve the intelligent service robot, visual perception is crucial. This paper presents human detection for a service robot. The camera is installed at the robot's uprisen angle of around 30 degrees. Two feasible algorithms for the real-time detection between the Haar cascade and Single Shot Detector (SSD) algorithm are compared. This research collects the training data as 1,000 images of the uprisen angle, the different views of the human within 60 cm of the user range from a camera to a human. The proposed method has a higher performance than traditional Haar cascade and SSD. Therefore, the proposed method is suitable and feasible to deploy to the service robot.



GS4-3 Synthesis of Drive Systems of Flapping and Feathering Motions for Bird-like Robot using Twist Drive Mechanism

Jun Iwao, Hiroshi Ohtake (Kyushu Institute of Technology, Japan)

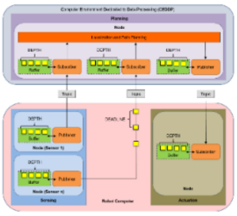
In the research field of flying robots, many studies on flapping-wing aircraft have been conducted in recent years. In this research, we focus on a bird-shaped flapping wing aircraft. In our previous research, we have developed a robot that mimics the musculoskeletal structure of an actual bird using a twist drive mechanism, and have achieved two types of wing motions, which are flapping and feathering, independently of each other. In this study, we have synthesized the drive systems of flapping and feathering, and have succeeded to interlocked two motions. In addition, we have covered the wings with a membrane to verify the wing motion under the condition of a membrane wing, which is strongly affected by air resistance. The flapping motion with twisting of the hands, such as actual birds do in flight, has been achieved.

©ICAROB 2023 ALife Robotics Corp. Ltd.

GS4-4 Optimization Algorithm for Balancing QoS Configuration in Aggregated Robot Processing Architecture

Abdul Jalil, Jun Kobayashi, Takeshi Saitoh (Kyushu Institute of Technology, Japan)

This study aims to design an optimization algorithm to find the optimal value of DEPTH and DEADLINE in the Quality of Service (QoS) configurations while the Robot Operating System 2 (ROS 2) nodes are communicated using the RELIABLE and KEEP_LAST options. We implemented this optimization to improve the quality of data transmission in Aggregated Robot Processing (ARP) architecture. Strict reliability is not guaranteed if ROS 2 uses the RELIABLE and KEEP_LAST options to transmit the message data between the publisher and the subscriber. In addition, unbalancing DEPTH and DEADLINE configurations can affect the high latency time of message data transmission and packet loss in RELIABLE connections. The result of this study shows that the optimization algorithm can determine the optimal value of DEPTH and DEADLINE to balance the QoS configuration to improve robot data transmission in the ARP architecture.



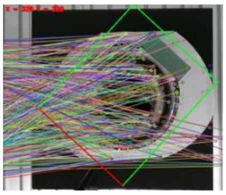
GS5 Robotics II (4)

Chair: Takayoshi Yamada (Gifu University, Japan)

GS5-1 Research on robotic assembly of gear motors (Stator recognition using keypoint matching and stator insertion using contact position estimation)

Yasumoto Imai, Takayoshi Yamada, Junya Sato (Gifu University, Japan)
Toshiki Hayashi, Shota Aono (Tsubakimoto Chain Co., Japan)

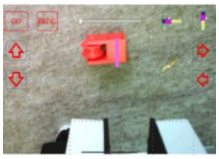
Assembly tasks in various fields, which require dexterous skills, are still often performed manually. Humans use information such as vision and hand sensation to accomplish these tasks. This paper describes a robotic assembly system for gear motors. Two main methods, key point matching and contact position estimation, are used to realize the gear motor assembly work, which is the insertion of the stator into the reducer. Keypoint matching is used to recognize the position and orientation of the part using a camera. Contact position estimation is used to detect contact between parts in the insertion process using a force sensor to prevent failure of the insertion operation. An experimental system using these methods is developed to achieve automatic stator insertion.



GS5-2 Robot Arm Operating Interface for Easy Grasping by Specifying the Gripping Width of End-effector

Rio Takeuchi, Laijun Yang, Norihiko Kato (Mie University, Japan)

In Japan, cervical cord injuries affect roughly half of all patients with spinal cord injuries. High-Level spinal cord injuries make it particularly difficult for patients to pick up objects from the ground because of the functional impairment of the trunk. To help these people become more independent, welfare robot arms have been developed recently. In our lab, an interface has been proposed to control the robotic arm by drawing a line on a touch screen in order to grasp an object on the floor and deliver it to the user. This technique has a fixed closing width for grasping, making it challenging to use when the object's sizes change. Therefore, we proposed a method that allows the adjustment of the closing width according to the drawn line on the touch panel interface. As a result, we found that the grasping was still possible even the object was altered.



©ICAROB 2023 ALife Robotics Corp. Ltd.

GS5-3 3D Real-Time Conversational Virtual Agents System: Do Facial Expressions and Camera Angles Persuade Human

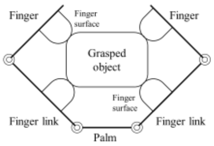
Pinkie Anggia, Kaoru Sumi (Future University Hakodate, Japan)

This paper aims to uncover state-of-the-art on persuasion through a dialogue interaction system between human and virtual agents. 37 participants interacted with a virtual agent system from facial expressions and camera angles in three conditions. Through empirical measurements using the Big Five theory, we discover that virtual agents' friendliness becomes a solid persuading factor to our predominantly extroverted participants. The experiment deduces that openness, conscientiousness, and extraversion personalities are easily persuaded by our virtual agents, unlike reversed-conscientiousness and neuroticism personalities.

GS5-4 Quasi-static Stability Analysis of Frictionless Planar Enveloping Grasps (Analysis of curvature effects at contact points)

Takayoshi Yamada, Junya Sato, Kazuaki Ito, Hidehiko Yamamoto (Gifu University, Japan)

Humans can grasp and manipulate various types of objects dexterously. In recent years, in order to obtain appropriate grasps, deep learning methods are introduced. On the other hands, in one of traditional ways, grasp stability based on potential energy of the grasp system replaced with an elastic system is investigated in order to derive grasp evaluation. In this paper, we discuss grasp stability of frictionless planar enveloping grasps. The stiffness matrix of the grasp system is derived. The grasp stability is evaluated by the eigenvalues of the matrix. We show that the matrix depends on grasp positions, grasp forces, local curvatures, etc., at contact points. Moreover, we analyze curvature effects on the grasp system by differentiating the matrix by the curvatures.



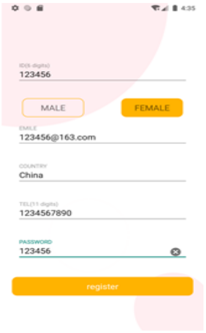
GS6 Applications I (5)

Chair Minoru Kumano (University of Miyazaki, Japan)

GS6-1 My Tally -A Personal Book Keeping Mobile Application

Li Zhihan[1], Abdul Samad Shibghatullah[1], Nur Hazirah Hamdan[2]
([1]UCSI University, Malaysia), ([2]Universiti Teknologi MARA,UiTM, Malaysia)

Bookkeeping has a long history and is charming and colorful. Throughout the ages, people have used the experience of abacus bookkeeping; Later, we used pens and books to keep accounts, supplemented by calculation results. In the information age of the 21st century, the popularization of computer software makes us turn to excel or other desktop software with powerful analysis function for computer bookkeeping. However, all the above methods are limited by time and space. My Tally focuses on the construction of Android environment, functionally designed many aspects involved in people's daily consumption. The personal bookkeeping application designed in this paper not only reflects the convenience of recording income and expenditure information, but also analyses the financial situation, so that people can record information faster and in real time and provide users with the ability to record and manage daily financial events. Compared with the traditional accounting methods, the personal bookkeeping application designed in this paper has the characteristics of security, portability and practicability.



ID(6-digits)
123456

EMILE
123456@163.com

COUNTRY
China

TEL(11 digits)
1234567890

PASSWORD
123456

©ICAROB 2023 ALife Robotics Corp. Ltd.

GS6-2 A Survey on Charitable Acts, Challenges and Using Charitable Mobile Application

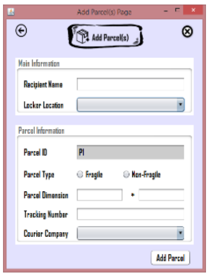
Nan Pepin[1], Abdul Samad Shibghatullah[1], Kasthuri Subaramaniam[1], Nur Hazirah Hamdan[2]
([1]UCSI University, Malaysia), ([2]Universiti Teknologi MARA, Malaysia)

A donation is a gift that is given voluntarily or sincerely without the expectation of receiving something in return. Donating in contemporary terminology is better known by the term philanthropy. Today, the emergence of charity platforms resulted in a modern society increasingly making donations in cyberspace, which are referred to as online charity. Although the use of digital technology allows for philanthropic activities to be carried out, there are still shortcomings in this online system. This paper aims to study mobile charitable applications among users. The survey has been conducted via Goggle form and distributed to 110 participants through the internet. The findings show that there are not many applications of this sort being produced by other developers and there is a need to develop a mobile application that can help donators and receivers. The proposed mobile applications should be faster, easy to manage and transparent.

GS6-3 Online Parcel Management System (PMS) for Small and Medium Company

Chew Cheng Jin[1], Kasthuri Subaramaniam[1], Abdul Samad Shibghatullah[1], Salini Devi Rajendran[2], Nur Hazirah Hamdan[3]
([1]UCSI University, Malaysia), ([2]Taylor's University, Malaysia),
([3]Universiti Teknologi MARA,UiTM, Malaysia)

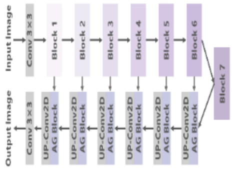
This research is going to gather more information of the recipient viewpoint on how Parcel Management System can help current system to reduce the workload and manpower to handle all these parcels, increase efficiency and save time and effort which make the staff of the reception counter easier to work and no need to use back the old method to record the details of the parcel. The result of the survey was analyzed in detail and find out the problem of the current parcel management system. The system development approach that the researcher used for this system is Rapid Application Development (RAD) Model. The reason that the researcher chooses this methodology is because there is lack of time and money to do this project. There is four of the stages in Rapid Application Development (RAD) methodology life cycle, which is requirement planning phase, user design phase, construction phase and transition phase. There are several phases in this project such as planning, requirement gathering and analysis, design and implementation and there is also have some deliverable in those phases.



GS6-4 Automatic Classification Method for Plastic Bottles and Caps Using Multi Attention Eff-UNet

Shunsuke Moritsuka, Tohru Kamiya (Kyushu Institute of Technology, Japan)

In Japan, the increase of general waste has become a problem. One solution is to recycle used plastic bottles. However, plastic bottles are thrown away with their caps still attached, requiring the caps to be removed manually. To solve this problem, we develop a method to automatically identify plastic bottles and caps using deep learning. In this paper, we add an attentional structure to Eff-UNet to build a model with reduced over-detection, as shown in the figure on the right. To extract features for objects of various sizes, we also propose a model that combines different numbers of blocks of the model shown on the right. In fact, experiments were conducted on an image of a plastic bottle, and this model succeeded in improving accuracy by 1.7% compared to Eff-UNet.

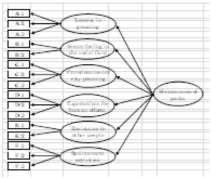


©ICAROB 2023 ALife Robotics Corp. Ltd.

GS6-5 Research on the structure of consciousness of people who maintain and manage parks

Jun Sanbuichi[1], Minoru Kumano[1], Toru Hiraoka[2]
([1]University of Miyazaki, Japan), ([2]University of Nagasaki, Japan)

Due to population decline, unused vacant land is expected to increase in the region in the future. So, it is considered important to have local residents maintain and manage local spaces. However, consciousness indicators of how to encourage maintenance and management are not clear. Therefore, the purpose of this study is to clarify the consciousness structure of people who maintain and manage park. The investigation method is a questionnaire to park volunteers. In the consciousness analysis of park volunteers, factor analysis and covariance structure analysis were performed. As a result, it was proved that six consciousness (Interest in greening etc.) are related to the comprehensive consciousness index of park maintenance behavior.



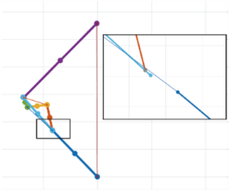
GS7 Applications II (3)

Chair Hiroaki Wagatsuma ([1]Kyushu Institute of Technology, [2]RIKEN CBS, Japan)

GS7-1 An Accuracy Evaluation of Multibody Dynamics for the Knee Support Exoskeleton Model with Respect to Implicit Methods for Numerical Integration

Shintaro Kasai, Hiroaki Wagatsuma (Kyushu Institute of Technology, Japan)

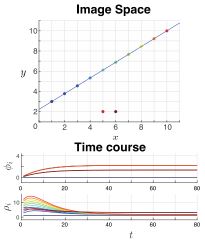
Numerical integration takes an important role to analyze models described by ordinary differential equations and it largely contributes to an assurance of the accuracy in displacement analyses of multibody dynamics if the model consists of several bodies especially with dynamic components such as springs and dampers. Exoskeletal assistive devices require flexible materials for absorbing reaction forces from human joints, which implies an inevitable necessity of an accurate evaluation of elastic effects in the model. In the present study, we introduced implicit methods for numerical integration to analyze the model under the formulation of multibody dynamics and computer experiments demonstrated results from explicit and implicit methods for the numerical integration as a comparison analysis. It can improve a degree of accuracy in inverse dynamics even in the theory of flexible Multibody Dynamics (fMBD).



GS7-2 A Basic Concept of the Nonlinear Oscillator-Based Hough Transform Implementation to Improve the Voting Procedure in the Scheme of Continuous Dual Spaces

Amarbold Purev, Hiroaki Wagatsuma (Kyushu Institute of Technology, Japan)

Hough Transform (HT) is a well-known algorithm to detect arbitrary lines in the image. The algorithm of HT consists of two processes mapping to the parameter space and voting to find a solution. The former process is executed in the continuous space, while the latter process relies on counting in the mesh as a discrete space, which increases susceptibility to noise. In the present study, we proposed attractor dynamics in coupled nonlinear oscillators instead of discrete voting. In our computer experiment, image pixels were mapped to the parameter spaces, and they were attracted to the crossing point of lines in the space. The approach may contribute not only to line detections in the image but also to other shape detections by using a consistent mathematical formulation.

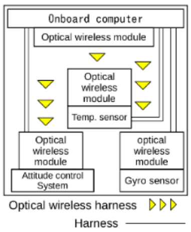


©ICAROB 2023 ALife Robotics Corp. Ltd.

GS7-3 Survey on Harness Design for CubeSats: Understanding the Constraints of CubeSats Design and Toward an Optical Wireless Bus for CubeSats

Masahiro Tokumitsu[1], Masatoshi Tsuji[2], Jun Nakaya[3]
([1]National Institute of Technology (KOSEN), Yonago College, Japan)
([2]National Institute of Technology (KOSEN), Kagawa College, Japan)
([3]Aichi Institute of Technology, Japan)

CubeSats of nano-satellites have attracted the attention of space scientists and engineers seeking to observe the space environment and develop innovative technologies in space engineering. The CubeSat is a class of miniaturized satellites with a form factor based on a 10 cm cube. However, the dimensional constraints of CubeSats restrict the embedding of relatively large mission devices, such as attitude control systems, into the satellites. Moreover, the harness used to transfer the data and supply power to the mission devices also occupies physical space to embed the mission devices. Therefore, this research surveys earlier studies on nano-satellite harness design. In addition, we consider the possibility of an optical wireless harness for the satellite bus system to achieve a more effective and reliable design for the CubeSats.


Onboard computer
Optical wireless module
Optical wireless module
Temp. sensor
Optical wireless module
Attitude control System
optical wireless module
Gyro sensor
Optical wireless harness
Harness

©ICAROB 2023 ALife Robotics Corp. Ltd.

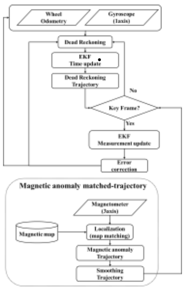
Poster

Chair Yong Hun Kim (Sejong University, Korea)

POS Magnetic Anomaly-Matched Trajectory and Dead Reckoning Fusion Mobile Robot Navigation

Yong Hun Kim, Bo Sung Ko, Jin Woo Song (Sejong University, Korea)

This paper presents a navigation algorithm based on the magnetic anomaly-matched trajectory for an indoor mobile robot. An environment with many similar magnetic field distortions cannot navigate stably with a magnetic anomaly-based navigation algorithm. In this study, we propose that stable navigation solution in various indoor environments through a fuse between magnetic anomaly-matched trajectory and mobile robot inertial trajectory. The proposed method uses dead reckoning(DR) as the primary navigation system and determines the similarity between magnetic anomaly-matched trajectory and DR trajectory to generate a trajectory key-frame. When the key-frame is detected, perform an extended Kalman filter(EKF) measurement update. In this study, an open dataset is used to verify the algorithm's performance, and we compared the existing algorithm. Since the method proposed uses only odometer, gyroscope, and magnetometer for indoor navigation, it has the advantage of building a system at a low cost.


Wheel Odometry
Gyroscope (1axis)
Dead Reckoning
EKF Time update
Dead Reckoning Trajectory
No
Key Frame?
Yes
EKF Measurement update
Error correction
Magnetic anomaly matched-trajectory
Magnetometer (3axis)
Magnetic map
Localization (map matching)
Magnetic anomaly Trajectory
Smoothing Trajectory

©ICAROB 2023 ALife Robotics Corp. Ltd.

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)

[A]

Ab. Aziz	Kamarulzaman	GS3-4	37/105
Ab. Aziz	Nor Azlina	GS3-4	37/105
Abbas	Samer	OS25-6	30/89
Abbyasov	Bulat	OS6-7	36/58
Abdul Aziz	Siti Zaleha	OS24-1	39/85
		OS24-2	39/86
Abdullah	Nor Fazilah	OS27-1	31/92
		OS27-5	31/93
Abe	Akinori	OS28-4	42/95
Abu Talip	Mohamad Sofian	OS26-1	34/90
Aburada	Kentaro	OS14-3	24/70
		OS14-4	24/70
Afanasyev	Ilya	OS6-6	36/58
Ahn	Jonghyun	OS20-3	25/78
Ai	Lian Lim	OS1-4	40/49
Ai Leen	Gan	OS29-3	27/96
Al-Badarneh	Aaa	OS1-7	40/50
Al-Talib	Ammar A.	OS27-1	31/92
		OS27-2	31/92
		OS27-3	31/92
		OS27-4	31/93
		OS27-5	31/93
		OS27-6	31/93
Al-quraan	Obaid	OS1-6	40/49
Alahmad	Raji	OS21-1	37/80
Alquran	Hiam	OS1-6	40/49
		OS1-7	40/50
Alraie	Hussam	OS21-7	38/82
AlRowaily	Mofleh Hannuf	OS29-5	27/97
Alsalatie	Mohammed	OS1-6	40/49
		OS1-7	40/50
Ang	Chun Kit	OS25-2	30/87
		OS25-3	30/88
		OS25-4	30/88
Ang	Koon Hian	OS25-4	30/88
Ang	Koon Meng	OS25-3	30/88
		OS25-4	30/88
Anggia	Pinkie	GS5-3	44/108
Anuar	Mohamad Izzuddin	OS26-1	34/90
Aoki	Kenji	OS18-3	40/76
Aono	Shota	GS5-1	44/107
Apurin	Artem	OS6-5	36/58
		OS6-7	36/58
Arima	Masakazu	OS20-1	25/78
Arita	Daisaku	OS20-5	25/79
Arof	Hamzah	OS26-1	34/90
		OS29-5	27/97
Asri	Muhammad Amirul Aiman	OS1-4	40/49
Atifah	Sarah	OS27-6	31/93
Atiqa	Ain	OS27-1	31/92
		OS27-4	31/93
Aung	IK Chu	OS27-2	31/92
Azizan	Muhammad Azizi	OS24-1	39/85
		OS24-2	39/86
		OS24-3	39/86

[B]

Bai	Yang	OS6-7	36/58
Biki	Bidesh Biswas	OS18-5	41/76
Blattler	Aran	GS4-2	43/106

©ICAROB 2023 ALife Robotics Corp. Ltd.

[C]
Cai Hanlin OS25-7 30/89
Cao Jing GS3-1 37/104
Chang Meng-Wei OS2-1 21/50
OS2-6 21/51
Cheah Dei Xuan OS26-3 34/90
Chebotareva Elvira OS6-1 35/56
OS6-2 35/57
OS6-4 35/57
OS6-5 36/58
Chen Chun -Nan OS5-3 22/55
Chen Hsiao-Fang OS5-5 22/56
Chen Jin-Yuan OS5-5 22/56
Chen Lu OS11-3 29/65
Chen Shilin OS25-8 31/89
Chen Shuai OS15-2 36/71
Chen Ting-Yi OS4-4 21/54
Chen Xiaoyan PS3 27/46
OS31-1 41/98
OS31-2 41/99
OS31-3 41/99
OS31-4 41/99
OS31-5 41/99
Chen Yu OS11-2 29/65
Chen Zhihui OS31-3 41/99
Cheng Wy-Liang OS25-1 30/87
OS25-2 30/87
Chew Cheng Jin GS6-3 44/109
Chit Su Mon GS2-2 43/103
Chou Wen-Huei OS4-1 20/53
Chow Li Sze OS25-6 30/89
Chu Qi OS11-4 29/65
Chumkamon Sakmongkon OS13-2 32/67
OS13-3 32/67
OS13-4 33/67
OS13-5 33/68
OS13-6 33/68
OS13-7 33/68
OS13-8 33/69
OS13-9 33/69

[D]
Dahari Mahidzal OS1-5 40/49
Dai Fengzhi OS8-2 32/61
OS9-1 28/63
OS11-2 29/65
Desa Hazry OS24-1 39/85
OS24-2 39/86
OS24-3 39/86
di Maria Enrico OS21-2 37/80
Dinara Gabdullina GS2-4 43/104
Dixit Ankur GS2-5 43/104
Dobrokvashina Alexandra OS6-3 35/57
OS6-7 36/58
Duan Suqing OS15-2 36/71
OS15-5 36/72

[E]
Edgar A. Martínez-García OS6-4 35/57
Egawa Haruka GS3-3 37/105
Enomoto Sho OS3-2 29/52
Evgeniy Zykov GS2-4 43/104

[F]
Fatahuddin Malla Abdul Jalil GS4-4 43/107
Fei Shumin OS8-2 32/61
Fujinaga Takuya OS20-3 25/78
OS20-5 25/79
Fukuda Riku OS20-2 25/78
Fukuda Yukiya OS17-4 24/74
Fukumoto Ryuma OS20-4 25/79
Fukumoto Yukihiro OS21-1 37/80
Fxyma Hiroki OS28-1 42/94

©ICAROB 2023 ALife Robotics Corp. Ltd.

[G]

Gamberov	Timur	OS6-3	35/57
Gamolped	Prem	OS13-2	32/67
		OS13-3	32/67
Gan	Jhui Ken	OS26-4	34/91
Gao	Yiting	OS10-4	28/64
Gautam	Ashish	OS12-1	22/66
Geiser	Sylvain	OS13-6	33/68
		OS13-8	33/69
Gong	Haoran	OS9-3	28/63
		OS11-4	29/65
Goud	Kaki Harshayrdhan	OS30-1	38/97
Grayson	Chua Ray Sern	OS27-4	31/93
Guo	Haoyu	OS11-2	29/65
		OS11-3	29/65
Guo	Jr-Hung	OS5-4	22/55

[H]

Hamada	Daiki	OS20-2	25/78
Hamakawa	Fumito	OS18-5	41/76
Hamamoto	Hiroki	OS7-3	23/59
Hamdan	Nur	GS6-1	44/108
		GS6-2	44/109
		GS6-3	44/109
Hara	Akiyoshi	OS16-1	35/72
Harada	Kensuke	OS16-1	35/72
		OS16-2	35/73
Hasebe	Kodai	OS18-5	41/76
Hayashi	Eiji	OS13-1	32/66
		OS13-2	32/67
		OS13-3	32/67
		OS13-4	33/67
		OS13-5	33/68
		OS13-6	33/68
		OS13-7	33/68
		OS13-8	33/69
		OS13-9	33/69
Hayashi	Toshiki	GS5-1	44/107
Hayashi	Yuki	OS28-3	42/94
Hayashida	Yuki	OS22-5	26/84
Higashi	Kazuki	OS16-1	35/72
Hirai	Hiroyasu	OS20-6	25/79
Hirakawa	Tomohiro	OS7-3	23/59
Hiramitsu	Tatsuhiro	OS16-3	35/73
Hiraoka	Toru	OS3-1	29/52
		OS3-2	29/52
		GS6-5	44/110
Hirata	Kensuke	OS20-4	25/79
Hirota	Masaharu	OS3-4	29/53
Hong	Yuhang	OS25-8	31/89
Hsia	Kuo-Hsien	OS5-4	22/55
		OS6-3	35/57
Hsiao	Jia-Ming	OS4-3	20/54
Hsu	Yu-Chi	OS5-1	21/54
Hu	Jiahao	OS29-1	27/95
Hu	Jianhao	OS15-1	36/71
Hu	Jiaqi	OS25-7	30/89
Huang	Meng-Fan	OS32-3	42/100
Huang	Min-Wei	OS2-4	21/51
Huang	Yu-Sheng	OS5-3	22/55
Hung	Chung-Wen	OS4-1	20/53
Husnain	Anees ul	OS1-4	40/49
		OS1-5	40/49

[I]

Iamcharoen	Phattharaporn	GS1-1	22/102
Ibrahim	Imanurfatiehah	OS26-1	34/90
		OS29-5	27/97
Ichikawa	Naoki	OS16-3	35/73
Ikeda	Mahiro	OS22-1	26/82
Ikeda	Satoshi	OS18-1	40/75
		OS18-2	40/75
		OS18-3	40/76

©ICAROB 2023 ALife Robotics Corp. Ltd.

		OS18-4	41/76
		OS18-5	41/76
Ikeda	Takuya	GS2-3	43/103
Imai	Shinichi	OS7-4	23/60
Imai	Yasumoto	GS5-1	44/107
Imam	Jakaria Mahdi	OS29-4	27/96
Inniyaka	Irmiya R.	OS20-2	25/78
		OS21-3	37/81
Iqbal	Javid	OS26-6	34/91
Ishak	Nurfadzillah	OS24-1	39/85
		OS24-2	39/86
		OS24-3	39/86
Ishigami	Ryoichi	OS14-2	23/70
Ishii	Kazuo	OS20-1	25/78
		OS20-2	25/78
		OS20-4	25/79
		OS20-5	25/79
		OS20-6	25/79
		OS20-7	26/80
		OS21-1	37/80
		OS21-2	37/80
		OS21-3	37/81
		OS21-4	37/81
		OS21-5	37/81
		OS21-6	37/82
		OS21-7	38/82
Ishikawa	Seiji	GS3-1	37/104
		GS3-3	37/105
Islam	Aminul	OS29-4	27/96
Islam	Shayla	OS26-3	34/90
		OS26-5	34/91
		OS29-2	27/96
Ito	Kazuaki	GS4-1	43/106
		GS5-4	44/108
Ito	Takao	OS18-1	40/75
		OS18-2	40/75
		OS18-3	40/76
		OS18-4	41/76

Ito	Tsutomu	OS18-1	40/75
		OS18-3	40/76
		OS18-4	41/76
Iwahashi	Masahiro	OS1-4	40/49
		OS1-5	40/49
Iwanaga	Takuro	OS18-2	40/75
Iwao	Jun	GS4-3	43/106

[J]

Jamal Deen	Ashraf Ali	OS1-1	39/48
Ji	Cuiying	OS15-3	36/71
Jia	Hongyan	OS9-4	28/63
Jia	Yingmin	OS8-4	32/62
		OS8-5	32/62
Jitviriya	Wisanu	GS1-1	22/102
Jothi	Neesha	OS1-1	39/48
		OS1-2	39/48
		OS1-3	39/48
		OS26-5	34/91
		OS26-6	34/91
		OS29-1	27/95
Jusman	Yessi	OS25-1	30/87
		OS25-3	30/88

[K]

Kamiya	Tohru	GS6-4	44/109
Kamon	Supaphon	OS21-2	37/80
Kanaoka	Daiju	OS17-3	24/74
Kang	Mingxia	OS10-2	28/64
Kang	Ruming	OS10-2	28/64
Kasai	Shintaro	GS7-1	33/110
Kasman		OS13-1	32/66
Katayama	Tetsuro	OS14-1	23/69
		OS14-3	24/70
		OS14-4	24/70
Kato	Norihiko	GS5-2	44/107
Kawada	Kazuo	OS23-1	34/84
		OS23-3	34/85

©ICAROB 2023 ALife Robotics Corp. Ltd.

		OS23-4	35/85
Kawaguchi	Natsuki	OS7-5	23/60
		OS7-6	23/60
		OS7-7	23/60
Kawajiri	Kazushi	OS20-5	25/79
Kawasaki	Yuki	OS22-4	26/83
Keoy	Kay Hooi	OS1-2	39/48
Khaled	Belal	OS30-3	38/98
Khan	M.K.A Ahamed	OS30-2	38/98
Kidiraliev	Elaman	OS6-8	36/58
Kim	Yong Hun	POS	26/112
Kimura	Kenji	OS20-6	25/79
		OS20-7	26/80
Kinioshita	Takuya	OS7-2	23/59
		OS7-3	23/59
Kita	Yoshihiro	OS3-3	29/52
		OS14-3	24/70
		OS14-4	24/70
Kitamura	Kentarou	GS1-3	22/102
Ko	Bo Sung	POS	26/112
Ko	Cheng-Lung	OS4-1	20/53
Ko	Chia-Nan	OS4-4	21/54
Ko	Ting-Ru	OS4-4	21/54
Kobayashi	Jun	GS4-4	43/107
Kobayashi	Tomoya	OS22-3	26/83
Kohno	Takashi	OS12-1	22/66
		OS12-2	22/66
Kolandaisamy	Raenu	OS1-1	39/48
		OS1-3	39/48
Koomklang	Nattapat	OS13-2	32/67
Kuang	Wentao	OS11-3	29/65
Kudo	Akihiro	OS18-6	41/76
Kumano	Minoru	GS6-5	44/110
Kumar	Mandigiri Golla Jagan	OS30-1	38/97
Kurogi	Shuichi	OS18-4	41/76
		OS18-5	41/76
Kuwata	Kaito	OS20-2	25/78
[L]			
Lai	Hsin-Yu	OS2-4	21/51
Lavrenov	Roman	OS6-6	36/58
		OS6-8	36/58
Lee	Chen-Yu	OS2-1	21/50
Lee	Meng-Huan	OS2-4	21/51
		OS2-5	21/51
Lee	Qi Jian	OS30-2	38/98
Li	Chu-Fen	OS2-3	21/50
Li	Fangyan	OS10-1	28/64
Li	Hongbing	OS6-4	35/57
Li	Jiaxin	OS11-1	29/65
Li	Jung-Shian	OS2-1	21/50
		OS2-3	21/50
		OS2-4	21/51
		OS2-5	21/51
		OS2-6	21/51
Li	Tengxiang	OS25-8	31/89
Li	Wei	OS9-4	28/63
Li	Yi-Yu	OS4-4	21/54
Li	Zheng	OS25-7	30/89
Li	Zhihan	GS6-1	44/108
Li	Zhiyang	OS11-2	29/65
		OS11-3	29/65
Li	Zongyi	OS11-4	29/65
Liang	Xiwen	PS3	27/46
Lim	Wei Hong	OS25-1	30/87
		OS25-2	30/87
		OS25-3	30/88
		OS25-4	30/88
		OS25-5	30/88
		OS25-6	30/89
		OS25-7	30/89
		OS25-8	31/89
		OS30-3	38/98
Lin	Chien-Tung	OS32-2	42/100

Lin	Mingrui	OS25-8	31/89
Lin	Shih-Chin	OS2-2	21/50
Liu	Hongze	OS15-4	36/72
Liu	I-Hsien	OS2-1	21/50
		OS2-3	21/50
		OS2-4	21/51
		OS2-5	21/51
		OS2-6	21/51
Liu	Jingwen	OS9-4	28/63
Liu	Mengyu	OS10-3	28/64
Liu	Qunpo	OS8-3	32/62
Liu	Siyuan	OS9-3	28/63
		OS11-4	29/65
Liu	Yong	OS11-1	29/65
Liu	Ziyue	GS3-1	37/104

[M]

M.Tahir	Noor Idayu	OS27-2	31/92
		OS27-6	31/93
Magid	Evgeni	PS2	20/46
		OS6-3	35/57
		OS6-4	35/57
		OS6-6	36/58
		OS6-7	36/58
Maike	He	OS16-3	35/73
Mansor	Fakhrul Mukmin	OS24-1	39/85
		OS24-2	39/86
Matsumoto	Ikuya	OS17-3	24/74
Matsumoto	Sho	OS14-1	23/69
Matsumoto	Takuya	OS13-7	33/68
Matsuo	Takayuki	OS20-1	25/78
		OS20-5	25/79
Matsuno	Seigo	OS18-1	40/75
Miao	Xia	OS31-2	41/99
Miura	Seiichiro	OS20-1	25/78
Miyashita	Takeaki	OS14-3	24/70
Mizoguchi	Nanako	OS18-7	41/77
Mizoguchi	Yusuke	OS20-2	25/78
Mochizuki	Ryuugo	OS21-5	37/81
Mohamed	Rizon	OS25-1	30/87
		OS25-2	30/87
		OS25-4	30/88
Mohammed Al-Qussari	Azam	OS1-3	39/48
Mohamed Shah	Noraisyah	OS1-5	40/49
Mohammad Shahrazel	Aisyah Arina	OS24-1	39/85
		OS24-2	39/86
Mohammadpour	Atefeh	OS30-1	38/97
Mohd Yamin	Muhammad Nadzree	GS3-4	37/105
Mokayef	Mastaneh	OS25-5	30/88
		OS30-1	38/97
		OS30-2	38/98
		OS30-3	38/98
Mokayef	Miad	OS30-2	38/98
		OS30-3	38/98
Mokhtar	Norrima	OS1-2	39/48
		OS1-3	39/48
		OS1-4	40/49
		OS1-5	40/49
		OS29-4	27/96
		OS29-5	27/97
Morita	Kairi	GS3-1	37/104
Moritsuka	Shunsuke	GS6-4	44/109
Motoda	Tomohiro	OS16-1	35/72
Mowshowitz	Abbe	OS13-7	33/68
Muhamad Hatta	S. F. W.	OS29-4	27/96
Mukunoki	Masayuki	GS2-3	43/103
Murai	Keita	OS23-1	34/84
Mustafa	Wan Azani	OS1-6	40/49
		OS1-7	40/50
Mustafin	Maksim	OS6-1	35/56
		OS6-4	35/57

©ICAROB 2023 ALife Robotics Corp. Ltd.

[N]

Nagamatsu	Masayasu	OS23-3	34/85
		OS23-4	35/85
Nagao	Kenta	OS7-7	23/60
Nagatomo	Yoki	GS2-3	43/103
Nagayoshi	Masato	GS1-2	22/102
Naim	Nani Fadzlina	OS29-6	27/97
Nakagawa	Masato	PS4	38/47
Nakahara	Takumi	OS18-3	40/76
Nakamura	Akira	OS16-2	35/73
Nakamura	Jun	OS28-2	42/94
Nakamura	Shingo	OS3-3	29/52
Nakaya	Jun	GS7-3	33/111
Nan	Pepin	GS6-2	1944/2
Nanami	Takuya	OS12-2	22/66
Naeem	Bushra	OS30-2	38/98
Ng	Wai Lam	OS26-5	34/91
Nishida	Yuya	OS20-2	25/78
		OS20-5	25/79
		OS21-1	37/80
		OS21-3	37/81
		OS21-5	37/81
		OS21-6	37/82
Nishimuta	Keisuke	OS20-2	25/78
Nitta	Masuhiro	GS3-1	37/104
Noma	Shogo	OS3-1	29/52

[O]

Ochiiwa	Takashi	OS7-3	23/59
Oda	Shunsuke	OS20-4	25/79
Ogata	Kouki	OS20-6	25/79
Ogata	Takashi	OS28-5	42/95
Ogawa	Shintaro	OS13-4	33/67
		OS13-8	33/69
Ohnaka	Kento	OS18-7	41/77
Ohnishi	Yoshihiro	OS23-2	34/84
Ohshima	Kenich	OS20-5	25/79
Ohtake	Hiroshi	GS4-3	43/106
Ohtani	Masahiro	OS22-4	26/83
Okada	Masayuki	OS20-5	25/79
Okazaki	Naonobu	OS14-3	24/70
		OS14-4	24/70
		GS2-3	43/103
Okuno	Hirotsugu	OS22-1	26/82
		OS22-2	26/83
		OS22-3	26/83
Ono	Jumpei	OS28-5	42/95
Ono	Taisei	GS3-3	37/105
Ono	Tomohiro	OS17-2	24/74
		OS17-4	24/74
		OS17-6	24/75
Ooya	Akihiro	OS13-3	32/67
Oshige	Takeru	OS20-3	25/78
Oshiumi	Wataru	GS2-5	43/104
Ozawa	Masayoshi	OS20-1	25/78

[P]

Pan	Li	OS25-1	30/87
		OS25-2	30/87
Pan	Wenxuan	OS11-2	29/65
Peng	Yizhun	OS15-1	36/71
		OS15-2	36/71
		OS15-3	36/71
		OS15-4	36/72
		OS15-5	36/72
Phunopas	Amornphun	GS4-2	43/106
Poojitha	Vytla	OS30-1	38/97
Promma	Sornsiri	OS13-5	33/68
Purev	Amarbold	GS7-2	33/110

[Q]

Qiao	Zhi	OS10-3	28/64
Qu	Feiyang	OS11-4	29/65

©ICAROB 2023 ALife Robotics Corp. Ltd.

[R]

Rahman	Hameedur	OS25-2	30/87
Rajagopal	Heshalini	OS1-1	39/48
		OS1-2	39/48
		OS1-3	39/48
		OS1-4	40/49
		OS1-5	40/49
		OS26-2	34/90
		OS26-3	34/90
		OS26-4	34/91
Rajapakse	R.P.C. Janaka	OS32-1	42/100
		OS32-2	42/100
		OS32-3	42/100
		OS32-4	42/101
Rajendran	Salini	GS6-3	44/109
Ramlee	Muhammad Syazni Ikmal	OS1-5	40/49
Reddy	Myadadha Goutham	OS30-1	38/97
Ren	Keying	OS31-1	41/98

[S]

Safin	Ramil	OS6-6	36/58
Saita	Satoko	GS1-3	22/102
Saitoh	Takeshi	GS4-4	43/107
Sakamoto	Makoto	OS18-1	40/75
		OS18-2	40/75
		OS18-3	40/76
		OS18-4	41/76
		OS18-5	41/76
		OS18-6	41/76
		OS18-7	41/77
Sanada	Atsushi	OS20-5	25/79
Sanbuichi	Jun	GS6-5	44/110
Saputra	Mochamad Chandra	OS14-4	24/70
Sarakon	Pornthep	GS2-1	43/103
		GS4-2	43/106
Sarawanan	Thivagar Chettiar	OS1-1	39/48
Saruchi	Sarah Atifah	OS27-3	31/92
Sato	Junya	GS4-1	43/106
		GS5-1	44/107
		GS5-4	44/108
Sato	Masanori	OS20-1	25/78
Sato	Ryoka	OS22-2	26/83
Sato	Takao	OS7-5	23/60
		OS7-6	23/60
		OS7-7	23/60
Sawada	Takuro	OS16-3	35/73
Seki	Hiroaki	OS16-3	35/73
Sethu	Devika	OS1-1	39/48
Shaik	Shabana Anjum	OS26-4	34/91
		OS29-2	27/96
		OS29-3	27/96
Sharma	Abhishek	OS25-1	30/87
		OS25-2	30/87
		OS25-3	30/88
		OS25-4	30/88
Sheng	Chang	OS10-1	28/64
Sheng	Yueqin	OS8-3	32/62
Shi	Yibo	OS8-1	32/61
Shiba	Tomoya	OS17-2	24/74
Shibghatullah	Abdul	GS6-1	44/108
		GS6-2	44/109
		GS6-3	44/109
Shigyo	Yuki	OS20-7	26/80
Shirahashi	Kanako	OS20-5	25/79
Siah	Cheong Lin	OS26-2	34/90
Silangren	Hattapat	GS1-1	22/102
Song	Jin Woo	POS	26/112
Song	Yunzhong	OS8-2	32/61
Sonoda	Takashi	OS20-5	25/79
Subaramaniam	Kasthuri	GS2-2	43/103
		GS6-2	44/109

©ICAROB 2023 ALife Robotics Corp. Ltd.

		GS6-3	44/109
Sudheer	Sayanth	OS1-2	39/48
		OS1-3	39/48
Sugino	Akihiro	OS20-2	25/78
Sugimoto	Taiki	OS28-2	42/94
Sulaiman	Shifa	OS6-2	35/57
		OS6-3	35/57
Sultanov	Ramir	OS6-2	35/57
Sumi	Kaoru	PS1	20/46
		GS5-3	44/108
Summakieh	MHD Amen	OS30-1	38/97
		OS30-2	38/98
		OS30-3	38/98
Sun	Qian	OS8-5	32/62
Sun	Yizhe	OS15-1	36/71
Sung	Jun Kyu	GS2-2	43/103
Susawa	Yuta	OS23-1	34/84
Suzuki	Akihiro	OS17-6	24/75
Suzuki	Hiroyuki Y.	OS23-1	34/84
		OS23-3	34/85
		OS23-4	35/85
Suzuki	Katsuaki	OS21-6	37/82
Suzuki	Yasuhiro	OS19-1	31/77
		OS19-2	31/77
Svinin	Mikhail	OS6-7	36/58

[T]

T. Hussain	Abadal-Salam	OS24-1	39/85
		OS24-2	39/86
Tangsuksant	Watcharin	GS1-1	22/102
		GS2-1	43/103
		GS4-2	43/106
Tahara	Shu	OS20-4	25/79
Takagi	Tomohiko	OS14-1	23/69
		OS14-2	23/70
Takatani	Atsushi	OS7-3	23/59
Takatsuka	Kayoko	GS2-3	43/103
Takei	Amane	OS18-3	40/76
		OS18-4	41/76
		OS18-5	41/76
		OS18-6	41/76
		OS18-7	41/77
Takemura	Kakeru	OS18-4	41/76
Takemura	Sota	OS7-2	23/59
Takemura	Yasunori	OS20-3	25/78
		OS20-5	25/79
Takeuchi	Rio	GS5-2	44/107
Takeuchi	Shoki	OS23-2	34/84
Tanaka	Rikuto	OS20-2	25/78
Tanaka	Yoshiki	OS20-2	25/78
Tamai	Teruyuki	OS23-2	34/84
Tamaki	Hisashi	GS1-2	22/102
Tamukoh	Hakaru	OS17-1	24/73
		OS17-2	24/74
		OS17-3	24/74
		OS17-4	24/74
		OS17-5	24/74
		OS17-6	24/75
Tan	Chi Jie	OS13-8	33/69
Tan	Gek Siang	GS3-4	37/105
Tan	Jia Wei	OS30-1	38/97
Tan	Joo Kooi	GS3-1	37/104
		GS3-2	37/105
		GS3-3	37/105
Tan	Yew Tian	OS24-2	39/86
Tanaka	Takaya	OS7-5	23/60
Tanaka	Yoshiki	OS20-2	25/78
Teh	Xi Hang	OS24-1	39/85
Teramoto	Mariko	GS1-3	22/102
Thong	Chee Ling	OS26-2	34/90
		OS29-1	27/95
		OS29-3	27/96

©ICAROB 2023 ALife Robotics Corp. Ltd.

Tiang Sew Sun OS25-1 30/87
OS25-2 30/87
OS25-3 30/88
OS25-4 30/88
OS25-5 30/88
OS25-6 30/89
OS25-7 30/89
OS30-3 38/98
Ting Chung-Te OS5-2 21/55
Tokumitsu Masahiro GS7-3 33/111
Tokuyama Toshimasa OS32-1 42/100
Tominaga Ayumu OS13-4 33/67
OS13-6 33/68
OS13-7 33/68
OS13-8 33/69
Tominaga Moeko OS20-3 25/78
OS20-5 25/79
Tomiyama Hideki OS7-3 23/59
Tomokawa Takumi OS13-6 33/68
OS13-8 33/69
Tsai Pang-Wei OS5-1 21/54
Tsoy Tatyana OS6-2 35/57
OS6-6 36/58
Tsuji Masatoshi GS7-3 33/111
Tsuji Tokuo OS16-3 35/73
Tsuji Tomofumi OS13-2 32/67
OS13-3 32/67
OS13-7 33/68

[U]
Uchida Noriyuki GS2-3 43/103
Ueda Takumi OS20-6 25/79
OS21-4 37/81
Ueno Takaaki OS13-9 33/69
Usuda Akio OS14-2 23/70

[V]
Vlada Kugurakova GS2-4 43/104

[W]
W. Muhamad Wan Norsyafizan OS29-6 27/97
Wadekar Suhaim Parvez OS25-3 30/88
Wagatsuma Hiroaki GS2-5 43/104
GS7-1 33/110
GS7-2 33/110
Wakitani Shin OS7-1 23/59
Wan Khairunizam OS1-6 40/49
OS1-7 40/50
Wan Wai Kit OS27-5 31/93
Wan Mohd Mahiyiddin Wan Amirul OS1-4 40/49
OS29-5 27/97
Wang Chaoli OS8-1 32/61
Wang Chun-Chieh OS4-2 20/53
Wang Hanyuan OS10-4 28/64
Wang Hucheng OS9-1 28/63
OS9-2 28/63
Wang Ji OS30-1 38/97
Wang Lu OS11-4 29/65
Wang Qian OS31-5 41/99
Wang Weifang OS10-2 28/64
Wang Weiqin OS25-8 31/89
Watanabe Yusuke OS17-5 24/74
Wei Hong-Yu OS5-1 21/54
Wen Haokang OS10-1 28/64
Wong Chin Hong OS25-1 30/87
OS25-3 30/88
OS25-4 30/88
OS25-5 30/88
OS25-6 30/89
OS25-7 30/89
OS25-8 31/89
OS30-3 38/98
Wong Wai Kiat OS25-5 30/88

©ICAROB 2023 ALife Robotics Corp. Ltd.

Wu	Chien-Hsin	OS2-3	21/50
Wu	Jiangyu	OS15-2	36/71
[X]			
Xian	Yap Chee	OS27-1	31/92
Xiang	Yande	OS10-3	28/64
		OS11-2	29/65
		OS11-3	29/65
Xiao	Huimin	OS8-2	32/61
Xu	Xin	OS25-8	31/89
[Y]			
Yahya	Siti Syuhaidah	OS24-1	39/85
		OS24-2	39/86
Yamaba	Hisaaki	OS14-3	24/70
		OS14-4	24/70
Yamada	Takayoshi	GS4-1	43/106
		GS5-1	44/107
		GS5-4	44/108
Yamamoto	Hidehiko	GS5-4	44/108
Yamamoto	Toru	OS7-2	23/59
		OS7-3	23/59
Yang	Laijun	GS5-2	44/107
Yang	They Kai	OS27-6	31/93
Yang	Wen-Fu	OS5-2	21/55
Yano	Yuga	OS17-4	24/74
Yao	Tomonori	OS7-6	23/60
Yasukawa	Shinsuke	OS20-4	25/79
		OS20-5	25/79
		OS22-4	26/83
		OS22-5	26/84
Yazid	Haniza	OS29-5	27/97
Yee	Chee Hong	OS26-6	34/91
Yew	Soo Yang	OS29-2	27/96
Yokomichi	Masahiro	OS18-4	41/76
Yokota	Hiroyuki	OS22-5	26/84
Yokote	Takumi	GS3-2	37/105
Yoshida	Keisuke	OS20-4	25/79
Yoshida	Syota	OS7-1	23/59
Yoshimoto	Yuma	OS17-1	24/73
Yuan	Tzu-Hsien	OS32-3	42/100
Yutao	Jin	OS31-4	41/99
[Z]			
Zagirov	Aidar	OS6-5	36/58
Zawawi	Hazwani	OS29-6	27/97
Zhang	Weicun	OS8-2	32/61
		OS8-5	32/62
Zhang	Yang	OS8-4	32/62
Zhang	Zongxuan	OS15-1	36/71
Zhao	Jichao	OS8-2	32/61
Zhao	Min	OS9-1	28/63
		OS9-2	28/63
		OS10-1	28/64
		OS11-1	29/65
Zhao	Yi	OS8-3	32/62
Zhao	Yuxi	OS8-3	32/62
Zhao	Zhihao	OS9-3	28/63
Zhong	Jianlin	OS32-4	42/101
Zhu	Tao	OS10-3	28/64
Zulkepli	Nik Noriman	OS24-1	39/85
		OS24-2	39/86
Zyout	Ala'a	OS1-7	40/50

Human-to-Human Interaction Using Virtual Agent Posing as Another Person

Kaoru Sumi
Future University Hakodate
Hakodate, Hokkaido 041-8655, Japan
Email: kaoru.sumi@acm.org
www.fun.ac.jp

Abstract

Opportunities for online distance education have greatly expanded in recent years, and agent-based interactions in virtual spaces have attracted attention in this context. In this talk, I will discuss the various educational possibilities of using virtual spaces and agents, by presenting examples from several studies. I will also introduce our own development of systems using game-based learning and a game-based story generation system that automatically generates scripts in real time on the basis of players' emotions and actions. Finally, I will discuss persuasive technology that systems can use to influence human behavior, along with impressions and applications of facial expressions and gestures, which are expressions of agents, and I will discuss what can happen when a virtual agent interacts with other users in a virtual space.

Keywords: virtual agents, virtual spaces, emotions, affective computing, persuasive technology, online education

1. Introduction

In recent years, remote teaching has rapidly become common in university education. However, it mainly entails watching a video of the speaker's face and viewing teaching materials as videos and slides.

In the field of artificial intelligence in education, a hot topic is learning systems that interactively handle learners by estimating their emotions. Our group's laboratory conducted research on a system that estimates a learner's emotions from his or her facial expressions, gestures, and operation history to assist in learning programming [1].

Another hot topic is metaverse systems, in which we become virtual agents that differ from ourselves and communicate in a virtual space. Regarding communication between virtual agents, in the past couple of years, there have been attempts for users to participate in conferences and open campuses as virtual agents from remote locations. These users choose prepared virtual agents to participate in events as different characters from themselves and communicate with other virtual agents.

This scenario poses several questions. For learning in a virtual space, what kind of learning effect is involved when an instructor and learner become virtual agents that differ from themselves? Does the kind of agent that will be more effective for learning depend on the learner's personality? Would estimation of the learner's current, real-world state be useful in a virtual world?

In this talk, we will examine various educational possibilities using virtual spaces and agents, while introducing examples from our group's research.

I will also introduce our own development of systems using game-based learning [2], and a game-based story generation system [3] [4] that automatically generates scripts in real time on the basis of players' emotions and actions.

In recent years, we have come to be able to use virtual spaces for distance learning. In the future, learning by experience in a virtual space is also expected to become feasible. When using virtual agents, the impact on communication should be considered. I will discuss persuasive technology [5] that systems can use to influence human behavior, along with impressions and applications of facial expressions and gestures, which are expressions of virtual agents [6]; and I will discuss what can happen when a virtual agent interacts with other users in a virtual space.

We need to understand the novel experience of using a virtual space and a virtual agent than differs from oneself. To provide preliminary knowledge and an example for the lecture, this paper describes a recently presented experiential learning system in a virtual space, which was developed in our laboratory. Please refer to the related papers [3][4] for a detailed survey.

2. Learning in a virtual space

The proposed system incorporates a player's real-time emotions and actions to facilitate game-based story generation. By using a webcam to acquire facial images of the player, the system performs real-time emotion recognition. Specifically, the player uses an Oculus Touch for motion recognition. The Oculus Touch controller uses two sensors to track its position in a three-dimensional virtual space. While the sensors respond to the controller's movements, the virtual space shows a virtual hand that moves just like the player's real hand. The system then classifies the player's actions by tracking his or her hand movements.

The system's virtual space comprises an office and a break room, like a closed space found in everyday life. As the office and break room aren't separated by a wall, the player can move freely between them.

This system was inspired by a short film called "The Black Hole." In the film, a company employee in an office prints a mysterious sheet of paper containing a "black hole." The hole turns out to be a "magical tool" that enables a human hand to pass through objects and take objects on the other side. For example, after placing the black hole on a vending machine, the employee could pass his hand through the machine and take items inside. The cunning employee used the hole to take snacks from the vending machine without paying. He then used it to pass his hand through a door and steal something.

In our proposed system, the player uses the Oculus Touch controller to perform various actions: "take out," "throw," "put," "get," and "eat/drink." The specific action is determined from the controller's acceleration and what the player touches. For example, if the player touches the desk in the office, he will take a document, book, or snack, depending on his emotion, which could be "fear," "anger," or "joy." As a result of the player operating the tool while experiencing emotions, the system generates a story.

The system uses Affdex with the Facial Action Coding System (FACS) [7] to perform real-time emotion recognition. FACS encodes emotional facial expressions by combining action units (AUs), which are the smallest units of facial expressions that can be distinguished visually. In this way, the system can acquire seven emotions: joy, fear, disgust, sadness, anger, surprise, and contempt. Because a player wearing a special sensor device to acquire emotions might be tense and unable to exhibit accurate emotions, the player's emotions are obtained via a camera.

The system generates story scripts by combining four components: (1) the action performed with the Oculus Touch, (2) the player's recognized emotion, (3) the object acquired in the virtual space, and (4) information about the player's emotion based on knowledge of the object. The text is generated by incorporating the above components into a sentence structure. In Table 1, rows 8–13 give examples of generated terms. Note that the system can also generate text automatically.

For example, suppose that the system generates and displays this sentence: "You nervously take the cola out of the vending machine." Here, it generates "You nervously" because it detects "fear" as the player's emotion. Similarly, "vending machine" is the target object of the player's virtual motion, "cola" is obtained from knowledge of the vending machine's contents, and "take out" is obtained from the player's action with the Oculus Touch, with modification based on the player's emotion.

The system visualizes and displays the results obtained by applying the knowledge-based method with the acquired actions and emotions. Here, I explain the sentence generation process in terms of the sentences listed in Table 1. Rows 1-7 contain introductory sentences that are displayed at the bottom of the screen to instruct the player how to use the system. In rows 8, 10 and 12, the player "takes out" an object from the desk, and the object corresponds to the emotion determined from knowledge of the target object (i.e., the desk). In rows 9, 11 and 13, the player "puts" the object somewhere after obtaining it. The story ends when the player performs a specific action a certain number of times within 5 min; for example, it may end when he or she "eats" three times. Likewise, the story may end if the player obtains an object a certain number of times. Finally, rows 14-16 contain ending sentences. Once the system has chosen the story's ending, it darkens the screen and outputs the ending.

This system outputs the user's game experience as a text history, and we believe that this reflects the learning experience itself with real-time emotions such as joy and surprise. With such a learning experience, we believe that reviewing the history of the experience later will also support the player's learning.

Table 1. *An Example of a Generated Story*

1	You had been sleeping for some time because you were tired from work.
2	You woke up at your own workplace with no one around.
3	You noticed that the copier is moving, and you go to the front of copier.
4	A black circle was printed on a piece of paper that emerges from the copier.
5	You felt a sense of strangeness and touched the black circle.
6	Suddenly, your hand was sucked into the circle.
7	You were surprised and remove your hand.
8	You took the book out of the disk with surprise.
9	You put the book with surprise.
10	You took out documents from the desk with a lack of interest.
11	You put the documents with a lack of interest.
12	You took the book out of the desk with surprise.
13	You put the book with surprise.
14	You were holding a paper in your hand, but you did not feel like doing anything.
15	You decided to spend the day at office today.
16	In the morning, the paper was back to its original state, and I told my colleagues what I had experienced.

3. Conclusion

In the future, systems will emerge that allow users to learn through experience in a virtual space, as in the system described here. Such systems still involve many unknown methodologies, and this is the crux of our research problem. There are also many unknown factors in learning through interaction using virtual agents, which means that there are research issues to address. Through this talk, I have sought to discuss various possibilities for using virtual spaces and agents in learning, by introducing several research cases.

References

1. Thomas James Z. Tiam-Lee and Kaoru Sumi: Analysis and prediction of student emotions while doing programming exercises. International Conference on Intelligent Tutoring Systems, Lecture Notes in Computer Science, 2019. (Best Full Paper Award).
2. Kaoru Sumi and Kodai Kasai: A Serious Game for Learning Social Networking Literacy by Flaming Experiences, Lecture Notes in Computer Science, Volume 178, Social Informatics and Telecommunications Engineering series, Springer, 2016.
3. Kaoru Sumi & Shusuke Sato: Experiences of Game-Based Learning and Reviewing History of the Experience Using Player's Emotions, Frontiers in Artificial Intelligence, 2022.
4. Shusuke Sato & Kaoru Sumi: Story Generation System Using Player’s Emotions for Review in Gamed-Based Learning. In 28th International Conference on Computers in Education, 2020. (Best Technical Design Paper award).
5. B. J. Fogg: Persuasive Technology –Using Computers to Change What We Think and Do-, Elsevier, 2003.
6. K. Sumi and M. Nagata, “Evaluating a virtual agent as persuasive technology, psychology of persuasion, janos csap ́o and andor magyar eds,” 2010.
7. Ekman P, Friesen W: Facial Action Coding System: A Technique for the Measurement of Facial Movement. Palo Alto: Consulting Psychologists Press, 1978.

Authors Introduction

Prof. Kaoru Sumi

Kaoru Sumi is a professor in Future University Hakodate, Japan. She received her Ph.D. in engineering from the University of Tokyo. She is currently working on Human Agent Interaction, Persuasive technology, Artificial Intelligence in Education, Affective Computing, Digital storytelling. She previously worked at ATR MI&C Research Laboratories, Communications Research Laboratory, and Osaka University, where she researched human-computer interaction, knowledge engineering, and the application of artificial intelligence. After Prof. Sumi worked on media informatics and human-agent interaction at the National Institute of Information and Communications Technology (NICT), and Hitotsubashi University. She was a visiting professor in British Columbia, Canada.

Simulation Tools for Urban Search and Rescue Robotics

Evgeni Magid
Department of Intelligent Robotics, Kazan Federal University, Kazan, Russian Federation
HSE University, Moscow, Russian Federation
E-mail: magid@it.kfu.ru
kpfu.ru/robolab.html

Abstract

Real world experiments are critical for validating performance of new concepts and algorithms in robotics field. Yet, experiments tend to be too expensive in terms of time and resources of a research team. Moreover, it is not feasible to conduct thousands of complex experiments with a physical robot in a real environment. To check new ideas, preliminary evaluate new algorithms and interaction protocols, on first stages of a research project it is reasonable to start within a simulation. To produce relevant results, a simulator should provide adequate models of robots and environments with realistic physical properties. This paper presents an overview of our experience in using robot operating system (ROS) with Gazebo and Webots simulators for urban search and rescue robotics projects and considers constructing new models of mobile robots and complicated environments, algorithm validation and comparative analysis.

Keywords: Robotics, modelling, simulation, USAR, ROS, Gazebo, Webots

1. Introduction

Urban search and rescue (USAR) robotics was introduced at the end of the 20th century as a research a field that studies mechanics of rescue robots[1], their navigation[2], mapping[3], interaction of a human with a rescue robot[4], and other classic tasks of robotics being viewed through a prism of rescue related tasks[5]. USAR robotics is employed for searching victims in partially damaged or completely destroyed man-made structures, reconnaissance and mapping, debris penetration (e.g., with small sized ground robots) that could improve the two previous tasks' efficiency, debris removal, telepresence, hazmat contamination, and other tasks[6]. In USAR scenarios rescue teams deal with collapsed buildings and victims (which are often trapped under urban environment ruins) and wireless communication may be severely disturbed because of large volume of steel and concrete debris. A typical USAR environment contains debris that are formed by damaged construction materials, furniture, household and office items that complicate environment observation, localization and mapping[7][8]. While such conditions are natural for a real world rescue scene, it is complicated, time, resources and space consuming to create a similar to real USAR scene for experiments. Yet, in order to check new ideas, preliminary evaluate new algorithms and interaction protocols for USAR, on first stages of a research project real world conditions could be approximated virtually within a simulation.

Nowadays, simulations are extensively used in robotics, including robot design and construction[9], control algorithms validation[10], interaction protocols development[11], education[12] etc. They allow building a broad variety of virtual models of robots and environments[13] while targeting to achieve a relatively realistic behavior. We highlight the following reasons of using simulation in robotics[14]: real experiments with robots are expensive while simulations are cheap; simulation experiments take less time than real experiments; simulations are always safer than experiments in real world; simulations allow testing novel concepts and algorithms even if a required hardware is not available for a user; simulations help to quickly detect and correct conceptual errors in algorithms; finally, simulations could provide fairly reasonable and statistically valuable testing of experimental setups. Those reasons apply to all robot types and tasks. A simulator became a very progressive tool that could reproduce complicated test environments [15]; they can use user-defined physics with dynamic changes, which makes it possible to construct and use a robot model behaving acceptably similar to a real robot.

Among a broad variety of popular simulators, our research group concentrates on the Gazebo simulator that was designed specifically for simulating robots and their environment[16]. Being fully compatible with Robot Operating System (ROS)[17] Gazebo simulator allows creating and using robot models with ROS-based control systems without additional efforts. In most cases the Gazebo simulator is used together with the RViz simulator[18] as they provide a good complementary level of sensory input and motion activities visualization. Recently, our research team started using Webots[19], which enables a faster modelling and prototyping while suffering from a lower level of realism. Occasionally we employ Matlab simulator, which allows fast drafting of ideas but in terms of USAR robotics suffers from a very high level of abstraction[20]. There exist a number of other popular robotics simulators that could be used for modelling USAR scenarios including USARSim[21] and VRep (recently known as CoppeliaSim[22]).

In the next sections we briefly share our experience in employing Gazebo, RViz and Webots simulators for a robot and environment modelling, a comparative analysis of various approaches, a preliminary validation of algorithms and robots' interaction.

2. Robot modelling

We constructed a number of new models of mobile robots that correspond to real robots, which are available for our research team within the laboratory. For the Gazebo simulator car-like Avrora Unior wheeled robot[23], Servosila Engineer crawler robot[24], and PX4-LIRS UAV[25] models were developed.

First models of the Servosila Engineer robot[26] suffered from an incompleteness and a high level of abstraction. A second generation of the model approximated tracks with a varying number of pseudo-wheels[24] and significantly improved the performance, yet had issues with a real time factor (RTF) of the simulation and a number of locomotion issues that caused seizing at sharp edges of an environment. Next, a comparative analysis[27] of an optimal number of pseudo-wheels for each track simulation allowed to find a tradeoff between the model complexity and performance in terms of RTF. To increase RTF further, Servosila Engineer model was improved via CAD files' analysis and their systematic simplification[28]. In[29] we proposed a novel approach of modelling Servosila Engineer robot's crawlers as a set of gear wheels, which had a better RTF and solved the seizing problem. While starting with ROS Indigo and Gazebo 2.2.3 simulator, gradually the robot model has been evolved to ROS Noetic and Gazebo 11 simulator[30]. Figures 1-5 demonstrate the evolution of the Gazebo model: the real robot (Fig.1), the first model in Gazebo and RViz (Fig.2), the track approximation with arrays of the pseudo-wheels (Fig.3) and the gear-wheels (Fig.4), the latest version of the model in Gazebo 11 (Fig.5).

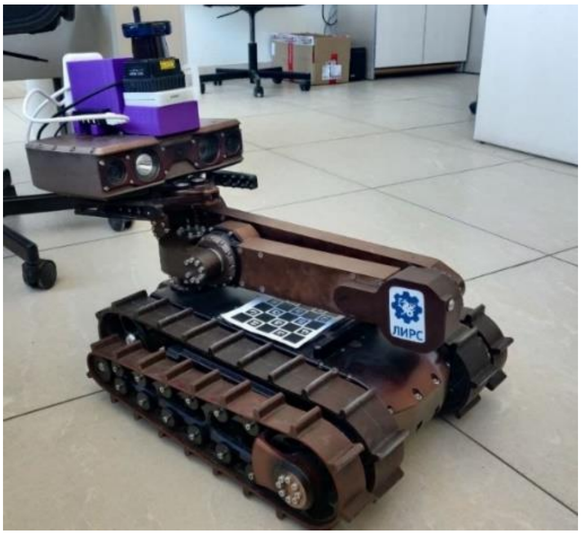


Fig.1. Servosila Engineer robot at Laboratory of Intelligent Robotic Systems (LIRS), Kazan Federal University

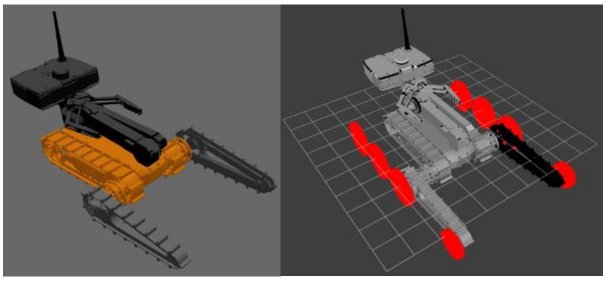


Fig.2. The first version of Servosila Engineer robot model in Gazebo 2.2.3 (left) and in RViz (right). Imaginary pseudo-wheels are shown in red[26].

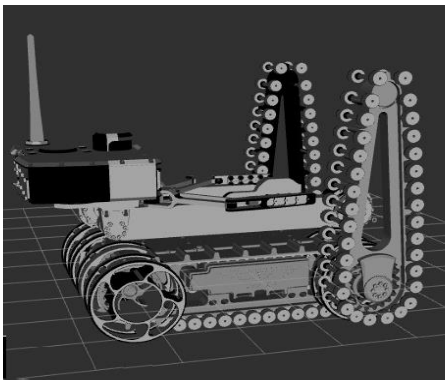


Fig.3. Track approximation with arrays of pseudo-wheels[27]

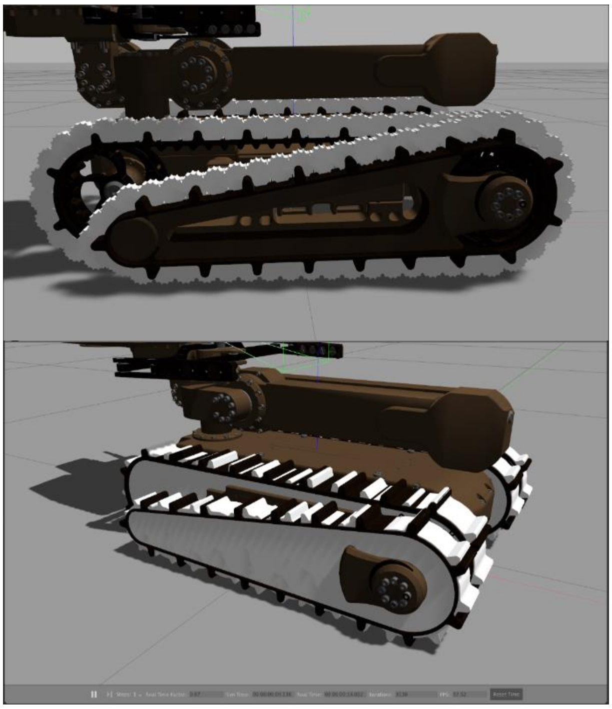


Fig.4. Track approximation with arrays of gear-wheels: small-sized (top) and large-sized wheels (bottom)[29]

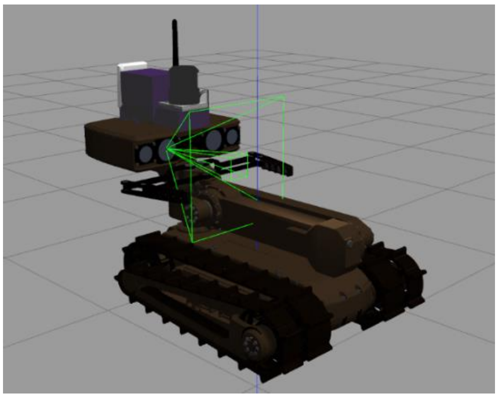


Fig.5. Gazebo 11 simulation model[30]

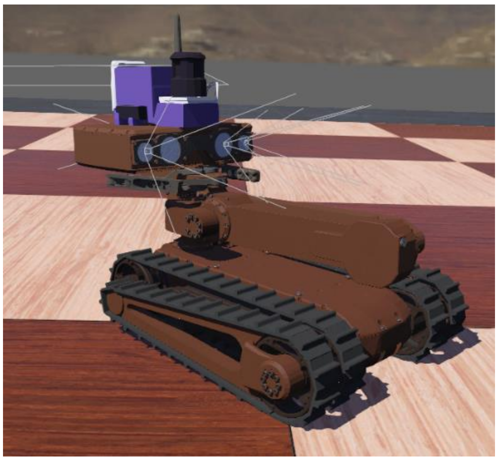


Fig.6. Webots simulation model[31]

Servosila Engineer robot[31] and its onboard sensory system[32] were modelled in Webots simulator in order to diverse virtual validation of algorithms that use the robot with a different simulator. Fig.6 demonstrates the Webots model of the robot.

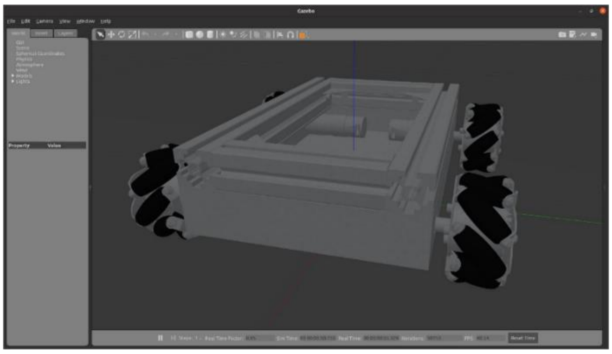


Fig.7. ArtBul chassis in the Gazebo[33]

While previously mentioned UGV and UAV models considered a recreation of existing real robots in the simulator, in[33] we presented a reverse approach of modelling chassis for omnidirectional wheeled robot *ArtBul* (Fig.7) that is still under development (Fig.8). The model was constructed using Blender software for Gazebo 11 and ROS Noetic. Next, an open-source plugin for omni-wheels was created[34].

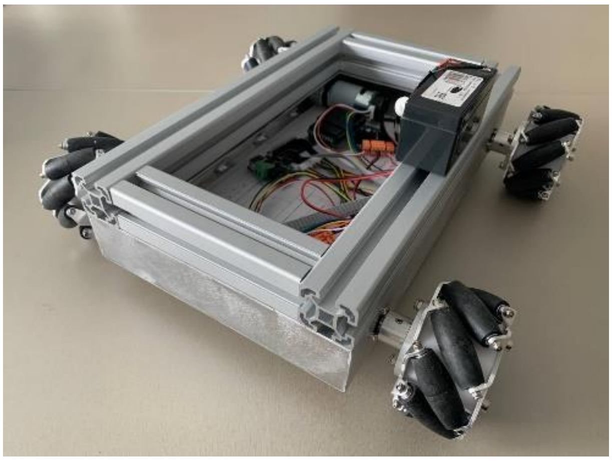


Fig.8. Current state of ArtBul chassis development[33]

3. Environment modelling

Two main approaches to environment modelling are a manual modelling and an automated modelling. For the manual modelling that targets, for example, for a raw testing of a navigation or mapping algorithm idea, a user can employ existing models of objects that come together with the Gazebo and construct rather simple flat landscapes filled with obstacles (Fig.9, [35]). A process of such environment construction is simple, but it could take several hours even for an average complexity and an average size of an environment.

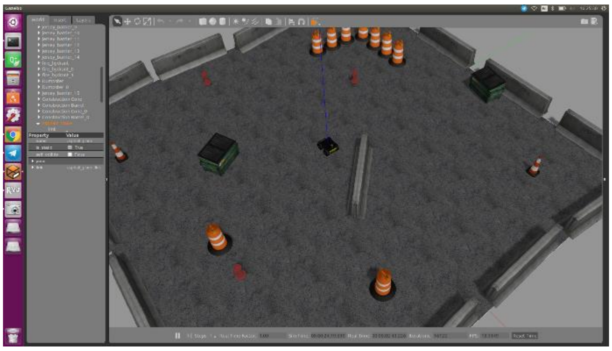


Fig.9. Robot Husky in a flat outdoor environment, which is filled with existing standard models of Gazebo objects[35]

If a user demands to recreate an existing in a real world environment (while targeting for a careful algorithm validation before transferring it from the simulator onto a real robot), precise measurements of the environment are important. The user spends at least several days making the measurements and then employs a software for constructing CAD models of the environment. In some cases, existing epura of a building could significantly speed up the process. Such approach produces lower complexity models, but it is extremely time and resource consuming. An example of manual construction based on an epura and manually performed measurements is presented in Fig.10 [36].

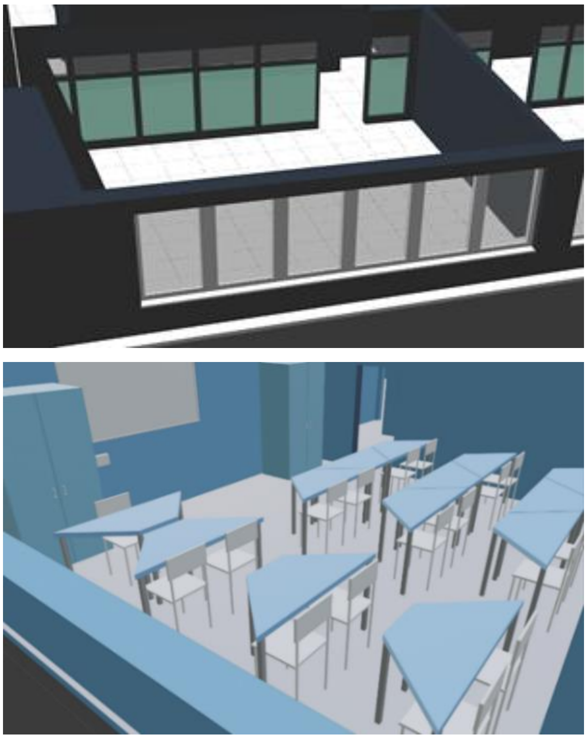


Fig.10. A part of 14-th floor environment of Department of Intelligent Robotics at Kazan Federal University: transparent glass and tinted glass windows (top) and a virtual classroom (bottom) [36]

Automated modelling is useful when a user needs to quickly create a large number of varying environments that will allow exhaustive testing of his/her algorithms in different situations and will provide statistically sufficient results for a deep further analysis. Our first attempts to automate environments' generation resulted in a low-efficiency tool that constructed voxel-environments with a high memory load and a low RTF [37]. Next, the tool was significantly modified and it is currently capable to automatically construct 3D solid model Gazebo worlds from a grayscale image and a user-selected texture[38]. Yet, not for every particular environment case a complicated software is required. For example, for a very high-structured so-called *random step environment* (RSE, [39]), which is broadly used for robot mobility evaluation within research and RoboCup rescue competitions. To automatically construct 3D RSE worlds for the Gazebo we created a user-friendly tool *LIRS-RSEGen* with a graphical user interface[40]. The tool is open-source and it allows constructing RSE worlds that could be further edited or directly used within the Gazebo. Moreover, LIRS-RSEGen can construct non-standard RSE models, for example, with different block sizes. Constructed by LIRS-RSEGen worlds were validated for their applicability in the Gazebo using virtual models of TurtleBot3 wheeled robot and Servosila Engineer crawler robot. Virtual tests demonstrated effectiveness of constructed RSE worlds in terms of the RTF, CPU and memory load, which had acceptable values even for a relatively large RSEs of 80x80 block size. Fig.11 demonstrates the Servosila Engineer robot within a generated by LIRS-RSEGen environment.

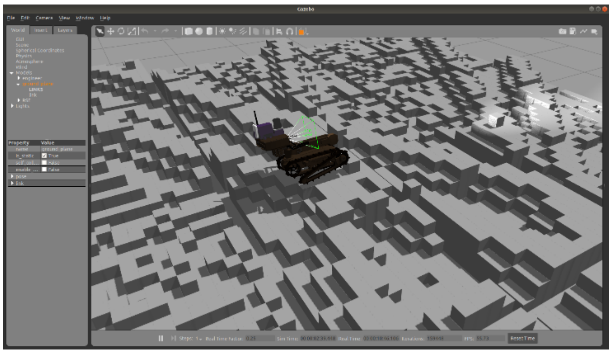


Fig.11. Gazebo 11 model of the Servosila Engineer robot within a generated by LIRS-RSEGen RSE[40]

Quality of automatically created models fells behind manual ones (e.g., large amount of generated triangles of a mesh increases a model's complexity and resource consumption, which in turn dramatically affect RTF), but automated modelling approach saves a huge amount of researchers' time and efforts.

4. Activities modelling in a virtual world of the Gazebo

This section demonstrates examples of employing simulators in order to compare and validate algorithms, methods and protocols. In[41] our team employed the Gazebo simulator for evaluation of visual SLAM methods in USAR applications. Validations of partially [42] or entirely[43] unknown environment exploration algorithms for a mobile robot were successfully performed in the Gazebo.

In[44] we presented a new concept of *Embedded ArUco* fiducial marker that allowed a high precision UAV landing. The marker performance was validated in the Gazebo. In[45] we proposed testing procedures architecture for establishing a fiducial marker recognition quality in UAV-based visual marker tracking task in the Gazebo. In general, the Gazebo simulator is critical when a large number experiments is required, e.g., we successfully employed it for exhaustive simulation approach for a virtual camera calibration evaluation[46] and automated fiducial marker comparison in the Gazebo environment[47].

The Gazebo could be successfully applied in human-robot related activities preliminary studies, e.g., human-following by a mobile robot in an indoor scenario[48] (which was initially tested in a virtual world and then validated in real world experiments[49]) or robot control using gestures[50]. The Gazebo simulator significantly facilitated a preparatory stage of real-world experiments in a humanoid robot assisted English language teaching[51], which allowed saving time and resources both for our research team and for experiments' participants. Moreover, the simulator is useful in large concepts' modelling, e.g., we used it to check an idea of a TurtleBot3 based delivery system for a smart hospital environment[52], which is expected to become a near future of pandemic mitigation approaches[53].

In[54] the Gazebo was used for comparative analysis of ROS-based centralized methods for conducting collaborative monocular visual SLAM using a pair of UAVs, which is important for large-size outdoor territory surveillance in a USAR scenario. In[55] we modelled a stairs recognition algorithm for a mobile robot navigation in 3D environments, which targets for stairs negotiation in indoor surveillance in a USAR mission. Two approaches for employing external infrastructure's IoT cameras were explored using the Gazebo: a single UGV based indoor localization[56] and an IoT cameras network application for an UAV Control via a GUI[57].

5. Conclusion

This paper overviewed our experience in using the Gazebo and Webots simulators for urban search and rescue (USAR) robotics projects. We presented particular examples of simulators' applications for constructing new models of existing mobile robots as well as modelling of a new robot that was still under development, capabilities of the simulators for manual and automated generation of virtual environments for virtual testing, their applicability for algorithm validation and comparative analysis. Overall, the Gazebo and Webots simulators demonstrated that they could significantly contribute to a broad variety of projects in USAR, human-robot interaction and other fields in terms of dramatically reducing research teams' time and resources consumption.

Acknowledgment

This paper has been supported by the Kazan Federal University Strategic Academic Leadership Program ("PRIORITY-2030").

References

1. Li Y, Li M, Zhu H, Hu E, Tang C, Li P, You S. Development and applications of rescue robots for explosion accidents in coal mines[J]. Journal of Field Robotics, 2020, Vol. 37(3): 466-489.
2. Magid E, Tsubouchi T. Static Balance for Rescue Robot Navigation-Translation Motion Discretization Issue within Random Step Environment[C]// Proceedings of the 7th International Conference on Informatics in Control, Automation and Robotics, 2010: 415-422.
3. Wang H, Zhang C, Song Y, Pang B, Zhang G. Three-dimensional reconstruction based on visual SLAM of mobile robot in search and rescue disaster scenarios[J]. Robotica, 2020, Vol. 38(2), 350-373.
4. Wagner A R. Robot-guided evacuation as a paradigm for human-robot interaction research[J]. Frontiers in Robotics and AI, 2021, Vol. 8, 701938.
5. Murphy, R R. Search and rescue robotics. Springer handbook of robotics[B], 2007.
6. Magid E, Pashkin A, Simakov N, Abbyasov B, Suthakorn J, Svinin M, Matsuno F. Artificial intelligence based framework for robotic search and rescue operations conducted jointly by international teams[C]//Smart Innovation, Systems and Technologies, 2019, Vol.154: 15-26.
7. Mingachev E, Lavrenov R, Tsoy T, Matsuno F, Svinin M, Suthakorn J, Magid E. Comparison of ROS-based monocular visual SLAM methods: DSO, LDSO, ORB-SLAM2 & DynaSLAM[C]//Lecture Notes in Computer Science, 2020, Vol. 12336: 222-233.
8. Malov D, Edemskii A, Saveliev A. Proactive localization system as a part of a cyberphysical smart environment[C]//International Conference on Industrial Engineering, Applications and Manufacturing. IEEE, 2019: 1–5.
9. Sagitov A, Gavrilova L, Tsoy T, Li H. Design of simple one-arm surgical robot for minimally invasive surgery[C]//International Conference on Developments in eSystems Engineering (DeSE). IEEE, 2019: 500-503.
10. Khusnutdinov K, Sagitov A, Yakupov A, Meshcheryakov R, Hsia K-H, Martinez-Garcia, E A, Magid E. Development and Implementation of Grasp Algorithm for Humanoid Robot AR-601M[C]//Proceedings of the 16th International Conference on Informatics in Control, Automation and Robotics, 2019, Vol.2: 379-386.
11. Pashkin A, Lavrenov R, Zakiev A, Svinin M. Pilot communication protocols for group of mobile robots in USAR scenarios[C]//International Conference on Developments in eSystems Engineering (DeSE). IEEE, 2019: 37-41.
12. Chebotareva E, Gavrilova L. Educational Mobile Robotics Project "ROS-controlled Balancing Robot" Based on Arduino and Raspberry Pi [C]// International Conference on Developments in eSystems Engineering (DeSE). IEEE, 2019: 209-214.
13. Simakov N, Lavrenov R, Zakiev A, Safin R, Martınez-Garcıa E A. Modeling usar maps for the collection of information on the state of the environment[C]//International Conference on Developments in eSystems Engineering (DeSE). IEEE, 2019: 918–923.
14. Shabalina K, Sagitov A, Su K L, Hsi K-H, Magid E. Avrora Unior Car-like Robot in Gazebo Environment[C]//International Conference on Artificial Life and Robotics, 2019: 116-119.
15. Rong G et al. Lgsvl simulator: A high fidelity simulator for autonomous driving[C]// IEEE 23rd International conference on intelligent transportation systems. IEEE, 2020: 1-6.
16. Foundation, O. S. R. (2021). Gazebo official site. http://gazebosim.org/
17. Safin R, Lavrenov R, Martinez-Garcia E A, Magid E. ROS-based Multiple Cameras Video Streaming for a Teleoperation Interface of a Crawler Robot[J]. Journal of Robotics, Networking and Artificial Life, 2018, Vol. 5(3): 184-189.
18. Pütz S, Wiemann T, Hertzberg J. Tools for visualizing, annotating and storing triangle meshes in ros and rviz[C]//European Conference on Mobile Robots. IEEE, 2019: 1-6.
19. Michel O. Webots: Symbiosis between virtual and real mobile robots[C]//International Conference on Virtual Worlds, 1998: 254-263.
20. Tsoy T, Safin R, Magid E, Saha S K. Estimation of 4-DoF manipulator optimal configuration for autonomous camera calibration of a mobile robot

using on-board templates[C]//Siberian Conference on Control and Communications, 2021: 9438925.

21. Balakirsky S, Scrapper C, Carpin S, Lewis M. Usarsim: a robocup virtual urban search and rescue competition[J]. Unmanned Systems Technology IX, 2007, SPIE, Vol. 6561: 498-508.
22. Rani P, Chauhan N R. Coal mine rescue robot simulation using V-rep and python[C]// Advances in Interdisciplinary Engineering. Springer, Singapore, 2019: 733-739.
23. Imameev D, Shabalina K, Sagitov A, Su K-L, Magid E. Modelling Autonomous Parallel Parking Procedure for Car-like Robot Avrora Unior in Gazebo Simulator[C]//International Conference on Artificial Life and Robotics, 2020: 428-431.
24. Moskvin I, Lavrenov R. Modeling Tracks and Controller for Servosila Engineer Robot[C]//Smart Innovation, Systems and Technologies, 2019, Vol.154: 411-422.
25. Khazetdinov A et al. RFID-based Warehouse Management System Prototyping Using a Heterogeneous Team of Robots[C]//International Conference on Climbing and Walking Robots and Support Technologies for Mobile Machines, 2020: 263-270.
26. Sokolov M, Afanasyev I, Lavrenov R, Sagitov A, Sabirova L, Magid E. Modelling a crawler-type UGV for urban search and rescue in Gazebo environment [C]//International Conference on Artificial Life and Robotics, 2017: 360-363.
27. Moskvin I et al. Modelling a Crawler Robot Using Wheels as Pseudo-Tracks: Model Complexity vs Performance[C]//IEEE International Conference on Industrial Engineering and Applications, 2020: 235-239.
28. Dobrokvashina A et al. Improving model of crawler robot Servosila Engineer for simulation in ROS/Gazebo [C]//International Conference on Developments in eSystems Engineering (DeSE). IEEE, 2020: 212-217.
29. Gabdrahmanov R, Tsoy T, Bai Y, Svinin M, Magid E. Gear Wheels based Simulation of Crawlers for Mobile Robot Servosila Engineer[C]//The 19th International Conference on Informatics in Control, Automation and Robotics, 2022: 565-572.
30. Dobrokvashina A, Sulaiman S, Gamberov T, Hsia K-H, Magid E. New Features Implementation for Servosila Engineer Model in Gazebo Simulator for ROS Noetic[C]//International Conference on Artificial Life and Robotics, 2023 (in press).
31. Dobrokvashina A, Lavrenov R, Magid E, Bai Y, Svinin M, Meshcheryakov R. Servosila Engineer Crawler Robot Modelling in Webots Simulator[C]//International Journal of Mechanical Engineering and Robotics Research, 11(6), 2021: 417-421.
32. Dobrokvashina A, Lavrenov R, Bai Y, Svinin M, Magid E. Sensors modelling for Servosila Engineer crawler robot in Webots simulator[C]//Moscow Workshop on Electronic and Networking Technologies, 2022: 1-5.
33. Apurin A. et al. LIRS-ArtBul: Design, Modelling and Construction of an Omnidirectional Chassis for a Modular Multipurpose Robotic Platform[C]//International Conference on Interactive Collaborative Robotics. Springer, Cham, 2022: 70-80.
34. Apurin A et al. Omniwheel Chassis' Model and Plugin for Gazebo Simulator[C]//International Conference on Artificial Life and Robotics, 2023 (in press).
35. Lavrenov R, Magid E, Matsuno F, Svinin M, Suthakorn J. Development and implementation of spline-based path planning algorithm in ROS/Gazebo environment[J]. Informatics and Automation, 2019, Vol. 18(1): 57-84.
36. Abbyasov B, Kononov K, Tsoy T, Martinez-Garcia E A, Magid E. Experience in Efficient Real Office Environment Modelling in Gazebo: a Tutorial [C]//International Conference on Artificial Life and Robotics, 2022: 673-677.
37. Lavrenov R, Zakiev A, Magid E. Automatic mapping and filtering tool: From a sensor-based occupancy grid to a 3D Gazebo octomap[C]// International Conference on Mechanical, System and Control Engineering, 2017: 190-195.
38. Abbyasov B, Lavrenov R, Zakiev A, Yakovlev K, Svinin M, Magid E. Automatic Tool for Gazebo World Construction: From a Grayscale Image to a 3D Solid Model[C]//International Conference on Robotics and Automation. IEEE, 2020: 7226-7232.
39. Magid E, Tsubouchi T, Koyanagi E, Yoshida T, Tadokoro S. Controlled Balance Losing in Random Step Environment for Path Planning of a Teleoperated Crawler Type Vehicle[J]. Journal of Field Robotics, 2011, Vol. 28(6): 932-949.
40. Gabdrahmanov R, Tsoy T, Bai Y, Svinin M, Magid E. Automatic Generation of Random Step Environment Models for Gazebo Simulator[C]// Lecture Notes in Networks and Systems, 2021, Vol.324: 408-420.
41. Safin R, Lavrenov R, Martinez-Garcia E A. Evaluation of Visual SLAM Methods in USAR Applications Using ROS/Gazebo Simulation[C]// Smart Innovation, Systems and Technologies, 2020, Vol. 187: 371-382.
42. Zakiev A et al. Partially unknown environment exploration algorithm for a mobile robot[J]. Journal of Advanced Research in Dynamical and Control Systems, 2019, Vol.11(8): 1743-1753.
43. Mavrin I, Tsoy T, Magid E. Modified E3 exploration algorithm for unknown environments with obstacles[C]//Asian Control Conference, 2022: 1413-1418.
44. Khazetdinov A et al. Embedded ArUco: a novel approach for high precision UAV landing [C]// Siberian Conference on Control and Communications, 2021: 9438855.

45. Kilin M. et al. Testing Procedures Architecture for Establishing a Fiducial Marker Recognition Quality in UAV-based Visual Marker Tracking Task in Gazebo Simulator[C]//International Conference on Artificial Life and Robotics, 2022: 691-694.
46. Tsoy T, Safin R, Martinez-Garcia E A, Roy S D, Saha S K, Magid E. Exhaustive simulation approach for a virtual camera calibration evaluation in Gazebo[C]//International Conference on Automation, Robotics and Applications. IEEE, 2022: 233-238.
47. Shabalina K, Sagitov A, Li H, Martinez-Garcia E A, Magid E. Virtual Experimental Stand for Automated Fiducial Marker Comparison in Gazebo Environment[C]//International Conference on Artificial Life and Robotics, 2018: 411-414.
48. Chebotareva E, Hsia K H, Yakovlev K, Magid, E. Laser rangefinder and monocular camera data fusion for human-following algorithm by PMB-2 mobile robot in simulated Gazebo environment[C]// Proceedings of 15th International Conference on Electromechanics and Robotics “Zavalishin's Readings”, 2021: 357–369.
49. Chebotareva E et al. Person-Following Algorithm Based on Laser Range Finder and Monocular Camera Data Fusion for a Wheeled Autonomous Mobile Robot[C]//Lecture Notes in Computer Science, 2020, Vol. 12336: 21-33.
50. Nikiforov N et al. Pilot studies on Avrora Unior car-like robot control using gestures[C]//Smart Innovation, Systems and Technologies, 2021, Vol.232: 271-283.
51. Gavrilova L, Kotik A, Tsoy T, Martinez-Garcia E A, Svinin M, Magid E. Facilitating a preparatory stage of real-world experiments in a humanoid robot assisted English language teaching using Gazebo simulator[C]//International Conference on Developments in eSystems Engineering (DeSE). IEEE, 2020: 222-227.
52. Safin R, Lavrenov R, Hsia K-H, Maslak E, Schiefermeier-Mach N, Magid E. Modelling a TurtleBot3 Based Delivery System for a Smart Hospital in Gazebo[C]//Siberian Conference on Control and Communications, 2021: 9438875.
53. Magid E, Zakiev A, Tsoy T, Lavrenov R, Rizvanov A. Automating pandemic mitigation[J]. Advanced Robotics, 2021, Vol. 35 (9), p. 572-589.
54. Abbyasov B et al. Comparative analysis of ROS-based centralized methods for conducting collaborative monocular visual SLAM using a pair of UAVs[C]//International Conference on Climbing and Walking Robots and Support Technologies for Mobile Machines, 2020: 113-120.
55. Mustafin M, Tsoy T, Martınez-Garcıa E A, Meshcheryakov R, Magid E. Modelling mobile robot navigation in 3D environments: camera-based stairs recognition in Gazebo. Moscow Workshop on Electronic and Networking Technologies, 2022: 1-6.
56. Kononov K, Lavrenov R, Gavrilova L, Tsoy T. External RGB-D camera based mobile robot localization in Gazebo environment with real-time filtering and smoothing techniques[C]//Smart Innovation, Systems and Technologies, 2021, Vol. 232: 223-234.
57. Dubelschikov A, Tsoy T, Bai Y, Svinin M, Magid E. Intelligent System Concept of an IoT Cameras Network Application for an Unmanned Aerial Vehicle Control via a Graphical User Interface[C]// International Conference on Information, Control, and Communication Technologies, 2022: 1-4.

Authors Introduction

Prof. 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. 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.

Enhancement methodology for low light image

Xiwen Liang

College of Electronic Information and Automation, Tianjin University of Science and Technology, No.
1038 Dagu Nanlu, Hexi District, Tianjin, China, 300222

Xiaoyan Chen

College of Electronic Information and Automation, Tianjin University of Science and Technology, No.
1038 Dagu Nanlu, Hexi District, Tianjin, China, 300222

E-mail: cxywxr@tust.edu.cn, 1540227453@qq.com

www.tust.edu.cn

Abstract

In order to solve the problems such as low brightness, high noise and poor contrast in weak illumination images, there are several methods proposed to address this issue. Usually, these methods are categorized into two different ways. One is based on traditional light-based technology, the other is based on machine learning technology. The low-light image enhancement is often a challenging task because the noises in dark areas are amplified with the overall brightness and contrast of the image. With the development of machine learning techniques, deep learning networks are becoming the popular research topics recently to overcome the disadvantages of noisy dots. Based on the deep analysis of the current research work, we proposed a novel network and carried out lots of comparison experiments to analysis the performances of these methods. By training, validation and testing on the datasets, the evaluation criterious are defined and utilized to analysis the efficiency of the methods. With the results, we draw the conclusion that the efficient low-light image method can make up for the shortcomings of the environment, bring better viewers' experience and provide preprocessing for subsequent high-level computer vision tasks, such as target recognition, face recognition, semantic segmentation, etc.

Keywords: Low-light Image Enhancement, Retinex Theory, Convolutional Neural Network

1. Introduction

Due to unavoidable environmental or technical constraints such as inadequate lighting and limited exposure time, images taken under sub-optimal lighting condition are dissatisfied with backlighting, non-uniform lighting and low light. Such images taken in low-light conditions are of poor visual quality, thus affecting the visual experience. Lacking of light results in missing details in the image, the images are insufficient for many computer vision-related tasks.

Low-light image enhancement is aimed at improving image quality and restoring details lost details due to poor or uneven lighting conditions. Various image enhancement schemes have been proposed to address the issue. Traditional low light enhancement methods include histogram equalization[1][2] and Retinex models[3][4][5][6]. The former often produce unnatural image because they do not take into account the relationship of pixels to their neighbors under natural light. The latter has received relatively more attention. A typical Retinex model-based approach decomposes low-illuminance images into reflected and illuminated components by some priori or regularization. The estimated reflection component is regarded as the result of enhancement. But this method can not deal with color distortion and noise effectively.

Since the first pioneering work[7], deep learning-based low-light image enhancement has achieved great success in recent years. Compared with traditional methods, deep learning-based solutions have better accuracy, robustness, and speed, which has attracted the attention of researchers. Therefore, various deep learning methods have been proposed for low-light image enhancement for different applications. Here, we aim to explore various deep learning-centered efforts to improve poor and uneven lighting images and provide comparative analysis.

2. Related Work

Deep learning has been used for a variety of tasks over the years, including de-noising[8], de-fogging[9] and super-resolution[10]. The impressive performance of the depth model in these tasks is conducive to the application of low-light image enhancement (LIE) tasks. Compared with the end-to-end network enhancement, deep Retinex-based methods achieves better enhancement performance in most cases because of the Retinex physically explainable theory[11][12].

Since Retinex theory can well simulate color perception in human vision, LIE method based on Retinex theory has attracted wide attentions. According to Retinex theory, the image can be decomposed into two components: reflectance R and illuminance L. Mathematically, the observed image I can be expressed as Eq. (1)

$$\boldsymbol{I} = \boldsymbol{R} \cdot \boldsymbol{L} \tag{1}$$

Here R, L, and · represent reflectance, illuminance, and multiply operation, respectively. However, such decomposition is an underdetermined problem, which requires several priors and regularizers to constrain the decomposition process [13][14][15]. Designing explicit priors before fitting the data is key to getting the model to perform well. However, since Eq. (1) is an ill-posed problem, it is difficult to design a constraint function suitable for multiple scenarios.

In order to avoid the complex implicit priors in the traditional Retinex method, researchers proposed some learning-based Retinex decomposition methods and achieved good results. Shen et al.[16] proposed a three-phase model called Retinex-Net, which firstly extracted two components from low-light image, then applied denoising and brightness adjustment techniques, and finally obtained enhanced results.

In the next section, this paper will discuss some deep-learning-centered low-light level image enhancement methods, and then propose a new LIE method and compare it with the existing advanced methods for quantitative analysis. Lore et al.[7] cleverly design an autoencoder for low-light enhancement. This method uses the stacked sparse denoising automatic coding method to identify the signal features of low-light images, and two modules are designed for contrast enhancement and denoising learning. Inspired by Retinex theory, KinD [17] divides the network model into two parts, one is for regulating light and the other for eliminating noise. Subsequently, the Kind++[19] network was developed to illuminate the dark area with removing the hidden artifacts and suppressing the noise. Lv et al.[18] design a low-light image enhancement network, which is composed of multiple branches and fused features of different levels extracted from multiple subnets to finally complete the task of image enhancement. Jiang et al.[20] train a powerful unsupervised generative adversarial network (EnlightGAN) for image enhancement by using information extracted from the input itself to normalize unpaired data. However, unsupervised models require careful selection of training data, and sometimes the image will lose details after enhancement. Guo et al.[21] proposed a new zero-reference depth curve estimation network (Zero-DCE), which uses the depth network to take low light enhancement as the task of image specific curve estimation. Zero-DCE adapts the dynamic range of a given image by training a network for estimating pixel-level and higher-order curves.

3. Methodology

The low light image enhancement problem can be regarded as consisting of several sub-problems, including enhancing dark areas, eliminating degradation and restoring details. The following are the core ideas of three current popular networks and the novel method we propose in this paper.

3.1. *LLNet*

LLNet presents a application that uses a class of deep neural networks-stacked sparse noise reduction autoencoders (SSDA) to enhance natural low-light level images. To the best of my knowledge, this is the first application of the depth architecture for natural low-light image enhancement. It is trained to capture the main signal features present in low-light images, and then denoise and brighten them in an adaptive manner. This method uses the local patch-wise contrast improvement to enhance contrast such that the improvements are done relative to local neighbors to prevent overamplifying the intensities of already brightened pixels. In addition, it uses the same network to learn about noise structure and goes further to give brighter and less noisy images. They propose a method for generating training data by modifing the images from the Internet in a synthetic manner to provide simulations of low-light conditions.

LLNet examines two types of deep architectures -- (i) simultaneous learning for contrast enhancement and denoising (LLNet) and (ii) sequential learning for contrast enhancement and denoising using two modules (staged LLNet and S-LLNet). In conducting the experiment, both natural and artificial images were taken into account to estimate the performance of the network in removing noise and adjusting contrast.

The training of LLNet is the error back propagation process to minimize the reconstruction loss, which is defined by following Eq. (2):

$$L_{DA}(D;\theta) = \frac{1}{N}\sum_{i=1}^{n}\frac{1}{2}\|(\mathrm{y_i} - \hat{y}(\mathrm{x_i})\|_2^2 + \beta\sum_{j=1}^{K}KL(\hat{y}_j\|y)$$
$$+(\|w'\|_F^2) + \frac{\lambda}{2}(\|W\|_F^2)$$
(2)

Where N is the number of patches, θ is the parameter of the model, and $KL(\hat{y}_j\|y)$ is the Kullback-Leibler divergence between y (target activation) and $\hat{y}_j$ (the empirical mean live of the j-th hiding unit), which can be expressed as Eq. (3)

$$KL(\hat{y}_j\|y) = \lambda log\frac{y}{\hat{y}_j} + (1-\lambda)log\frac{1-y}{1-\hat{y}_j} \qquad (3)$$

Where $\hat{y}_j$ is defined by following Eq. (4)

$$\hat{y}_j = \frac{1}{n}\sum_{i=1}^{n}h_j(x_i) \qquad (4)$$

After the decoder weight is initialized, the error back propagation algorithm is used to fine-tune the whole pre-training network. Experiments show that the depth autoencoder can effectively learn noise details from low-light images and understand important signal features.
This method can enhance the contrast and reduce the noise well, but the processing capacity of the encoder is limited, and it is only suitable for small size image.

3.2. *KinD*

KinD fuses Retinex structures into efficient deep network designs to absorb the benefits of both Retinex (i.e., good signal structures) and deep learning (i.e., generally useful priors extracted from large datasets). KinD can be divided into three modules, namely layer decomposition, reflectivity recovery and illumination adjustment.

KinD first decomposed the low-light image into noisy reflectance and gradient-smooth illuminance, then uses U-Net to recover reflectance from noise and color distortion. KinD uses a shallow U-Net as a decomposition net, but recovering two components from an image without guidance from ground truth information is a very ill-posed problem. Since there is no real information, the decomposition network uses paired low-light and normal-light images $[I_l, I_h]$as constraints. The training loss function is given by following Eq. (5)

$$L^{LD} = L_{rec}^{LD} + 0.01L_{rs}^{LD} + 0.08L_{is}^{LD} + 0.1L_{mc}^{LD} \qquad (5)$$

Where, L_{rec}^{LD} is the reconstruction error obtained by comparing the result of recombining the output reflectance and illuminance with the normal light image. L_{rs}^{LD} represents the L2 norm of reflection similarity of two reflection components $[R_l, R_h]$. L_{is}^{LD} measures the smoothness of illumination. L_{mc}^{LD} represents the mutual consistency of the output illumination of low and normal light images.

The second stage is to improve the degradation problems in the reflectance, such as noise and color distortion problems. For this purpose, the KinD uses the reflectance component obtained from the input image of normal light as the ground-truth benchmark to train the restoration network, and its loss function is given by following Eq. (6)

$$L^{RR} = \left\|\hat{R} - R_h\right\|_2^2 - SSIM\left(\hat{R}, R_h\right) + \left\|\nabla\hat{R} - \nabla R_h\right\|_2^2 \,(6)$$

Where, SSIM(·,·) is the structural similarity, and R corresponds to restored reflectance. The third item focuses on the tightness of the texture.

In the final stage, KinD designed a illumination adjustment net to flexibly convert one lighting condition to another in order to meet different needs. It uses a lightweight network with three convolutional layers and one sigmoid layer to improve lighting conditions, and its loss function is given by following Eq. (7)

$$L^{IA} = \left\|\hat{L} - L_t\right\|_2^2 + \left\||\nabla\hat{L}| - |\nabla L_t|\right\|_2^2 \tag{7}$$

$\hat{L}$ and L_t are the adjusted illuminance map and the target illuminance map, respectively. The illumination adjustment net is more sensitive to real dark parts by adding light to relatively dark areas while leaving already-bright areas almost unchanged. Compared with traditional gamma correction, the luminance information can be adjusted more flexibly.

3.3. *Zero-DCE*

Zero-DCE does not carry out the image-to-image transformation task, but redefines the enhancement as an image-specific curve estimation problem. Specifically, Zero-DCE takes low-light images as input and produces higher-order curves as output. These curves are then used to make pixel-level adjustments to a dynamic range to enhance the input image.

The method uses L-E curves to learn the mapping between low-light images and improved quality images, where the curve parameters depend only on the input images. A quadratic L-E curve can be expressed as Eq. (8)

$$LE(I(x);\beta) = \beta I(x)(1 - I(x)) + I(x) \tag{8}$$

Where x represents the position of each pixel in the image. β is a trainable curve parameter that can adjust the curvature of the L-E curve to control the exposure level. In this method, a seven-layer simple convolutional neural network is used to learn the mapping relationship between the input image and its optimal curve parameter map. To train DCE-Net in a zero-reference manner, it uses a compound non-reference loss function as shown in Eq. (9)

$$L_{total} = L_{exp} + L_{spa} + W_{col}L_{col} + W_{tv_A}L_{tv_A} \tag{9}$$

Spatial consistency loss L_{spa} prevents loss of original information by maintaining differences between adjacent areas of the input image and the enhanced image. Exposure control loss L_{exp} can solve the common problem of underexposure in low light images. Color constant loss L_{col} is used to correct potential color deviation in enhanced images. L_{tv_A} is the total variation loss. W_{col} and W_{tv_A} are the weights of the corresponding loss function.

In order to verify the effect of this algorithm on improving application performance, the latest FACE detection algorithm DSFD[22] was used as the basic model to conduct verification experiments on DARK FACE[23] data set. After Zero-DCE enhancement, the brightness and visibility of faces in dark areas are significantly improved, and the performance of face detection is also greatly improved.

3.4. *Ours module*

We propose a dual attention-guided generative adversarial network named DAGAN for fully unsupervised low-light image enhancement. Figure 1 shows the overall architecture of the proposed model DAGAN. We adopt the encoder-decoder architecture for the generator. A Double attention module is embedded in the generator to guide image enhancement and de-noising. The global discriminator D_g is a fully convolutional network composed of seven convolutional layers. It takes the entire image of the enhanced image (I_{fake}) and the normal light image (I_{real}) as input, and outputs the discriminant result with one channel. The local discriminator is similar to it. It is a fully convolutional network composed of six convolutional layers. It takes the local image blocks of the enhanced image (I_{fake}) and the normal illumination image (I_{real}) as input, and outputs a discriminative result with one channel

For discriminators, PatchGAN is used for global and local true/false authentication. For both global and local discriminators, we use relative discriminator LSGAN as the adversarial loss which is defined by following Eq. (10)

$$L_G = W_g L_{G_g} + W_l L_{G_l} + W_c L_c \tag{10}$$

Where, L_{G_g} and L_{G_l} is the generator of global adversarial loss and local adversarial loss. L_c represents content loss. W_g, W_l and W_c are the weights of the above loss functions, respectively.

L_{G_g} is given by following Eq. (11)

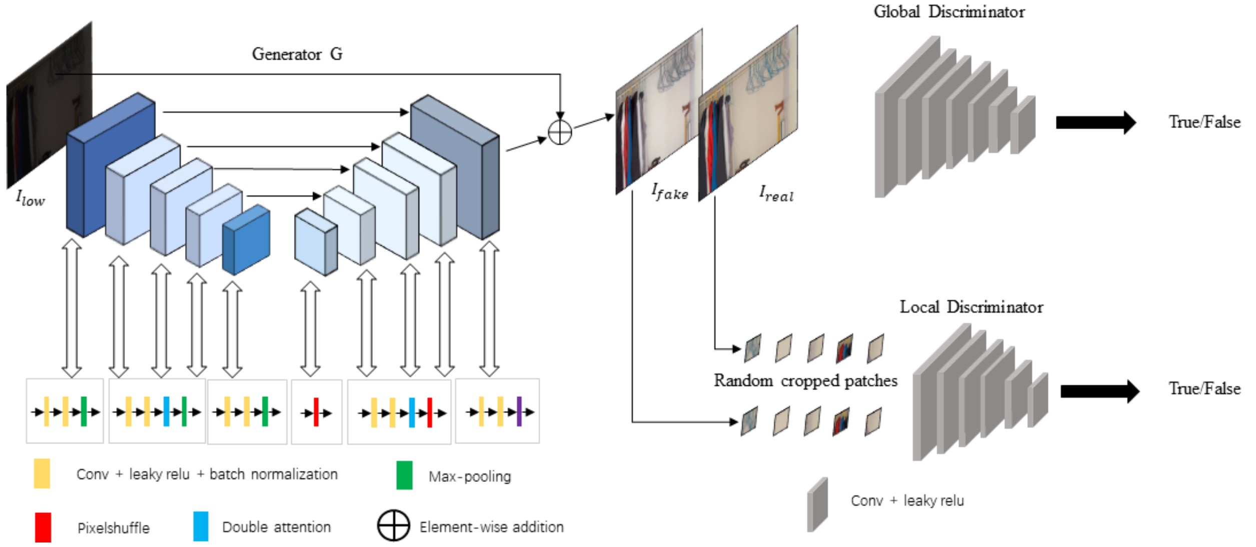

Generator G
Global Discriminator
True/False
I_{low}
I_{fake}
I_{real}
Local Discriminator
Random cropped patches
True/False
Conv + leaky relu + batch normalization
Max-pooling
Pixelshuffle
Double attention
Element-wise addition
Conv + leaky relu

Fig. 1 An overview of DAGAN consisting of one generator and two discriminators. The generator is an encoder-decoder architecture that contains a double attention module layer. The global discriminator takes the entire image as input, while the local discriminator takes patches cropped randomly from both the output image and the real reference image as input.

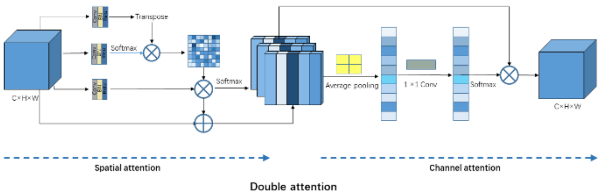

Transpose
Softmax
Softmax
C×H×W
Average pooling
1×1 Conv
Softmax
C×H×W
Spatial attention
Channel attention
Double attention

Fig. 2 Double attention module layer structure diagram. It consists of two attention modules: spatial attention module and channel attention module

$$L_{Gg} = E_{x_r \in P_{real}}\left[(D_g(x_r) - E_{x_f \in P_{fake}}(D_g(x_f)))^2\right] + E_{x_f \in P_{fake}}[(D_g(x_f) - E_{x_r \in P_{real}}(D_g(x_r)) - 1)^2] \quad (11)$$

Where, D_g is the global discriminator, E(·) is the mean calculation, P_{real} is the distribution of real natural light image data distribution, P_{fake} is the distribution of image data generated for the network, and x_r and x_f are the samples in the corresponding data distribution.

L_{D_l} is given by following Eq. (12)

$$L_{D_l} = E_{x_f \in P_{fake\_patch}}[(D_l(x_f) - 1)]^2 \quad (12)$$

Where D_l is the local discriminator, and P_{fake_patch} is the image block data distribution generated by the network. In this article, the size of the image block is 32×32.

L_c is given by following Eq. (13)

$$L_c = \frac{1}{WH}\sum_{i=1}^{W}\sum_{j=1}^{H}(\varphi_l(I(i,j)) - \varphi_l(R(i,j))) \quad (13)$$

Where, W and H represent the width and height of the feature map, respectively. $\varphi_l(\,)$ represent the output of the first convolution layer after the fifth maximum poolings layer of the pre-trained VGG-16 model. I represents the input image and R represents the image generated by the network.

In this work, we have improved the double attention mechanism[24] and cleverly embedded it in the generator for low light enhancement. In a low-light image, there are some darker or under-exposed areas. Channel attention can reassign different weights to different areas, and spatial attention can better extract spatial information of images. Therefore, the attention mechanism can extract global information from low-light images and expand reception domain. This method skillfully integrates the redistributed channel feature information and spatial feature information. Figure 2 shows the structure of the double attention module layer, where H, W and C represent the dimensions of the feature map.

In this paper, the spatial attention module[25] is used to expand the current feature diagram, as shown in the left half of Figure 2. In the right half of Figure 2, the

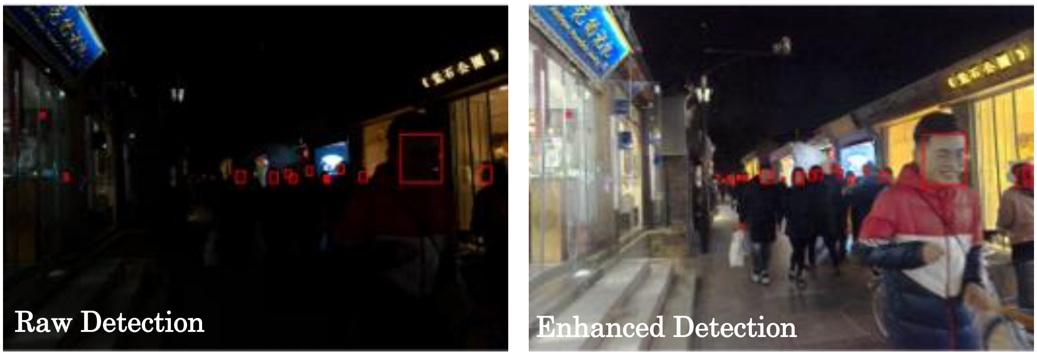


Fig. 3 Face detection results of a group of low light images before and after enhancement

interdependencies between channel feature maps are used to build the channel attention module[26].

4. Experimental Evaluations

In this section, we will conduct quantitative and qualitative experiments on the models mentioned above. It verifies the effect of the proposed DAGAN on the DSFD dataset for face detection under dark conditions.

4.1. *Dataset and Metrcis*

We trained DAGAN on the unpaired dataset collected in reference[20], which contained 914 low-light images and 1016 normal-light images. To evaluate the performance of DAGAN, we conduct quantitative analysis on three public datasets (LOL[27], DICM[28] and NPE[29]). The LOL dataset contains the low/normal light image pairs, while the other two DICM and NPE don't contain paired images.

We compare the performance of LLNet[7], KinD[17], MBLLEN[18], Zero-DCE[21], Kind++[19], Enlight GAN[20] and the proposed DAGAN.

4.2. *Quantitative Assessment*

As the LOL dataset has paired images, we conducted a quantitative comparison between PSNR, SSIM and LPIPS[30] on the LOL dataset. Since neither DICM nor NPE datasets have paired reference images, we calculated no-reference metrics NIQE[31]. Among these metrcis, PSNR and SSIM are widely used image quality assessment metrics in low-level visual tasks to evaluate the similarity between enhanced results and real reference images.

Method\ Metrics	PSNR	SSIM	LPIPS
LLNet	28.12	0.51	0.34
KinD	27.99	0.77	0.16
MBLLEN	28.07	0.78	0.21
Zero-DCE	27.82	0.66	0.31
Kind++	28.16	0.76	0.18
EnlightGAN	27.80	0.73	0.29
DAGAN(Ours)	**28.31**	**0.79**	**0.15**

Comparison of LOL Datasets based on PSNR, SSIM and LPIPS metrics

Method	DICM	NPE
LLNet	4.50	4.29
KinD	4.79	4.45
MBLLEN	3.98	4.30
Kind++	4.57	4.13
DAGAN(Ours)	**3.77**	**3.90**

Comparison of DICM and NPE Datasets based on NIQE metrics

Learning perceptual image patch similarity (LPIPS) is also known as perceptual loss. Compared with traditional metrics, LPIPS is obtained by computing the distances between features, which is more suitable for human visual perception of texture. The lower LPIPS indicates that the enhanced image has higher perceptual similarity with the corresponding ground truth.

Natural Image Quality Evaluator (NIQE) is a well-known non-reference image quality assessment metric for evaluating image restoration performance without ground truth. The lower the NIQE value, the closer the enhanced image is to the natural image.

5. Conclusions

Through a detailed analysis of LLNet, KinD, MBLLEN, Zero-DCE, Kind++, EnlightGAN and DAGAN, it can be found that each method has its advantages and disadvantages. Some give better results in terms of visual appeal, but at the cost of fuzzy details. LLNet, KinD and KinD++ generally fall into this category. On the other hand, some methods focus on details, such as EnlightGAN and MBLLEN, but bring color distortion and noise amplification problems due to over-focus on detail restoration. In extremely dark conditions, Zero-DCE can enhance images quickly and achieve good results. From the calculation results, the proposed DAGAN can effectively improve image quality and restore color and detail better. In order to further prove the practicability of DAGAN, the application experiment in this paper is carried out on the DARK FACE dataset. The DARK FACE dataset consists of

10,000 outdoor images taken in the dark. This paper randomly selects 100 images from the validation set for evaluation. After DAGAN preprocessing, the average face detection (AP) accuracy of the detector DSFD increases from 7.1% to 45.0%, which shows that DAGAN can improve the performance of computer vision tasks. Figure 3 shows an example of the face detection results. It can be seen that DAGAN algorithm improves the brightness and visibility of faces in dark areas, thus improving the performance of face detection.

Acknowledgment

This work was supported by the “Tianjin University of Science and Technology-Ansoft Technology Intelligent edge computing Joint Laboratory”

References

1. H. Ibrahim and N. S. P . Kong, “Brightness preserving dynamic histogram equalization for image contrast enhancement,” TCE, vol. 53, no. 4, pp. 1752–1758, 2007.
2. M. Abdullah-AI-Wadud, M. H. Kabir, M. A. A. Dewan, and O. Chae, “A dynamic histogram equalization for image contrast enhancement,” TCE, vol. 53, no. 2, pp. 593–600, 2007.
3. S. Wang, J. Zheng, H. Hu, and B. Li, “Naturalness preserved enhancement algorithm for non-uniform illumination images,”TIP, vol. 22, no. 9, pp. 3538–3548, 2013.
4. X. Fu, Y . Liao, D. Zeng, Y . Huang, X. Zhang, and X. Ding, “A probabilistic method for image enhancement with simultaneous illumination and reflectance estimation,” TIP, vol. 24, no. 12, pp.4965–4977, 2015.
5. X. Guo, Y . Li, and H. Ling, “LIME: Low-light image enhancement via illumination map estimation,” TIP, vol. 26, no. 2, pp. 982–993, 2016.
6. S. Park, S. Yu, B. Moon, S. Ko, and J. Paik, “Low-light image enhancement using variational optimization-based retinex model,” TCE, vol. 63, no. 2, pp. 178–184, 2017.
7. K. G. Lore, A. Akintayo, and S. Sarkar, “LLNet: A deep autoencoder approach to natural low-light image enhancement,” PR, vol. 61, pp. 650–662, 2017.
8. K. Zhang, W. Zuo, Y . Chen, D. Meng, and L. Zhang, “Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising,” IEEE Transactions on Image Processing, vol. 26, no. 7, pp. 3142-3155, 2017.
9. B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, “Dehazenet: An end-to-end system for single image haze removal,” IEEE Transactions on Image Processing, vol. 25, no. 11, pp. 5187-5198, 2016.
10. C. Dong, C. C. Loy, K. He, and X. Tang, “Image super-resolution using deep convolutional networks,” IEEE transactions on pattern analysis and machine intelligence, vol. 38, no. 2, pp. 295-307, 2015.
11. E. H. Land, “An alternative technique for the computation of the designator in the retinex theory of color vision,” National Academy of Sciences, vol. 83, no. 10, pp. 3078–3080, 1986.
12. D. J. Jobson, Z. ur Rahman, and G. A. Woodell, “Properties and performance of a center/surround retinex,” TIP, vol. 6,no. 3, pp. 451–462, 1997.
13. Shijie Hao, Xu Han, Yanrong Guo, Xin Xu, and Meng Wang, “Low-light image enhancement with semi-decoupled decomposition,” IEEE transactions on multimedia, 22(12):3025– 3038, 2020.
14. Jie Hu, Li Shen, and Gang Sun, “Squeeze-and-excitation networks,” In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 7132–7141, 2018.
15. D. J. Jobson, Z.-u. Rahman, and G. A. Woodell, “A multiscale retinex for bridging the gap between color images and the human observation of scenes,” IEEE Transactions on Image processing, vol. 6, no. 7, pp. 965–976, 1997.
16. L. Shen, Z. Yue, F. Feng, Q. Chen, S. Liu, and J. Ma, “Msr-net: Low-light image enhancement using deep convolutional network,” arXiv preprint arXiv:1711.02488, 2017.
17. Zhang, Y., Zhang, J., Guo, X. “Kindling the Darkness: A Practical Low-light Image Enhancer,” In:Proceedings of the 27th ACM international conference on multimedia. 1632-1640 (2019).
18. Lv F, Lu F, Wu J, et al. “MBLLEN: Low-Light Image/Video Enhancement Using CNNs,” In British Machine Vision Conference, 2018: 220–232.
19. Zhang, Y.H., et al. “Beyond brightening lowlight images,” International Journal of Computer Vision. 129(4), 1013-1037 (2021)
20. Y . Jiang, X. Gong, D. Liu, Y . Cheng, C. Fang, X. Shen, J. Yang, P . Zhou, and Z. Wang, “EnlightenGAN: Deep light enhancement without paired supervision,” TIP, vol. 30, pp. 2340–2349, 2021.
21. C. Guo, C. Li, J. Guo, C. C. Loy , J. Hou, S. Kwong, and R. Cong, “Zero-reference deep curve estimation for low-light image enhancement,” in CVPR, 2020, pp. 1780–1789.
22. Jian Li, Yabiao Wang, Changan Wang, Ying Tai, Jianjun Qian, Jian Yang, Chengjie Wang, Jilin Li, and Feiyuen Huang. Dsfd, “Dual shot face detector,” In CVPR, 2019.
23. Shuo Yang, Ping Luo, Chen-Change Loy, and Xiaoou Tang, “Wider face: A face detection benchmark,” In CVPR, 2016.
24. Fu, J. , et al, “Dual Attention Network for Scene Segmentation,” In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 3146–3154.
25. Zhang, H. , et al. “Self-Attention Generative Adversarial Networks,” In International conference on machine learning, 2019: 7354–7363.

26. Wang Q, Wu B, Zhu P, et al, "ECA-Net: efficient channel attention for deep convolutional neural networks," In CVF Conference on Computer Vision and Pattern Recognition(CVPR). IEEE, 2020.
27. C. Wei, W. Wang, W. Yang, and J. Liu, "Deep retinex decomposition for low-light enhancement," arXiv preprint arXiv:1808.04560, 2018.
28. Lee, C., Lee, C., Kim, C.S, "Contrast enhancement based on layered difference representation of 2D histograms," IEEE transactions on image processing. 22(12), 5372-5384 (2013)
29. Shuhang., et al. "Naturalness preserved enhancement algorithm for non-uniform illumination images" IEEE Transactions on Image Processing. 22(9), 3538-3548 (2013)
30. Zhang, R., et al. "The unreasonable effectiveness of deep features as a perceptual metric," In: Proceedings of the IEEE conference on computer vision and pattern recognition. 586-595 (2018)
31. Mittal, A., et al. "Making a "completely blind" image quality analyzer," IEEE Signal processing letters. 20(3), 209-212 (2012)

Authors Introduction

Mr.Xiwen Liang

He received his bachelor's degree from the school of electronic information and automation of Tianjin University of science and technology in 2021. He is acquiring for his master's degree at Tianjin University of science and technology.

Ms.Xiaoyan Chen

She, professor of Tianjin University of Science and Technology, graduated from Tianjin University with PH.D (2009), worked as a Post-doctor at Tianjin University (2009.5-2015.5). She had been in RPI, USA with Dr. Johnathon from Sep.2009 to Feb.2010 and in Kent, UK with Yong Yan from Sep-Dec.2012. She has researched electrical impedance tomography technology in monitoring lung ventilation for many years. Recently, her research team is focus on the novel methods through deep learning network models.

A New Style of Research and Development from the EU Perspective

Masato Nakagawa
Denso Corporation, Fellow
1-1 Showa, Kariya, Aichi, 448-8661, Japan

Hiroshima University, Guest Professor
1-3-2 Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8511, Japan

E-mail: masato.nakagawa@jp.denso.com

Abstract

This paper introduces a new style of research and development with a unique process of engineering development based on the EU perspective. Two different fields will be separated in the new style: competition field and non-competition field. The former is a specific area encouraging companies to develop their unique technology as differentiation strategy, and the later stresses collaborations among different companies and organizations for spreading the standardization of common technologies. This new style with two different fields shows us a new direction of the engineering development in various engineering industry. In addition, this paper explains robot technologies in the manufacturing of automotive sector in terms of smart manufacturing concept.

Keywords: competition field, non-competition field, Factory-IoT, collaborative robots

1. Introduction

This paper will cover two aspects of topics. One is the manufacturing (MONOZUKURI) technology including IoT technology. The other is a new style of research and development with a unique process of engineering development based on the EU perspective[1] [4].

With respect to manufacturing technology, in Europe, Smart-Manufacturing is the key word in terms of MONOZUKURI. Compared to Europe, there is similar approach of manufacturing in Japan. This paper describes the DENSO's "Factory-IoT". DENSO Corporation is one of the automotive Tier-1 suppliers.

2. DENSO's "Factory-IoT"

1.1 Roadmap of Production System

DENSO is a manufacturing company mainly for automotive components and systems. DENSO started using the single automated manufacturing process like a spot machine in the 1950's. After that DENSO expanded automation to the production line unit and then, further expanded to factory unit. Furthermore, DENSO expanded the automation to global unit like a global network.

One of the features of DENSO manufacturing is the in-house development and fabrication of the robot machines since the early 1970's.

Figure 1 represents the roadmap of the DENSO production system.

With respect to automation and robotics, DENSO uses automation technology, mainly in Assembly and Visual Inspection processes and for in-plant logistics by using in-house machines and robots. There is a however, still human-based manufacturing process left for these three fields. DENSO has been working on these fields by using the intelligent technology of robots. One of the features of this technology is "collaborative robots" which means that two-robots collaborate and co-operate each other. It contributes to working efficiency and quality in the manufacturing plant.

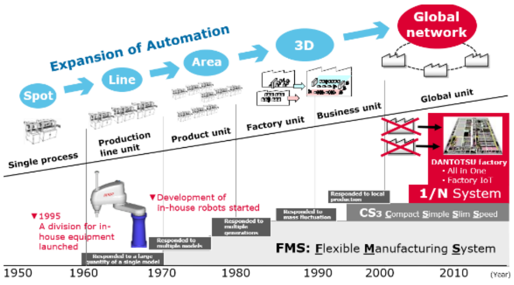

Expansion of Automation
Spot
Line
Area
3D
Global network
Single process
Production line unit
Product unit
Factory unit
Business unit
Global unit
DANTOTSU factory
・All in One
・Factory IoT
1/N System
▼1995
A division for in-house equipment launched
▼Development of in-house robots started
Responded to a large quantity of a single model
Responded to multiple models
Responded to multiple generations
Responded to mass fluctuation
Responded to local production
CS3 Compact Simple Slim Speed
FMS: Flexible Manufacturing System
1950
1960
1970
1980
1990
2000
2010
(Year)

Figure 1 Roadmap of Production System

1.2 *Principe of "Factory-IoT"*

It is said that IoT technology is a tool for keeping maintenance of stable production in the field of manufacturing (MONOZUKURI). DENSO's unique point of IoT is that human is involved in this process. By utilizing human skill and knowledge, sustainable growth and continuous evolution can be achieved. This means that both human and machine can provide the optimal solutions thanks to co-creation by human and machine. DENSO pursuits the "Factory-IoT" technology for all of plants globally by connecting each plant.

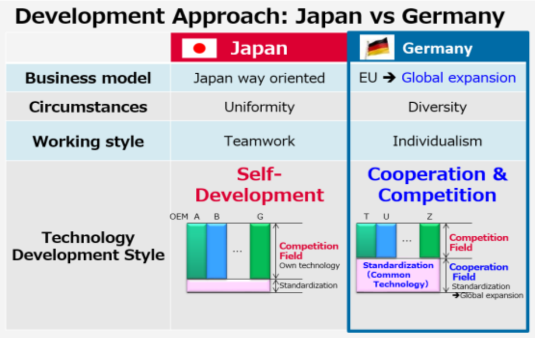
3. EU-Way vs JP Way Development

3.1. *Comparison between EU- Way and JP-Way*

Figure-2 represents the comparison between EU and JP development ways[2]. This chart summarizes the main features development ways, working styles and business models based on the Automotive sector of Europe and Japan. In this chart, for EU, German development way is used as a typical example. There are significant and remarkable differences in the EU development way compared to that of Japan-Way. In Germany, there is a clearly two fields of development style. One is "Competition Field" and the other is "Non-competition Field" for the engineering development. They co-operate and collaborate among the same industry domain in the base technology field. They establish the common technology field like a standardization and/or regulations. Thanks to this system, they can concentrate on their resources on the development within the unique technology field. Then, they are able to create the differentiated technology. On the other hand, in Japan, OEMs develop almost all technologies by themselves.

In this approach, they are able to create unique technologies, however, recently there are various new technologies to be developed like a CASE fields (**C**onnected, **A**utonomous Driving, **S**haring, **E**lectrification) .

Therefore, there are concerns in terms of the lack of resources for the various field of engineering development in the Automotive sector.


Development Approach: Japan vs Germany

	Japan	Germany
Business model	Japan way oriented	EU ➔ Global expansion
Circumstances	Uniformity	Diversity
Working style	Teamwork	Individualism
Technology Development Style	Self-Development	Cooperation & Competition

OEM A B G
...
Competition Field
Own technology
Standardization
T U Z
...
Competition Field
Standardization (Common Technology)
Cooperation Field
Standardization
➔Global expansion

Figure 2 Comparison between EU-Way and JP-Way

3.2. *Collaboration Concept between Germany and Japan*

There are good points to be learnt from German-Way engineering development. For example, efficient and rational development by using industry standard too-chains, model-based development, cutting-edge simulation, automated calibration etc. In addition, Germany OEMs cooperates with R&D Engineering Companies for their vehicle and engine development in not only the research phase but also for application engineering area. There are remarkable R&D Engineering Companies in Europe who have sophisticated development tools, testing facilities and high-talented engineering resources and its know-how.

In Japan also, there are good points for their engineering development way like teamwork and harmonized development approach with dignity[3].

4. Conclusion

The key point is to respect each other of their engineering development ways, then, reflect and accept the good points into their development ways. Figure 3 represents the concept of collaboration between Europe and Japan. The important point is that both good points to be well-arranged like a fusion. In this case, a new development way which is a fusion between Europe and Japan could become a global competitive engineering approach.

Both parties should respect their engineering development ways and inspire each other in research and development field.

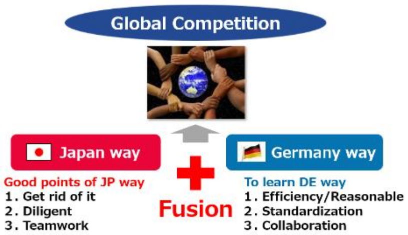

Global Competition
Japan way
Good points of JP way
1. Get rid of it
2. Diligent
3. Teamwork
Fusion
Germany way
To learn DE way
1. Efficiency/Reasonable
2. Standardization
3. Collaboration

Figure 3 Concept of Collaboration

References

1. M. Nakagawa, *26th International AVL Conference Engine 2020: Sparks versus compression ignition in a new environment* (Graz, Austria, September 2014)
2. M. Nakagawa, Dr.-Ing. O. Hermann and Dipl. Ing. Sebastian Visser, Diesel Powertrain Energy Management via Thermal Management and Electrification (Detroit SAE April 2017)
3. M. Nakagawa, DENSO’s Contribution for Future Predictive Powertrain Control including Electrification (Korea, FISITA2016)
4. M. Nakagawa, Plenary Speech on ICAROB2018 (Beppu, Oita, Japan Feb. 2018)

Author Introduction

Mr. Masato Nakagawa

He graduated the mechanical engineering at Hiroshima University, Japan. After he entered DENSO Corporation, he had been working in Europe consecutive 14 years in Germany, The United Kingdom and the Netherlands in the field of automotive sector. During his stay in Europe, he leans lots of European engineering development way.

Arduino Based Smart IoT Food Quality Monitoring System

Ashraf Ali Jamal Deen, Thivagar Chettiar Sarawanan

Department of Electrical and Electronic Engineering, Manipal International University, Malaysia

Heshalini Rajagopal
Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia

Devika Sethu
Department of Electrical and Electronic Engineering, Manipal International University, Malaysia

Neesha Jothi, Raenu Kolandaisamy
Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia
E-mail: heshalini@ucsiuniversity.edu.my
www.ucsiuniversity.edu.my

Abstract

Food safety and hygienic as well as health are significant issues to stop food wastage. The high quality of the food requires to be kept track of and it should be also protected against deteriorating and decaying by the climatic variables like temperature level, humidity, and dark. In this paper, a comparable food quality monitoring tool will be created that will keep watch of ecological factors like temperature level, moisture, alcohol web content as well as exposure to light for fruits and vegetables. The system is built on Arduino UNO where it is interfaced with various sensors like DHT-22 to keep track of temperature level and humidity, MQ3 to identify alcohol material as well as LDR to gauge direct exposure to light. It sends the measured sensor data to an IoT system via ESP8266 Wi-Fi Module. The IoT system will certainly be made use of for logging and checking sensing unit data and this is beneficial in monitoring the food storage from anywhere and anytime.

Keywords: Food Quality, Monitoring System, Internet of Things (IoT), Arduino

1. Introduction

In terms of both economic and social considerations, food safety is very essential. Many implications might result from a company's failure to meet food safety and security standards. Failure to provide adequate food safety throughout the production process may have a significant impact on the lives of those who are affected by the lack of food safety measures [1]. Contaminated foods may enter the food supply chain if proper food safety and health protocols are not followed. When a faulty product is discovered, the food service industry is vulnerable to major disruptions in their operations as they attempt to handle and remember the item's quality. The importance of food safety in today's society cannot be overlooked [2]. Problems with food safety and security are a major cause of more than 200 avoidable illnesses throughout the globe. In the United States, one in ten persons suffers from foodborne illness or injury each year. More than a quarter of the estimated 420,000 individuals who die each year as a consequence of eating poisoned food are kids [3]. Therefore, it is necessary to create a system that can assist consumers in determining whether or not food is fresh and also of high quality. In this paper, a food quality monitoring system is proposed. Several studies were done on this area. B.Yu, et al. (2020) proposed a monitoring system which integrated smart contracts and evaluation models for the automatic evaluation of the quality of fruit juice samples generated in each production stage[4]. A.Popa, et al. (2019) proposed a food quality monitoring system for vegetables stored in vacuum-packed foods[5]. However, to the best of our knowledge, there is no food quality monitoring system developed for fruits, vegetable and cooked food such as curry and milk in a single system.

2. Hardware structure

The block diagram of the design is shown in Fig.1. The system consists of power supply units, Arduino UNO microcontroller, WIFI module, Gas sensor (MQ3), LDR, Ph value sensor, DHT22 sensor, and LCD display. The flow line shows the inputs and outputs of the system. The main controlling unit of the system is Arduino Uno and it will be powered by a 12V battery. Temperature, humidity, alcohol, light exposure and moisture are monitored using the Arduino board, DHT-22, MQ3, LDR and pH sensors. Measurements can be conducted as quickly as feasible because to its high sensitivity and quick reaction time. The ESP8266 Wi-Fi Modem is connected to the internet via a Wi-Fi router via the Arduino. The sensor data is also shown on an Arduino UNO-connected character LCD. Sensor data is logged and monitored using this IoT platform. Several of these devices can be put at a site for better monitoring and quality control. The Arduino Sketch that runs on the device does a variety of tasks for the project, including receiving sensor data, converting it to strings, displaying it on a character LCD, and transmitting it to the IoT platform, Blynk Application.

2.1. *Flowchart of the system*

The DHT-22 sensor measures temperature and humidity[6]. The Humidity detecting component and the NTC temperature sensor are the two major components of the DHT-22 sensor (or Thermistor). Thermistors are variable resistors that change resistance in response to temperature changes. They both sense the surrounding region's temperature and humidity and communicate the information to the IC.

The LDR is used to detect light intensity. The LDR creates an analogue voltage, which is converted to a digital readout by the built-in ADC. The 16X2 LCD display is connected to the Arduino board by linking its data pins to the Arduino board.

The ESP8266 Wi-Fi Module is an (SOC) self-contained system on chip with that can connect to a Wi- Fi network and has an integrated TCP/IP protocol stack. The ESP8266 may either install apps or representative all Wi-Fi networking functions to a separate application processor. Each ESP8266 module comes with pre-programmed AT instruction set software. The ESP-01 and ESP-12 variants of the module are available. The ESP-12 has 16 pins accessible for the interface, whereas the ESP-01 only has 8 pins.

pH is the measurement unit that we use to determine the acidity of a chemical. The undesirable log of the hydrogen ion concentration is defined as "H." The pH values range from 0 to 14 on a scale of one to fourteen. A pH of 7 is measured neutral since distilled water has an exact pH of 7. Essential or alkaline levels are more than 7, whereas acidic values are less than 7. The pH scale is used to assess the amount of acidity and basicity in a fluid. It can have results ranging from 1 to 14, with 1 indicating the most acidic fluid and 14 indicating the most standard fluid. The pH of 7 refers to things that are neither acidic nor basic.

The purpose of an analog pH sensing device is to determine the pH value of a substance as well as the acidity or alkalinity of the substance. It is commonly used in a variety of applications such as farming, wastewater treatment, industry, environmental monitoring, and so on. The component contains an on-board voltage controller authority chip that supply a wide voltage range of 3.3-5.5V DC, making it well-matched with both 5V and 3.3 V control boards such as Arduino.

The system flowchart for this project is shown in Fig. 2. A dedicated link is established between the sensing units, the Wi-Fi component, and the microcontroller. The information from the three sensing devices is pre-processed by the microcontroller before being sent to the web server. AT commands will be given to the Wi-Fi component to interact with the pre-processed data and send it to the web server for further processing and visualization.

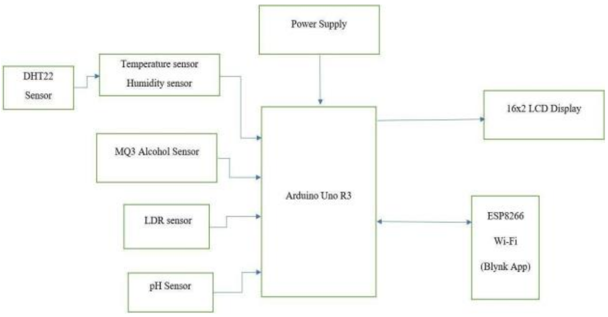


Fig. 1. Block diagram of the system

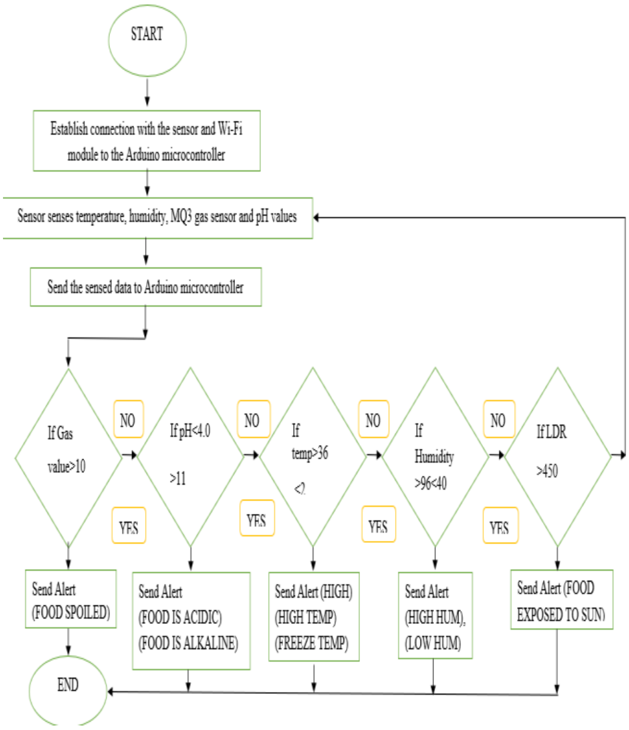

START
Establish connection with the sensor and Wi-Fi module to the Arduino microcontroller
Sensor senses temperature, humidity, MQ3 gas sensor and pH values
Send the sensed data to Arduino microcontroller
If Gas value>10
NO
If pH<4.0 >11
NO
If temp>36 <?
NO
If Humidity >96<40
NO
If LDR >450
YES
YES
YES
YES
YES
Send Alert (FOOD SPOILED)
Send Alert (FOOD IS ACIDIC) (FOOD IS ALKALINE)
Send Alert (HIGH) (HIGH TEMP) (FREEZE TEMP)
Send Alert (HIGH HUM), (LOW HUM)
Send Alert (FOOD EXPOSED TO SUN)
END

Fig.2. Flowchart of the system

The Wi-Fi module's status will be checked on a regular basis to ensure network connectivity schedule and dependability. A link reset should be performed if the status of the Wi-Fi and web server connections is discovered to be disconnected at any time. The information from the Wi- Fi module will undoubtedly be analyzed and considered by the web server. As soon as the device is turned on, the microcontroller will begin refining data over the Internet.

With the help of the Blynk Application, Arduino can be controlled via the Internet. A visual user interface is created using a digital dashboard to ease the monitoring of the sensors' readings.

3. Results and Discussion

3.1. *Detecting the Quality of Fruits*

Each fruit and vegetable have a different expiration date, fragrance, and color. Also, different food has different parameters such as humidity, temperature, and pH value. Food rotting may also be influenced by changes in temperature. When it comes to fruits and vegetables, storing them at the correct temperature is critical to extending their shelf life. It is possible for rotting to be accelerated by extreme temperatures. Water in plant cells freezes into ice crystals when frozen, causing cell wall ruptures and discoloration as well as a slimy appearance. Temperature and humidity can be measured using the DHT22 sensor. -40 to $+125^0$C, with a $+-0.5^0$C accuracy, is its temperature range. With a range of 0 to 100 percent, the DHT22 sensor can measure humidity with a 2 to 5 percent accuracy. It is discovered with the help of the DHT 22 sensor, it is possible to keep produce fresher and more flavorful by storing it at lower temperatures. The growing of pathogenic fungi that cause spoilage of fruits and vegetables in storage is also slowed by low temperatures.

Continues with MQ3 Sensor which has been used to detect alcohol parts per million. This sensor plays an important role in this system where this sensor has been used to detect fruit and vegetable whether the fruit or vegetables is spoiled or not by detecting the gas of ripening fruit emitting which is ethylene gas. Heavy exposure to direct sunlight and light air flow shows a uniform effect of accelerating ripening and decay so LDR sensor has been used to sense the light intensity in the storage box. Therefore, observation is done from the storage box of the fruit and vegetables. The results have found that the healthy fruit has strong cell walls that must contain the humidity level, preventing the fruit from decaying faster. When fruits are exposed to either high temperature or direct sunlight, its cell walls can lose their natural moisture quickly which could then lead them into a condition of dry storage. For this, we have placed an LDR sensor to measure the lux value of direct sunlight. From the LDR sensor that has been used, it has been found that in direct sunlight the value is shown 450 when it reaches 450 or more than that, it will notify the user in both Blynk App and LCD screen to alert the users. In addition, warm days will create an environment that is perfect for fungi growth as well as increase susceptibility to insects harmlessly feeding on the fleshy areas. Sunlight will also promote fungal growth due to its ability to damage cells at a cellular level through a process known as photodynamic degradation.

In this study, two types of fruits, Apple and Banana and a vegetable, tomato were used. Based on Fig.3, the ppm values for the fruits and vegetables kept increasing every day. The banana and tomato spoilt on the 4th day and 6th day, respectively. The ppm value for banana and tomato were recorded as 10 ppm and 9 ppm, respectively. The apple also started to spoil with the value above 10ppm on the 6th day. This shows that the fruit and vegetables release Ethylene gas as they decay.

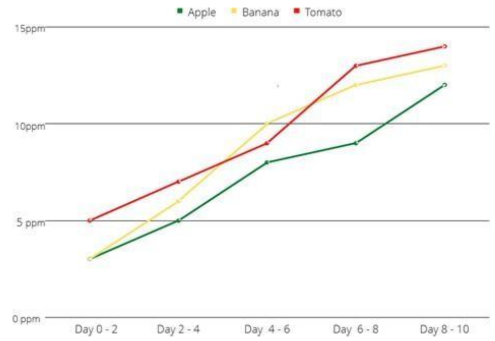


Fig.3: PPM values for Apple, Banana, and Tomato

3.2. *Monitoring System*

We used an IoT app to monitor the food in this system (in the storage box). In order to monitor temperature, humidity, light intensity, ethylene gas concentration, and pH, there are five gauge- typewidgets (see Fig.4). The gauge widget's measure has the greatest importance when seen as an average value against a maximum value. However, multiple values have been set and ranged and some to the highest value has been shown on blynk's widget app and the lowest value has been used. This will show data in the form of a gauge. The widget will show the most current data that has been gathered. When the widget is set to "gauge" mode, the data series configuration field (which displays the gathered data's maximum, minimum, or mean) has no effect on how the data is shown.

3.3. *Quality of Cooked Food*

The humidity and temperature of cooked food are effortless to detect in order to determine the quality of it. The quality of cooked food can be determined with pH value. When the food's pH value is detected from 0 to 6, the foods are acidic, if from 8 to 14 the foods are alkaline and if the pH is 7 the foods are neutral. Several foods have hard surfaces while the rest have smooth surfaces. For example, the vegetables with hard surfaces will be long-lasting whereas the vegetables with smooth surfaces will expire early. The spoiled milk, result of an over-growth of bacteria that compromises the quality, flavor, and texture of milk. Once you open a carton of milk, it's exposed to extra bacteria from the environment. Over time, these small microbial communities can reproduce and eventually cause the milk to spoil. Signs the milk has spoiled, it develops an unpleasant, acid odor where it can be found using the pH meter. Based on Fig. 5, the pH value of milk and curry decreases every day. The curry spoiled on the 2nd day and its pH value was below 6. For milk, the pH value dropped to below 4 when it is spoiled on the 3rd day. This shows that the acidity of milk and curry increases as the food is spoiled. The proposed system could measure and monitor the real time values using Blynk app for monitoring the condition of the food.

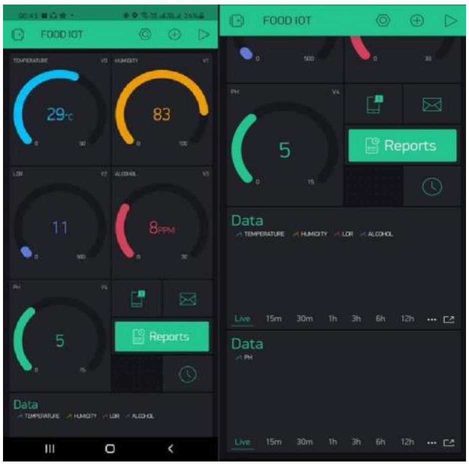


Fig. 4. Blynk Application

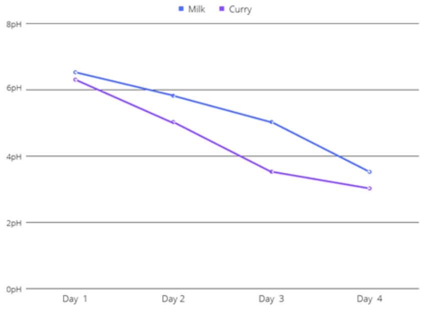


Fig. 5. pH Value for Milk and Curry

4. Conclusion

In conclusion, the system is capable to monitor the of temperature, humidity, alcohol, and light exposure in fruits and vegetables storage using DHT 22, MQ3 and LDR sensors. In addition, the proposed system could also detect if the cooked food such as curry and milk has spoilt using pH sensor. The proposed system is capable to send the data from the sensors to the IoT platform, Blynk Application. This enables the users to monitor the condition of the food from anywhere and at any time. This system is beneficial in food industries as the system could monitor the condition of the food can be monitored all the time automatically. This system can be further improved by using image processing to enhance the efficiency of the detection.

References

1. Srivastava A, Gulati A. iTrack: IoT framework for Smart Food Monitoring System[J]. International Journal of Computer Applications, 2016, 148(12): 1–4.
2. Tort Ö Ö, Vayvay Ö, Çobanoğlu E. A Systematic Review of Sustainable Fresh Fruit and Vegetable Supply Chains[J]. Sustainability (Switzerland), 2022, 14(3): 1–38.
3. Schlein L. Un: Contaminated Food Sickens 600 Million, Kills 420,000 Every Year[J]. VOA News: Science & Health, 2021.
4. Yu B, Zhan P, Lei M. Food Quality Monitoring System Based on Smart Contracts and Evaluation Models[J/OL]. IEEE Access, 2020, 8: 12479–12490.
5. Popa A, Hnatiuc M, Paun M. An intelligent IoT-based food quality monitoring approach using low-cost sensors[J]. Symmetry, 2019, 11(3).
6. Bogdan M. How to Use the DHT22 Sensor for Measuring Temperature and Humidity with the Arduino Board[J/OL]. ACTA Universitatis Cibiniensis, 2016, 68(1): 22–25.

Authors Introduction

Mr. Ashraf Ali Jamal Deen

He is currently pursuing his Bachelor’s Degree from the Department of Electrical and Electronic Engineering, Manipal International University, Malaysia.

Mr. Thivagar Chettiar Sarawanan

He is currently pursuing his Bachelor’s Degree from the Department of Electrical and Electronic Engineering, Manipal International 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. 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.

Madam Devika Sethu

She received her Master's degree from the Liverpool John Moores University (LJMU), UK in 2006. She received the B.Tech in 1991 from College of Engineering Trivandrum (CET), University of Kerala, India. Currently, she is an Assistant Professor in Manipal International University (MIU), Malaysia. Her research interest includes robotic & automation, HVAC , AI , IoT, image processing and energy management.

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. 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.

Development of Image Quality Assessment (IQA) For Haze Prediction

Heshalini Rajagopal
Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia

Sayanth Sudheer
Department of Electrical and Electronic Engineering, Manipal International University, Malaysia

Neesha Jothi, Keoy Kay Hooi
Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia

Norrima Mokhtar
Department of Electrical Engineering, Faculty of Engineering, University of Malaya, Malaysia
E-mail: heshalini@ucsiuniversity.edu.my
www.ucsiuniversity.edu.my

Abstract

Haze is a term that is widely used in image processing to refer to natural and human activity-emitted aerosols. It causes light scattering and absorption, which reduce the visibility of captured images. This reduction hinders the proper operation of many photographic and computer vision applications, such as object recognition/localization. Therefore, an approach for haze density estimation is highly demanded. This paper proposes a model that is known as the haziness degree evaluator to predict haze density from a single image without reference to a corresponding haze-free image. The proposed model quantifies haze density by optimizing an objective function comprising haze-relevant features that result from correlation and computation analysis.

Keywords: Haze, Image Quality Assessment (IQA), Hazy Images, Haze Density

1. Introduction

Haze is traditionally an atmospheric phenomenon in which dust, smoke, and other dry particulates obscure the clarity of the sky. Haze particles can sometimes affect the heart and lungs, especially in people who already have chronic heart or lung disease e.g., asthma, Chronic Obstructive Pulmonary Disease (COPD), or heart failure. There may be up to one to three days of time between exposure to haze and health effects or symptoms[1]. Haze density (or usually referred to as see through quality) is measured with a narrow angle scattering test in which light is diffused in a small range with high concentration. This test measures the clarity with which finer details can be seen through the object being tested. This is the method currently used by most of the industry to measure haze, the main apparatus used are ASTM E430, ASTM D4039, ISO 13803 these are shaped like microscope and can identify the haze density[2]. Most of the times photos taken outdoors are often degraded by haze, an atmospheric phenomenon produced by small floating particles which absorbs and scatter light from its multiple direction. Haze influences the visibility of the picture as it generates loss of contrast of the distant object in the image[3]. In computer vision applications, dehazing is applied to enhance the visibility of outdoor images by reducing the undesirable effects due to scattering and absorption caused by atmospheric particles. Dehazing is needed for human activities and in many algorithms like objects recognition, objects tracking, remote sensing and sometimes in computational photography. In bad visibility environments, such applications require dehazed images for a proper performance[3]. Therefore, studies have been conducted to remove haze and then improve the quality of degraded images[3]. In addition, a quality scale is an important indicator of the degrees of degradation and improvement. Image Quality Assessment (IQA) is well known in the scope of image processing. The IQA is classified into two methods, namely, subjective and objective methods[4]. The subjective method is used to obtain human's perception on the image quality[5]. This method simulates the perception technique of a people, a visual system to estimate it. Then, their perception is converted in the form of Mean Opinion Score (MOS) which were used to model the objective method which is the proposed IQA metric.

2. Methods

This paper proposed a model that is known as the haziness degree evaluator to predict haze density from a single image without reference to a corresponding haze-free image. The proposed model quantifies haze density by optimizing an objective function comprising haze-relevant features that result from correlation and computation analysis. First the mean of score (MOS)[4] was obtained for the hazy images from twenty (20) human subjects, then the MOS values will be validated by comparing them IQA metrics such as Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE)[7], Naturalness Image Quality Evaluator (NIQE)[8], Structural Similarity Index (SSIM)[9], Feature Similarity (FSIM)[10], Gradient Magnitude Similarity Deviation (GMSD)[11]. Next, gaussian features were extracted from the hazy images. Then, these features together with the validated MOS were fed into the Support Vector Machine (SVM) Regression (SVR) model to train the machine to map the features with the MOS where an optimized model will be obtained. The optimized model was then used to predict the quality score of the test hazy images. Then, the quality scores were used to predict the haze density on the images.

Firstly, the Mean Subtracted Contrast Normalized (MSCN) of the hazy images were calculated[6]. Then, two types of Gaussian distribution functions were incorporated in this study to accommodate the diverse characteristics of MSCN coefficient, namely the Generalized Gaussian Distribution (GGD) and Asymmetric Generalized Gaussian Distribution (AGGD)[5]. The GGD, where α represents the shape of the distribution and σ^2 represents the variance and AGGD parameters were calculated using the Eqs. (1) – (3):

$$GGD(x;\alpha,\sigma^2) = \frac{\alpha}{2\beta\Gamma\left(\frac{1}{\alpha}\right)}\exp\left(-\left(\frac{|x|}{\beta}\right)^{\alpha}\right) \tag{1}$$

Where x represents the MSCN and

$$\beta = \sigma\sqrt{\frac{\Gamma\left(\frac{1}{\alpha}\right)}{\Gamma\left(\frac{3}{\alpha}\right)}} \tag{2}$$

$$\Gamma(a) = \int_0^\infty t^{a-1}\, e^{-t}\, dt, a > 0 \tag{3}$$

The AGGD, AGGD, namely v, which represents the shape of the distribution, σ_l^2 and σ_r^2 which represent the left and right-scale parameters, respectively, and η which represents the mean of the distribution. using the Eqs. (4) – (7):

$$AGGD(x;v,\sigma_l^2,\sigma_r^2,\eta) = \begin{cases} \frac{v}{(\beta_l+\beta_r)\Gamma\left(\frac{1}{v}\right)}\exp\left(-\left(\frac{-x}{\beta_l}\right)^{v}\right) & x < 0 \\ \frac{v}{(\beta_l+\beta_r)\Gamma\left(\frac{1}{v}\right)}\exp\left(-\left(\frac{x}{\beta_r}\right)^{v}\right) & x \geq 0 \end{cases} \tag{4}$$

Where x is the MSCN calculate at four neighborhood pixels and

$$\beta_l = \sigma_l\sqrt{\frac{\Gamma\left(\frac{1}{v}\right)}{\Gamma\left(\frac{3}{v}\right)}} \tag{5}$$

$$\beta_r = \sigma_r\sqrt{\frac{\Gamma\left(\frac{1}{v}\right)}{\Gamma\left(\frac{3}{v}\right)}} \tag{6}$$

$$\eta = (\beta_r - \beta_l)\frac{\Gamma\left(\frac{2}{v}\right)}{\Gamma\left(\frac{1}{v}\right)} \tag{7}$$

The flowchart of the proposed haze prediction system is shown in Fig.1

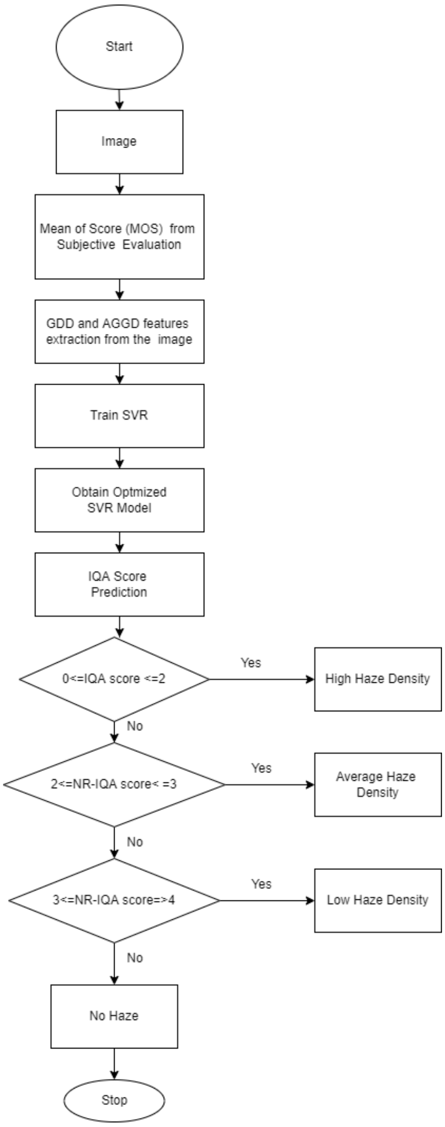

Start
Image
Mean of Score (MOS) from Subjective Evaluation
GDD and AGGD features extraction from the image
Train SVR
Obtain Optmized SVR Model
IQA Score Prediction
0<=IQA score <=2
Yes
High Haze Density
No
2<=NR-IQA score< =3
Yes
Average Haze Density
No
3<=NR-IQA score=>4
Yes
Low Haze Density
No
No Haze
Stop

Fig.1.Flowchart of the system

3. Results and Discussion

3.1. *Validation of MOS*

The MOS values obtained from the twenty human subjects were validated by using well-established IQA metrics, namely NIQE, BRISQUE, SSIM, FSIM and GMSD. The correlation between the MOS and these IQAs were calculated using Pearson's Linear Correlation Coefficient (PLCC). The PLCC values between MOS and IQAs are shown in Table 1.

Table 1. PLCC between MOS and NIQE, BRISQUE, SSIM, FSIM and GMSD

FR-IQAs	NIQE	BRISQUE	SSIM	FSIM	GMSD
PLCC	0.819	0.895	0.832	0.788	0.900

According to Taylor R. (Taylor, 1990), two datasets are said to have high correlation if the correlation coefficient values are between 0.68 to 1.0[5]. Since all the PLCC values are more than 0.68, this shows that the MOS values that were obtained from the subjective evaluation is valid and can be used to train SVR.

3.2. *Performance of the Proposed System*

The performance of the proposed system using the PLCC between the MOS, NIQE, BRISQUE, SSIM, FSIM, GMSD and the proposed IQA metrics as shown in Table 2. Based on Table 2, the proposed IQA metric recorded the highest PLCC value compared to the other IQA metrics such as BRISQUE, NIQE, SSIM, GMSD and FSIM. Hence, the proposed IQA metric is the most suitable metric to evaluate hazy images as it is very close to the MOS values.

Table 2. PLCC between MOS and NIQE, BRISQUE, SSIM, FSIM, GMSD and proposed IQA.

IQA	NIQE	BRISQUE	SSIM	FSIM	GMSD	Proposed IQA
PLCC	0.819	0.895	0.832	0.788	0.900	**0.970**

4. Conclusion

In this paper, an IQA has been developed to detect the haze density automatically from an image. The proposed system is beneficial and cost efficient where sensors are not required to detect the haze density. Furthermore, this system is capable to detect the haze density without the need of a reference images.

References

1. Othman K A, Li N, Abdullah E H. Haze monitoring system in city of kuala lumpur using zigbee wireless technology implementation[C]//Proceedings of the World Congress on Engineering 2013.
2. International A. Standard Test Method for Reflection Haze of High-Gloss Surfaces[R/OL].
3. Ancuti C, Ancuti C O, Timofte R. I-HAZE: A Dehazing Benchmark with Real Hazy and Haze-Free Indoor Images[J]. Lecture Notes in Computer Science (including subseries Lecture Notes in

Artificial Intelligence and Lecture Notes in Bioinformatics), 2018, 11182 LNCS: 620–631.

4. Rajagopal H, Mokhtar N, Khairuddin A S M. Gray level co-occurrence matrix (GLCM) and gabor features based no-reference image quality assessment for wood images[J]. Proceedings of International Conference on Artificial Life and Robotics, 2021, 2021: 736–741.
5. Heshalini Rajagopal, Norrima Mokhtar T F T M N, Izam W K W A. No-reference quality assessment for image- based assessment of economically important tropical woods[J/OL]. PLoS ONE, 2020: 1–15.
6. Rajagopal H, Mokhtar N, Khairuddin A S M. A No-Reference Image Quality Assessment Metric for Wood Images[J/OL]. Journal of Robotics, Networking and Artificial Life, 2021, 8(2): 127.
7. Mittal A, Moorthy A K, Bovik A C. No-Reference Image Quality Assessment in the Spatial Domain[EB/OL](2012–12).
8. Mittal A, Soundararajan R, Bovik A C. Making a ‘ Completely Blind ’ Image Quality Analyzer[J]. : 1–4.
9. Wang Z, Simoncelli E P, Bovik A C. Multi-Scale Structural Similarity for Image Quality Assessment[C]//Asilomar Conference on Signals, Systems and Computers. IEEE, 2003: 1398–1402.
10. Zhang L, Zhang D, Mou X. FSIM: a feature similarity index for image quality assessment[J]. Image Processing, IEEE Transactions on, 2011, 20(8): 2378–2386.
11. Xue W, Zhang L, Mou X. Gradient magnitude similarity deviation: A highly efficient perceptual image quality index[J]. Image Processing, IEEE Transactions on, 2014, 23(2): 684–695.

Authors Introduction

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.

Mr. Sayanth Sudheer

He received his Diploma from the Department of Electrical and Electronic Engineering, Manipal International University, Malaysia in 2022. Currently, he is an Intern in Splatz Zone Sdn Bhd, Malaysia.

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. Keoy Kay Hooi

He received his PhD from the Sheffield Plallam University, UK. He is currently an Associate Professor and Director of Institute of Computer Science and Digital Innovation (ICSDI), UCSI University, Malaysia. His research interest areas are Computer Networks and Wireless Communication, Network Mobility in Heterogeneous Network, Cyber Physical Systems Security & Internet of things (IoT)

Dr. Norrima Mokhtar

She received the B.Eng. degree from University of Malaya, the M.Eng. and the Ph.D. degree from Oita Univerity, Japan. She is currently a Senior Lecturer in the Department of Electrical Engineering, University of Malaya. Her research interests are signal processing and human machine interface.

Development of IoT based Key Finder

Sayanth Sudheer
Department of Electrical and Electronic Engineering, Manipal International University, Malaysia

Heshalini Rajagopal
Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia

Azam Mohammed Al-Qussari
Department of Electrical and Electronic Engineering, Manipal International University, Malaysia

Norrima Mokhtar
Department of Electrical Engineering, Faculty of Engineering, University of Malaya, Malaysia

Neesha Jothi, Raenu Kolandaisamy
Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia
E-mail: heshalini@ucsiuniversity.edu.my
www.ucsiuniversity.edu.my

Abstract

Typically, we misplace our keys and hunt for them throughout the home, eventually locating them with much difficulty after a long search. Therefore, we propose to develop a basic IoT-based Key finder utilizing NodeMCU, Buzzer, and Battery in this paper. In this paper, the development of key chain which can be attached to the keys will be explained. The paper also includes the creation of a website dedicated to the search for the missing keys. The missing keys can be located using a mobile phone's Google Chrome browser. When the webpage is enabled to find the missing keys, the designed IoT-based keychain is equipped with a buzzer that will make a beep sound. The developed IoT based key finder is important in terms of keeping track of the keys and could also save time.

Keywords: Key Finder, Internet of Things (IoT), NodeMCU, Google Chrome

1. Introduction

Generally, people misplace their keys and search the entire home for them, only to discover them after a long search. It is very easy to misplace keys, and it can be a stressful situation if they do not have any spares. This is the twenty-first century, and there are tools that can assist anyone in quickly locating missing keys. When they are within a certain distance, some of them make loud noises, while others utilize GPS and Bluetooth to show information on your smartphone. However, adding modules like GPS and Bluetooth to the gadget can make it overly big and heavy[1]. Hence, this study aims to build a simple IoT-based Key finder by using NodeMCU, Buzzer, and Battery. This project will be built in a form of a key chain which can be attached to your keys. The project also aimed to develop a webpage dedicated to find the missing keys. The missing keys can be found using google chrome webpage from the mobile phone. The developed IoT based keychain is equipped with a buzzer which will produce beep sound when the webpage is activated to find the missing keys.

According to a survey of 1,000 U.S. adults conducted by Bluetooth tracking startup Pebblebee.com, one in every five people loses or misplaces personal items every week [2]. Car keys are the most frequently misplaced items. Twenty-eight percent said they look for an item for less than a week before accepting it as lost forever. This shows that everyone misplaces personal items, particularly keys. This is a big issue because no one has time to seek for a missing key in today's world. These misplaced keys might sometimes end up in the hands of the wrong individuals, causing everyone to suffer as a result of their oversight. Therefore, this paper has proposed an IoT-based Key to tackle this challenge. With this idea, the misplaced key will make a buzzer sound to get our attention and find our lost key.

2. Prior Study

Several studies have been conducted on IoT based detection and tracking system. One of the examples is the IoT based fire alarm system. Asma Mahgoub, Abdullah al-Ali, Nourhan Tarrad, (2019) has designed an IoT based fire alarm system. The author has proposed an IoT based wireless fire alarm system that is easy to install. The proposed system used an ESP8266 microcontroller. Each microcontroller is connected to different kind of sensors (temperature, smoke, humidity). The system also consists of a buzzer to alert the user. An Arduino was used to program the microcontroller. The ESP8266 has been programmed to send a SMS to the fire department and the user, also call the user and alert the house by producing an alarm sound. An app was created by author to send message to the user. The proposed system was able to achieve its goals which were mainly building an IoT based fire alarm system that was capable of detecting the presence of fire and call or sent SMS to the necessary parties when a fire is detected[3]. The drawback of this system of the study is that it could have used a rechargeable battery source instead of a power supply.

IoT was also incorporated widely to Human Activity Recognition (HAC) by monitoring the vital signs remotely. Diego Castro, William coral, Camila Rodriguez (2017) has designed a wearable-based human activity recognition using IoT based devices. The proposed system used an ESP8266 microcontroller. Each microcontroller program to detect the data from the sensors (heart rate, respiration rate, skin temperature). An application has been created to record the data that were taken from the sensors. These data can be checked by a medical professional daily. If the patient lives far away from a medical assistance a call or a SMS will be sent to a hospital to monitor the patient any vital sign are showing any problem. The proposed system was able to achieve its objective which was to build a human activity recognition IoT based device to monitor and record their vital signs. Also, to send a SMS or call the hospital if any sign doesn't see normal[4]. The drawback for this system is that it can be more lightweight and instead of putting the device in waist it can be build smaller so it can fit in the arm.

IoT technology was also used in the child safety and tracking device where it is developed to help the parents to locate and monitor their children. M Nandini Priyanka, S Murugan, K N H Srinivas, T D S Sarveswararao, E Kusuma Kumari (2019). The proposed system is developed using a linkit One board and the system is interfaced with temperature, heartbeat, touch sensors and also GPS, GSM & digital camera modules. The system will automatically alert the parents by sending SMS, when immediate attention is required for the child during emergency. The parameters such as touch, temperature &heartbeat of the child are used for parametric analysis and results are plotted for the same. An application was created by the author to record these results obtained. The proposed system was able to achieve its objective which was to build an IoT based device for child safety and tacking[5]. The drawback for this system is that there will be charges for SMS services.

Moreover, IoT was also used to perform automated irrigation by using Android Application. The irrigation robot was equipped with eight wheels to facilitate the movement of the robot[6]. The robot was also equipped with two ultrasonic sensors of any obstacle at the front and base of the robot to detect obstacles in the front and under the robot. The robot lifts up before moving forward when it detected obstacles in the front and comes back down with the assistance of the base ultrasonic sensor when the obstacle has been passed. This also enabled the system to water the tall plants. Soil moisture and DHT22 sensors are fixed on the land to sense the soil moisture, humidity and temperature. The readings of these parameters are sent to the user vis SMS every one minute. In addition, the user can view the readings of the parameters via Android App. The robot moves forward once and perform the seeding at one point, irrigate the seed and moves to the next point for the seeding process. The robot dispenses seeds automatically at a fixed distance (7 cm)[7].

3. Methods

The IoT based key finder system proposed in this paper is built in a form of a key chain which can be attached to the keys. The proposed system also aimed to develop a webpage dedicated to find the missing keys. The missing keys can be found using google chrome webpage from the mobile phone. The developed IoT based keychain is
The block diagram and the flowchart of the proposed system is shown in Figs. 1 and 2, respectively.

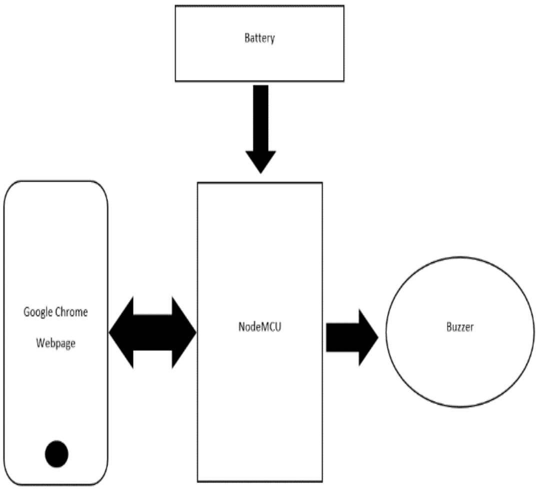

Battery
Google Chrome
Webpage
NodeMCU
Buzzer

Fig.1. Block Diagram of the system

4. Results and Discussion

4.1. *Hardware of IoT based Key Finder*

The circuit of the IoT based key finder assembled on PCB is shown in Fig. 3. The circuit is covered with a plastic cover as shown in Fig. 3. The key finder has been designed to be small and light. The weight of the key finder is only 20g and the size of the key finder is only 9 cm by 6cm. Hence, it is feasible to bring it everywhere.

4.2. *Google Chrome Webpage*

A webpage dedicated to find the missing keys is developed. The missing keys can be found using google chrome webpage from the mobile phone as shown in Fig.4.

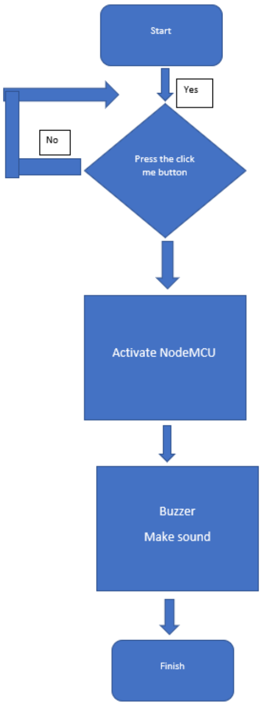

Start
Yes
No
Press the click me button
Activate NodeMCU
Buzzer
Make sound
Finish

Fig.2. Flowchart of the system

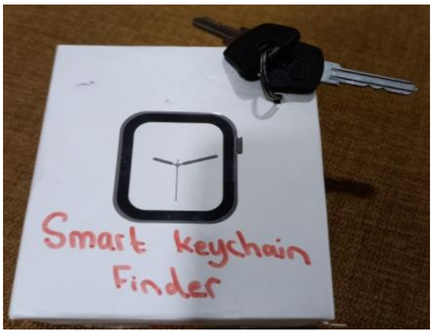
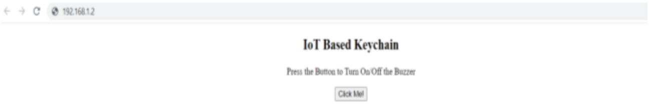

Smart keychain Finder

Fig. 3. Hardware of IoT based key finder


192.168.1.2

IoT Based Keychain

Press the Button to Turn On/Off the Buzzer

Click Me!

Fig. 4. GUI of the Webpage

First, the user has to connect the NodeMCU to user device using device hotspot. In the next stage, enter Wi-Fi credentials like the username and password for the Wi-Fi router to which your NodeMCU should connect with. Once connected the user can type in the IP address (192.168.1.2) of the device to open the webpage shown in Fig. 4. This webpage is created using HTML code. This code is then programmed into the NodeMCU. When the 'click me' icon is clicked in case where the key is missing, the buzzer will be turned on. Once the user has found the key, the user can disable the buzzer by clicking the 'click me' icon again. The system has been tested in various detection distances. The detection range of the key from the mobile is 1m to 15m. The frequency of the sound produced by the buzzer is 6kHz and the user will be able to hear the buzzer sound in 15m range.

5. Conclusion

In this paper, IoT-based Key finder by using ESP8266-01, Buzzer, and Battery has been developed. This system was built in a form of a key chain which can be attached to the keys. The system has developed a webpage dedicated to find the missing keys. The missing keys can be found using google chrome webpage from the mobile phone. The developed IoT based keychain is equipped with a buzzer which will produce beep sound when the webpage is activated to find the missing keys. A number of improvements can be made to the project. A GPS module could have been used in the smart keychain finder. With the help of this keychain finder the user could have easily tracked the missing keys. A message system could have been programmed to alert the user about finding there missing keys. Instead of a webpage the user can use an android app to track the missing keys.

References

1. Ometov A, Shubina V, Klus L. A Survey on Wearable Technology: History, State-of-the-Art and Current Challenges[J/OL]. Computer Networks, 2021, 193: 108074.
2. What Is a Bluetooth Tracker?[EB/OL](2022).
3. Mahgoub A, Tarrad N, Elsherif R. IoT-Based Fire Alarm System[C/OL]//2019 Third World Conference on Smart Trends in Systems Security and Sustainablity (WorldS4). IEEE, 2019: 162–166.
4. Cabra J, Castro D, Colorado J. An IoT Approach for Wireless Sensor Networks Applied to e-Health Environmental Monitoring[C/OL]//2017 IEEE 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). IEEE, 2017: 578–583.
5. Nandini Priyanka M, Murugan S, Srinivas K N H. Smart IOT device for child safety and tracking[J]. International Journal of Innovative Technology and Exploring Engineering, 2019, 8(8): 1791–1795.
6. Kumar A, Rajagopal H. Automated Seeding and Irrigation System using Arduino[J]. Journal of Robotics, Networking and Artificial Life, 2022, 8(4): 259–262.
7. Kumar A, Rajagopal H. Design and development of automated seeding and irrigation system[J/OL]. Proceedings of International Conference on Artificial Life and Robotics, 2021, 26: 742–746.

Authors Introduction

Mr. Sayanth Sudheer

He received his Diploma from the Department of Electrical and Electronic Engineering, Manipal International University, Malaysia in 2022. Currently, he is an Intern in Splatz Zone Sdn Bhd, 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. 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.

Mr. Azam Mohammed Al-Qussari

He received his Diploma from the Department of Electrical and Electronic Engineering, Manipal International University, Malaysia in 2022.

Dr. Norrima Mokhtar

She received the B.Eng. degree from University of Malaya, the M.Eng. and the Ph.D. degree from Oita Univerity, Japan. She is currently a Senior Lecturer in the Department of Electrical Engineering, University of Malaya. Her research interests are signal processing and human machine interface.

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. 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.

Quality assessment for microscopic parasite images

Muhammad Amirul Aiman Asri, Norrima Mokhtar
Department of Electrical Engineering, Faculty of Engineering, University of Malaya, Malaysia

Heshalini Rajagopal
Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia

Wan Amirul Wan Mohd Mahiyiddin
Department of Electrical Engineering, Faculty of Engineering, University of Malaya, Malaysia

Yvonne Ai Lian Lim
Department of Parasitology, Faculty of Medicine, University of Malaya, Lembah Pantai, Kuala Lumpur, Malaysia.

Masahiro Iwahashi
Department of Electrical, Electronics and Information Engineering,Nagaoka University of Technology, Japan

Anees ul Husnain
Department of Electrical Engineering, Faculty of Engineering, University of Malaya, Malaysia

E-mail: norrimamokhtar@um.edu.my, heshalini@ucsiuniversity.edu.my
www.um.edu.my

Abstract

Microscopy lens distortion will cause errors during parasites image acquisition for water sample inspection. Since water sample inspection is crucial for treated water monitoring, the quality of microscopic parasite images such as Giardia and Cryptosporidium need to be monitored as well to avoid errors in treated water inspection. In this work, the subjective and objective evaluation of parasite images were performed. The parasite species studied were Cryptosporidium and Giardia (oo)cysts. Parasite image database consisting of 20 reference images and 360 distorted images were used in the evaluation. The distorted images were generated from the reference images by applying distortion to the reference images with Gaussian White Noise and Motion Blur, at 9 levels of distortions. Twenty subjects were selected to assess the distorted images for the subjective evaluation. The scores obtained from the subjects were transformed into Mean Opinion Score (MOS). In the objective evaluation, six Full Reference-IQA (FR-IQA) metrics, namely MSSIM, SSIM, FSIM, IWSSIM, GMSD and VIF were used to evaluate the distorted images. The subjective MOS scores were used as the benchmark to determine the most suitable objective IQA to assess parasite images. The relationship between the subjective MOS and objective IQAs are examined using performance metrics namely PLCC and RMSE. It was found that MSSIM is the most suitable IQA to assess parasite images distorted with Gaussian White Noise and Motion Blur.

Keywords: Parasite images, Giardia, Cryptosporidium, Full-Reference Image Quality Assessment (FR-IQA).

1. Introduction

Inspection via object recognition has become the topic of interests for many researchers in image processing fields that apply automation to replace human expert. Inspection of treated water sample under microscope requires good quality of images in order to count average number of parasites like Giardia and Cryptosporidium that captured in the sample which determine the safety of the treated water[1]. For robust object recognition performance, good quality of images is required for both manual and automatic inspection. To date, many image quality assessment models have been developed. However, image quality assessment of microscopic images especially for Giardia and Cryptosporidium parasites, has not been reported yet to the best of our knowledge.

This work carries significant contribution as many researchers investigated and reported on distortion correction of a microscopy lens system which causing non-uniform geometric distortion, displacement errors and others[2][3][4][5]. However, the focus of these works and reports was only on the distortion correction. Our work focused on image quality assessment on microscopic parasite images for Giardia and Cryptosporidium, which eventually will open more interests on other microscopic images as well for further studies. Generally, Image Quality Assessment (IQA) has two categories which are subjective and objective evaluations. Subjective evaluation setup is done when the images are evaluated by human, who provides rating based on their visual assessment on the image quality. For objective evaluation setup, the rating calculation for the images is conducted via mathematical algorithms. For justification, the gold standard of IQA is subjective evaluation by human, however, human always has fatigue eyes which eventually lead to errors, as well as cost and time consumption [6]. Hence, objective assessment is the way forward for improvement of treated water inspection process.Recently, IQA objective assessment, has evolved from Full Reference-IQA(FR-IQA) to No-Reference-IQA(NR-IQA) [7][8]. FRIQA is an objective evaluation by comparing the image with its reference image. Meanwhile, Reduced Reference-IQA(RR-IQA) is an objective evaluation using partial information of the reference image. On contrary, NR-IQA is an objective evaluation without information of reference image.

To this study, six Full Reference-IQA (FR-IQA) metrics, namely Structural Similarity Index (SSIM)[9], Multiscale SSIM (MS-SSIM)[9], Feature SIMilarity (FSIM)[10], Visual Information Fidelity (VIF)[11], Information Weighted SSIM (IW-SSIM)[12] and Gradient Magnitude Similarity Deviation (GMSD) [13]. These metrices are chosen as they are widely used to study the image quality of various types of images such as MRI, wood, underwater and natural images [14][15][16][17]. After investigation of these objectives assessment, the performance are measured using human mean opinion scores(MOS) and metrics which in our study, Pearson’s Linear Correlation Coefficient (PLCC) [18] and Root Mean Square Error (RMSE) [19] were used.

2. Methods

2.1. *Parasite Images*

Twenty parasite images from two parasite species, namely Giardia and Cryptosporidium. The images were obtained from Department of Parasitology, University of Malaya, 50603 Kuala Lumpur, Malaysia. The twenty parasite images are shown in Fig. 1. The images were converted to grayscale and the pixel values were normalized to the range 0 - 255 for ease of applying the same levels of distortion across all the reference images. The images consisted of a matrix of 1376 x 1320 pixels. These twenty reference parasite images were the distorted by Gaussian white noise and motion blur, which represent image distortions typically encountered in parasite images. Gaussian white noise often arises in during acquisition of parasite images by the capturing apparatus [20]. On the other hand, parasite images are subjected to motion blur when there is a relative motion between the capturing apparatus namely microscope lens and the parasite specimen [21]. These distortions cause the quality of the parasite image to be low [21]. Thus, the features of the parasite could not be differentiated. This could cause misclassification of the parasite species as the feature extractor will not be able to extract distinctive features from the parasite texture images effectively [21]. The Gaussian white noise with standard deviation, σ_{GN} and motion blur with standard deviation, σ_{MB} were applied to the reference images at nine levels of distortion of the reference images, i.e.: σ_{GN} = 10, 20, 30, 40, 50, 60, 70, 80 and 90 for Gaussian white noise and σ_{MB} = 10, 20, 30, 40, 50, 60, 70, 80 and 90 for motion blur. This produces 380 parasite images, twenty reference images, 180 images distorted by Gaussian White Noise and 180 images distorted by Motion Blur.

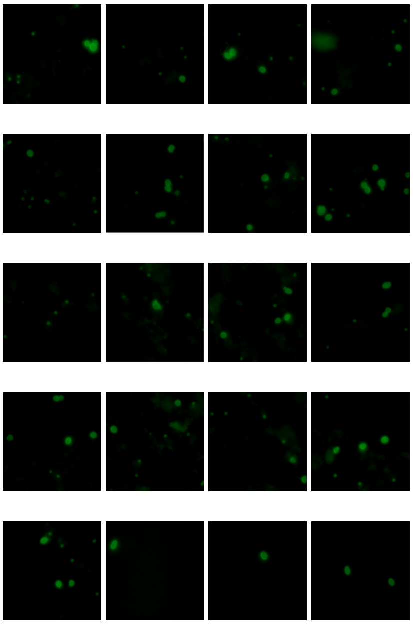


Fig. 1. Twenty reference Giardia and Cryptosporidium parasite images

2.2. *Subjective Evaluation*

Twenty students aged 20-25 years from the Department of Electrical and Electronics Engineering from Manipal International University (MIU) and University of Malaya, Malaysia volunteered to evaluate the parasite images. The evaluation was performed based on the procedures recommended in Rec. ITU-R BT.500-11 [22] in an office environment using a 21 inch LED monitor with resolution of 1920 x 1080 pixels. Uncorrected near vision acuity of every subject was checked using the Snellen Chart before the subjective evaluation in order to confirm their fitness to perform the evaluation task. The subjective evaluation was performed based on the Simultaneous Double Stimulus for Continuous Evaluation (SDSCE) methodology [15,22]. Here, the reference and distorted images are displayed on the monitor screen side-by-side, where the reference image is displayed on the left and the distorted image is displayed on the right. Each subject evaluates the distorted image by comparing the quality of the images (right side) with its reference image (left side). The subject rates either Excellent (5), Good (4), Fair (3), Poor (2) or Bad (1) for each image displayed. The numerical scores were not revealed to the subjects as it could cause bias between the subjects [23] . The evaluation process takes 15 to 20 minutes. The ratings obtained from the subjects were used to calculate MOS [23].

2.3. *Objective Evaluation*

The MOS is compared with six FR-IQA metrics: Structural Similarity Index (SSIM) [9], Multiscale SSIM (MS-SSIM) [9], Feature SIMilarity (FSIM) [10], Visual Information Fidelity (VIF) [11], Information Weighted SSIM (IW-SSIM) [12] and Gradient Magnitude Similarity Deviation (GMSD) [13]. The FR-IQA metrics are descripted in Table 1.

Table 1. FR-IQA Metrics used in this study

IQA Algorithm	Description
Structural Similarity Index Metrics (SSIM)	Captures the loss in the structure of the image.
Multiscale SSIM (MS-SSIM)	Mean of SSIM that evaluates overall image quality by using a single overall quality.
Feature SIMilarity (FSIM)	A low-level feature-based image quality assessment which used two types of features: Phase Congruency (PC) and Gradient Magnitude (GM).
Visual Information Fidelity (VIF)	Measures image information by computing two mutual information quantities from the reference and distorted images.
Information Weighted SSIM (IW-SSIM)	Obtained by combining content weighting with MS-SSIM.
Gradient Magnitude Similarity Deviation (GMSD)	Computes the pixel-wise similarity between the gradient magnitude maps of the reference and distorted images to create a Local Quality Map (LQM) of the distorted image.

2.4. *Performance Evaluation*

Two well established performance metrics were used to study the correlation between the objective IQA and subjective MOS values. The first performance metric chosen is PLCC. The PLCC values between the FR-IQA metrics and MOS were calculated after the nonlinear regression. The second performance metric is RMSE. RMSE is a statistical metric commonly used to evaluate a model’s performance [24]. Similar to PLCC, the RMSE values were calculated after the nonlinear regression.

3. Results and Discussion

The relationship between MOS and ten distortion levels of Gaussian white noise and Gaussian blur were shown in Fig. 2 (a) and (b). Generally, as the distortion levels increase, the quality of the image will be poorer. As the quality of the images gets poorer, the MOS value decreases.

Based on the scatter plot in Fig. 2 (a), the MOS scores decreased at the range of 4.3 to 1.6 as the gaussian white noise’s distortion levels increased from 10 to 90. This shows that the human subjects could differentiate images with different distortion levels. In Fig. 2 (b), the MOS scores did not deviate much even when the level of Gaussian blur increased from 10 to 90. This shows that there are no significant changes in the quality of parasite images when the images were blurred.

The calculated PLCC and RMSE values between MOS, FR-IQA metrics (FSIM, IWSSIM, MSSIM, GMSD, SSIM and VIF) and NR-IQA metrics (BRISQUE, NIQE and PIQE) are shown in Table 2. PLCC values close to 1 indicate that the MOS correlates well with the IQA metric, whereas lower RMSE values indicate that the MOS correlates with the IQA metric. Table 2 shows that the highest PLCC and lowest RMSE values for Gaussian white noise and the overall database were obtained for the MSSIM. The highest PLCC and lowest RMSE were obtained for BRISQUE for Gaussian Blur.

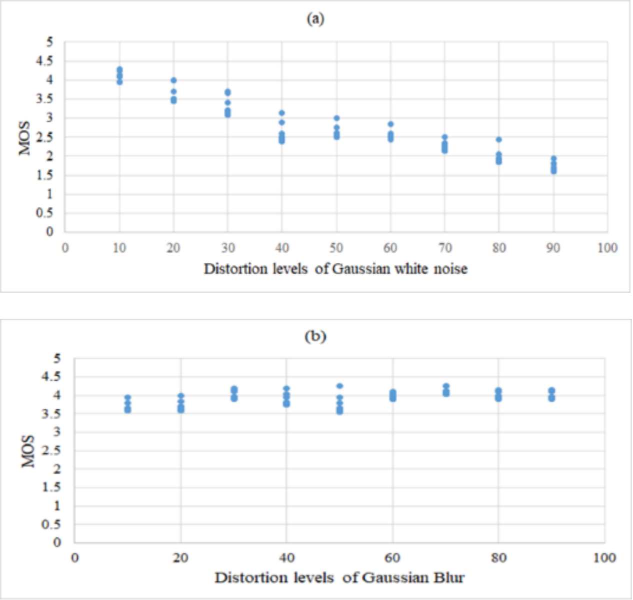


Fig. 2. Scatter plot of MOS versus (a) Gaussian white noise, (b) Gaussian blur.

Table 2. PLCC and RMSE values between MOS, FR-IQA and NR-IQA.

		FSIM	IWSSIM	MSSIM	GMSD	SSIM	VIF
PLCC	**GWN**	**0.939**	**0.966**	**0.975**	**0.755**	**0.957**	**0.881**
	MB	**0.294**	**0.060**	**0.047**	**0.056**	**0.048**	**0.104**
	All	**0.946**	**0.960**	**0.961**	**0.866**	**0.955**	**0.936**
RMSE	**GWN**	**0.256**	**0.191**	**0.165**	**0.488**	**0.216**	**0.352**
	MB	**0.170**	**0.177**	**0.177**	**0.177**	**0.177**	**0.176**
	All	**0.263**	**0.228**	**0.224**	**0.407**	**0.240**	**0.287**

4. Conclusions

A database which consists of 380 parasite images (20 reference and 360 distorted images) was generated. The reference images were distorted with Gaussian White Noise and Motion Blur which commonly occur during the acquisition of parasite images. The database also contains the subjective MOS, six types of objective FR-IQAs and three NR-IQAs' evaluation.

The relationship between the subjective MOS and objective IQAs are examined using performance metrics namely PLCC and RMSE. Both performance metrics showed that MSSIM is the most suitable IQA to assess parasite images. This is because MSSIM is a more direct way to compare the structures of the reference and the distorted parasite images as it measures the structural information change between the reference and distorted images. This study has also shown that IQA is important in assessing parasite images as feedback method prior to inspection procedure.

References

1. Mittal A, Moorthy A K, Bovik A C. No-Reference Image Quality Assessment in the Spatial Domain[EB/OL](2012–12).
2. Yoneyama S. Lens distortion correction for digital image correlation by measuring rigid body displacement[J/OL]. Optical Engineering, 2006, 45(2): 023602.
3. Jin P, Li X. Correction of image drift and distortion in a scanning electron microscopy[J/OL]. Journal of Microscopy, 2015, 260(3): 268–280.
4. Rajagopal H, Mokhtar N, Khairuddin A S M. A No-Reference Image Quality Assessment Metric for Wood Images[J/OL]. Journal of Robotics, Networking and Artificial Life, 2021, 8(2): 127.
5. Maraghechi S, Hoefnagels J P M, Peerlings R H J. Correction of Scanning Electron Microscope Imaging Artifacts in a Novel Digital Image Correlation Framework[J/OL]. Experimental Mechanics, 2019, 59(4): 489–516.
6. Wang Z . Applications of objective image quality assessment methods[J]. Signal Processing Magazine, IEEE, 2011, 28(6): 137–142.
7. Gulame M, Joshi K R, Kamthe R S. A Full Reference Based Objective Image Quality Assessment[J]. International Journal of Advanced Electrical and Electronics Engineering (IJAEEE), 2013, 2(6): 13–18.
8. Chandler D M. Seven Challenges in Image Quality Assessment: Past, Present, and Future Research[J/OL]. ISRN Signal Processing, 2013, 2013: 1–53.
9. Wang Z, Simoncelli E P, Bovik A C. Multi-Scale Structural Similarity for Image Quality Assessment[C]//Asilomar Conference on Signals, Systems and Computers. IEEE, 2003: 1398–1402.
10. Zhang L, Zhang D, Mou X. FSIM: a feature similarity index for image quality assessment[J]. Image Processing, IEEE Transactions on, 2011, 20(8): 2378–2386.
11. Sheikh H R, Bovik A C. Image information and visual quality.[J]. IEEE transactions on image

processing : a publication of the IEEE Signal Processing Society, 2006, 15(2): 430–444.

12. Wang Z, Li Q. Information content weighting for perceptual image quality assessment[J]. IEEE Transactions on Image Processing, 2011, 20(5): 1185–1198.
13. Xue W, Zhang L, Mou X. Gradient magnitude similarity deviation: A highly efficient perceptual image quality index[J]. Image Processing, IEEE Transactions on, 2014, 23(2): 684–695.
14. Heshalini Rajagopal, Norrima Mokhtar T F T M N, Izam W K W A. No-reference quality assessment for image- based assessment of economically important tropical woods[J/OL]. PLoS ONE, 2020: 1–15.
15. Chow L S, Rajagopal H, Paramesran R. Correlation between subjective and objective assessment of magnetic resonance (MR) images[J/OL]. Magnetic Resonance Imaging, 2016, 34(6): 820–831.
16. Shi S, Zhang X, Wang S. Study on Subjective Quality Assessment of Screen Content Images[C]//Picture Coding Symposium (PCS).
17. Hou G, Zhao X, Pan Z. Benchmarking Underwater Image Enhancement and Restoration, and beyond[J]. IEEE Access, 2020, 8: 122078–122091.
18. Song X-K. Correlated data analysis: modeling, analytics, and applications[M]. Springer Science & Business Media, 2007.
19. Chai T, Draxler R R. Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature[J]. Geoscientific Model Development, 2014, 7(3): 1247–1250.
20. Haider S A, Cameron A, Siva P. Fluorescence microscopy image noise reduction using a stochastically-connected random field model[J/OL]. Scientific Reports, 2016, 6(1): 20640.
21. Cheng Jiang, Jun Liao, Pei Dong, Zhaoxuan Ma, De Cai, Guoan Zheng, Yueping Liu, Hong Bu J Y. Blind deblurring for microscopic pathology images using deep learning networks[J]. arXiv preprint arXiv:2011.11879, 2020.
22. Recommendation I T U R B T (2002). 500-11. Methodology for the Subjective Assessment of the Quality of Television Pictures[R]//Recommendation ITU-R BT. 500-11. ITU Telecom. Standardization Sector of ITU.
23. Bindu K, Ganpati A, Sharma A K. A Comparative Study of Image Compression Algorithms[J]. International Journal of Research in Computer Science, 2012, 2(5): 37–42.
24. Chow L S, Rajagopal H, Paramesran R. Correlation between subjective and objective assessment of magnetic resonance (MR) images[J]. Magnetic Resonance Imaging, 2016, 34(6): 820–831.

Authors Introduction

Mr. Muhammad Amirul Aiman Bin Asri

He received the B.E (Electrical) from the Department of Electrical Engineering, University of Malaya, Malaysia in 2020 Currently, he is a Research Assistant in University of Malaya, Malaysia. His research interest includes image processing, artificial intelligence and machine learning.

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. 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. Wan Amirul Bin Wan Mohd Mahiyidin

He received the M.Eng. degree from the Imperial College London in 2009, the M.Sc. degree from University of Malaya in 2012, and the Ph.D. degree from University of Canterbury in 2016. He is currently a senior lecturer at the Department of Electrical Engineering, University of Malaya. His research interests are multiple antennas system, cooperative MIMO, channel modelling and positioning system.

Dr. Yvonne Lim Ai Lian

Dr. Yvonne Lim Ai Lian is a professor at the Department of Parasitology, Faculty of Medicine, Universiti Malaya (UM), Kuala Lumpur and a council member of the Academy of Sciences Malaysia. She has taken on leadership roles as the Deputy Dean (Research) at the Faculty of Medicine and the Director of International Relations Office (IRO), Universiti Malaya. She has also served on the University Senate and was the former president of the Malaysian Society of Parasitology and Tropical Medicine (MSPTM).

Her research focuses on host-parasite interactions and the epidemiology and control of neglected tropical diseases primarily among the indigenous communities. Universiti Malaya. She has also served on the University Senate and was the former president of the Malaysian Society of Parasitology and Tropical Medicine (MSPTM). Her research focuses on host-parasite interactions and the epidemiology and control of neglected tropical diseases primarily among the indigenous communities.

Dr. Masahiro Iwahashi

Masahiro Iwahashi (Senior Member, IEEE) received the B.Eng., M.Eng., and D.Eng. degrees in electrical engineering from Tokyo Metropolitan University, Tokyo, Japan, in 1988, 1990, and 1996, respectively., In 1990, he joined Nippon Steel Company Ltd. Since 1993, he has been with the Nagaoka University of Technology, Nagaoka, Japan, where he is currently a Professor with the Department of Electrical, Electronics and Information Engineering. His research interests include the areas of digital signal processing, multirate systems, and image compression, Dr. Iwahashi is a Senior Member of the IEICE and a member of the Asia Pacific Signal and Information Processing Association (APSIPA) and the Institute of Image Information and Television Engineers (ITE).

Anees ul Husnain

Anees is currently pursuing his Doctoral in Electrical Engineering at Universiti 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.

A Study on the Impact of Hardware Limitations in Multi-Rotor UAVs on Coverage Path Planning Models

Anees ul Husnain, Norrima Binti Mokhtar, Noraisyah Binti Mohamed Shah, Mahidzal Bin Dahari, Muhammad Syazni Ikmal Bin Ramlee
Applied Control and Robotics Laboratory, Level 4 Blok M, Faculty of Engineering, University of Malaya,
Kuala Lumpur, 50603, Kuala Lumpur

Heshalini Rajagopal
Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia

Masahiro Iwahashi
Information, Telecommunication and Control System Group, Nagaoka University of Technology, Niigata, Japan

E-mail: anees.husnain@iub.edu.pk, norrimamokhtar@um.edu.my, noraisyah@um.edu.my, mahidzal@um.edu.my,
ikmalramlee336@gmail.com, heshalini@ucsiuniversity.edu.my, iwahashi@nagaokaut.ac.jp,
www.um.edu.my

Abstract

Search and explore missions through patrolling UAVs need effective strategies for area coverage. Various methodologies for coverage path planning were explored and analyzed through the ROS-Gazebo simulation environment using the Hector quadrotor model. Considering the impact of hardware limitations, simulations were conducted, for such missions where the UAV is needed to switch frequently between search and localize modes. This study investigated raster-scan exploration, expanding spiral search and zigzag pattern coverage to analyze the impact of limitations from Hector UAV on these models. The evaluation parameters were percentage of covered area, number of turns and time taken by the UAV.

Keywords: Area Coverage, Path Planning, Coverage Path, UAVs, ROS, Gazebo.

1. Introduction

The problem of Coverage Path Planning (CPP) means to find a route that passes through all the points of interest in a certain region. Most of the time, CPP pris considered in isolation, meaning, the nature of vehicle or robot is not considered as an influential factor that may influence the ideal path. Several others considered ground vehicles as primary mobile units while Unmanned Aerial Vehicles (UAVs) were considered as an extension [1]. However, there is a crucial perspective to this problem: the design, structure, and certain components of a UAV's subsystems, influence their maneuverability and moveability. Consequently, the coverage paths generated from the coverage model do not match the actual path, at times.

1.1. *Problem Statement*

Since, this work targets the coverage area problem for a UAV in search and explore emissions, the same might also need to localize a sub-region as well. Subsequently, a UAV might have to switch between search and localize modes of operation, quite frequently. Therefore, the impact of slight differences in the ideal path generated from the CPP model and the actual path executed by the UAV, can be substantial. It was felt needed to investigate the reasons and study the impact of hardware limitations on the coverage path. To study the impact, a specific UAV with a realistic simulation model was needed that considers its subsystems limitations while following a path generated from an algorithm. 'Hector quadcopter UAV' is one of the most suitable choices to meet this requirement. Hector quadrotor UAV simulation project is developed for Robot Operating System (ROS)-Gazebo simulation framework. One of the greatest advantages selecting ROS is that the codes written in ROS run 'as it were' when needed to be transferred on hardware. Moreover, Hector emulates those quadcopters that well-suite opensource hardware and software technologies for academic research purposes. The project is still active and available under creative common licensing, permissible to modify and use [2, 3].

The rest of this article is organized as, Section 2 briefly presents background of CPP problem, Section 3 describes the simulation setup, Section 4 introduces CPP models under consideration, Section 5 explains the results with brief analysis while Section 6 concludes this article.

2. Formulating a CPP problem

There can be more than one useful solution for any CPP problem that considers factors like shape of coverage area, obstacles in the area and percentage of area needed to be covered. However, the limitations in maneuverability and mobility from its sensors and structures have been overlooked. Let us first consider how a coverage problem is formulated for a given area.

2.1. *The shape of Coverage Area*

The primarily factor is the shape of the coverage area which can either be in the form of regular structures (like, square, rectangle, triangle) or irregular and complex. The implications of the shape-complexity are beyond the scope of this work, since the prime focus is to study the impact of limitations of a specific on the path generated by a CPP model. Therefore, a 10x10 meter square-shaped area is considered for this work.

2.2. *Cell Definition and Area Decomposition*

Once the shape of a coverage area is identified, one of the optional next steps is to decompose the target area into unit cells, however, some algorithms might not need it. These cells define the resolution of the overall map, drastically affect computation overhead, and precision of information gathered by the exploring vehicle. For a smaller cell size, the vehicle would require to just go through once in a cell needing more flight time to cover the area. However, if the cell size is bigger, multiple exploration trips may be needed within a cell as well [1]. Fundamentally, a cell size depends on the UAV's ability (sensor subsystems) to explore a unit of region at a time.

2.3. *Availability of Information for Coverage Area*

If some prior information related to the coverage is available, this can improve the overall performance of CPP, *e.g.*, a piece of information related to the target location in a search mission can avoid unnecessary exploration. Moreover, prior information can also suggest more sophisticated patterns that may prioritize a certain sub-area. It is suggested that an attempt to acquire any prior information must be made as part of the planning process.

2.4. *Area of Interest*

Next thing in line comes the technique for cell decomposition and the primary consideration is most probable or favorite region within the coverage area. As an instance, closer to the boundary of the coverage area may be of a lesser significance as compared to central ones, depending on the objective to attain. A technique to place 'dot-points' within the coverage can address this concern. The concentration of dot-points can either be equally distributed or uneven. The sub-areas with higher interests may have a higher dot-point density.

2.5. *Performance Metrics*

Performance metrics associate a measurable outcome when an activity is executed to meet an objective. There are a few aspects that need to be considered in performance metrics. The total travelled distance or route length [4, 5], the time to accomplish a mission [6], the area coverage maximization [7], and the number of turning maneuvers [8]. However, as the total area increases, so the total distance travelled by UAV increases and vice versa. Furthermore, the area of interest may also affect the time taken by UAV to complete the coverage mission. Thus, area coverage maximization seems difficult if the area of interest is too huge due to limitations of drone technology. This brings the metrics to the basic form of measuring the percentage of area covered in a unit of time, considering flight duration requirements.

2.6. *Pattern Identification / Selection*

The final stage of CPP here is the pattern identification and selection that covers the dot-points (area of interest). This depends upon the nature of distribution of the dot-points. In most of the studies, the enhancements are fundamentally attempted through improved dot-points distribution (area of interests), novelty in constrained-optimization (minimizing time or maximizing coverage), and so on. For this study, three useful CPP patterns are handpicked since the objective here is to study the impact of hardware limitations.

3. Simulation Setup

3.1. *ROS (Melodic)-Gazebo Framework*

ROS Melodic Morenia is originally designed for the Ubuntu 18.04 (Bionic) version, although it also works on other Linux distributions, as well as Mac OS X, Android, and Windows [9]. Gazebo is the most accurate and efficient simulator for simulating robot populations in complicated indoor and outdoor situations with a powerful physics engine, high-resolution visuals [10].

3.1.1. HECTOR Quadrotor

Hector is also an opensource ROS project that contains programs for quadrotor UAV modeling, control, and simulation developed by Meyer *et al.* [11]. These come in the forms of packages that develop its systems design, operation, simulation, and a Unified Robot Description Format (URDF) model as well as versions that include different sensors.

3.1.2. UAV's Field of View (FOV)

Hector's Field of View (FOV) of the camera is the region that is covered when the UAV flies at a certain height h. camera's dimension can be obtained as the equation below and illustrated in Fig. 1:

$$W = 2h \times \tan\left(\frac{\alpha}{2}\right) \quad (1)$$

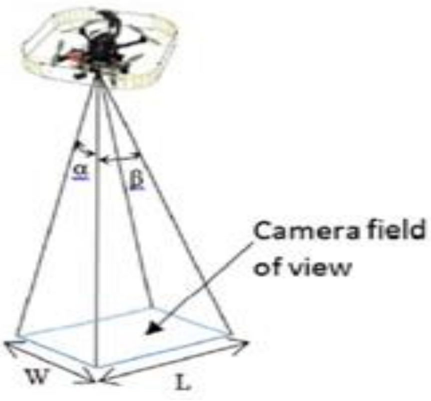


Fig. 1. Field of View (FOV)

$$L = 2h \times \tan\left(\frac{\beta}{2}\right) \quad (2)$$

Where:
W = width of the FOV
L = length of FOV
h = height of altitude
α = camera vertical degree
β = camera horizontal degree

For this project, the UAV flies at a height of 1.2m, the value of α given is 45 degrees and β is also set to 45 degrees.

3.1.3. Hokuyo Laser Range Finder

Hokuyo laser range finder is used to detect any boundary or obstacles that is available near the UAV. Hokuyo Laser consists of a 1081 set array of beam light that covers the left side, front side, and right side of the UAV. The longest distance for the Hokuyo Laser Range Finder for the drone to detect any obstacle or boundary is set to 2 meters. A visualization of range finder is presented in Fig. 2 alongside parameters used in algorithms.

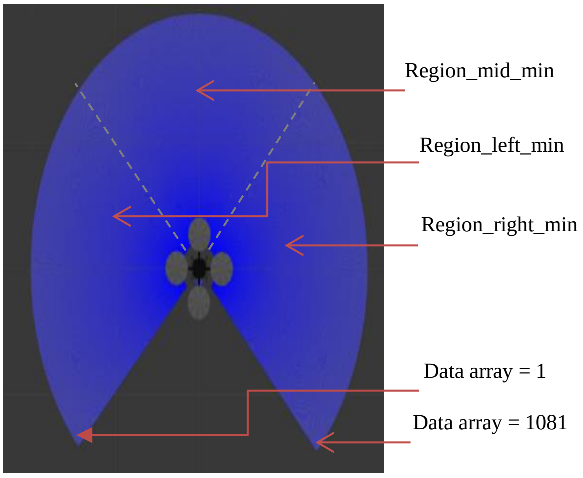


Fig. 2. Data variables through Hokuyo Laser Range Finder

4. CPP Models

4.1. *Workspace Setup*

The workplace is divided into one-by-one meter square with no gaps among them and is called a unit cell. Each part of cell is put with one point at the center to mark the region of interest, in this study. Moreover, if the UAV reaches any point in a cell, it is anticipated as that cell has been covered. Camera field of view calculates the size of the cell in the mosaicking process. Hence, a 2D Grid graph is formed with cells having and a dot has been marked as the center point of each cell as shown in Fig. 3.

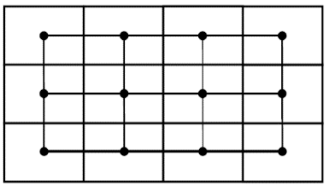


Fig. 3. Cellular decompositions of proposed workspace with equidistance dot-points

4.1.1. Patterns Generation

One of the advantages of UAVs like quadcopters is that through the varying combinations of their motors, using motor mixing algorithms, a substantial number of patterns can be generated by programming simple back-and-forth. Each cell is connected to eight neighboring cells. The connectivity of grid-cells, also known as Moore neighborhood, has been found effective for coverage area planning through UAVs. Through this capability, to rotate to any yaw angle value, which are $\pm 45, \pm 90$ and ± 135, series of such combinations can generate numerous patterns by switching the velocity of each motor, illustrated in Fig. 4.

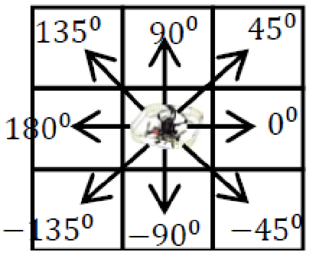

135^{0}
90^{0}
45^{0}
180^{0}
0^{0}
-135^{0}
-90^{0}
-45^{0}

Fig. 4. Moore Neighborhood of quadcopter

4.1.2. Raster-Scan Exploration

Considering this mechanism of pattern generation, the raster-scan search pattern was considered as one of the key techniques that is widely used for exploration missions. The algorithm is presented through a flowchart in Fig. 5, while the simulation results and brief analysis are presented in Section 5.2 and 5.2.1, respectively.

4.1.3. Expanding Spiral Search

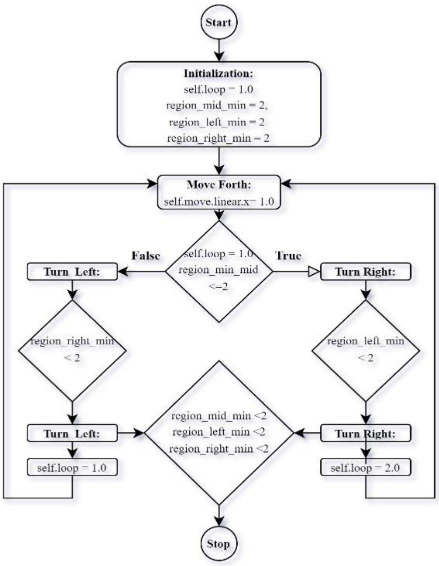

Start
Initialization:
self.loop = 1.0
region_mid_min = 2,
region_left_min = 2
region_right_min = 2
Move Forth:
self.move.linear.x= 1.0
self.loop = 1.0
region_min_mid
<=2
False
True
Turn Left:
Turn Right:
region_right_min
< 2
region_left_min
< 2
region_mid_min <2
region_left_min <2
region_right_min <2
Turn Left:
self.loop = 1.0
Turn Right:
self.loop = 2.0
Stop

Fig. 5. Flowchart depicting Raster-Scan Exploration

In the spiral pattern, UAV starts the coverage mission by flying at the center point of an area. In this stage, the distance travelled by the UAV will be counted and increased for each iteration. One iteration is considered when the UAV finishes the straight moving and starts to turn. For the next iterations, the distance that needed to be travelled by the UAV will be increased to make the UAV fly longer than the path travelled before. Fig. 6 is presented to illustrate the expanding spiral algorithm while Section 5.3 and 5.3.1 presents simulation and brief analysis.

4.1.4. Zig-Zag Coverage

The zigzag pattern is similar as the back-and-forth pattern except the difference is this that the current trajectory is not parallel to the previous trajectory. The difference between two trajectory paths is set to 60 degrees which is 150 degrees to -150 degrees. The flowchart for the zigzag algorithm is presented in Fig. 7. Section 5.4 and 5.4.1

visualizes simulation and presents analysis, respectively.

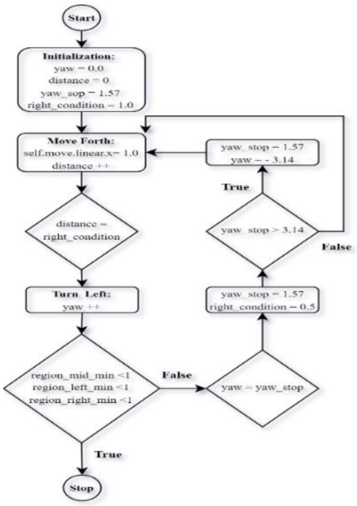

Start
Initialization:
yaw = 0.0
distance = 0
yaw_sop = 1.57
right_condition = 1.0
Move Forth:
self.move.linear.x= 1.0
distance ++
distance =
right_condition
Turn Left:
yaw ++
region_mid_min <1
region_left_min <1
region_right_min <1
True
Stop
False
yaw = yaw_stop
yaw_stop = 1.57
right_condition = 0.5
yaw_stop > 3.14
False
True
yaw_stop = 1.57
yaw = - 3.14

Fig. 6. Flowchart presenting Expanding Spiral Exploration

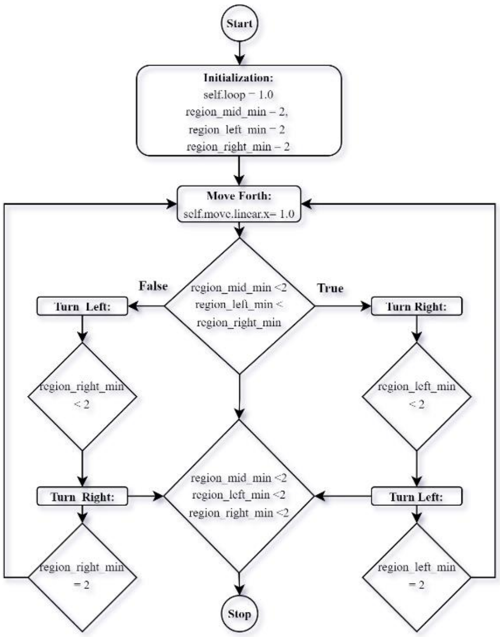

Start
Initialization:
self.loop = 1.0
region_mid_min = 2,
region_left_min = 2
region_right_min = 2
Move Forth:
self.move.linear.x= 1.0
False
region_mid_min <2
region_left_min <
region_right_min
True
Turn Left:
Turn Right:
region_right_min
< 2
region_left_min
< 2
region_mid_min <2
region_left_min <2
region_right_min <2
Turn Right:
Turn Left:
region_right_min
= 2
region_left_min
= 2
Stop

Fig. 7. Flowchart for zigzag coverage

5. Results and Discussions

5.1. *Coverage Area Vs Exploration Time*

In the RVIZ, the workspace has been built with numbers of cells and each cell has an area of one-by-one m^2. The total area of the workspace is $100m^2$. For coverage maximization purposes, one cell is considered entirely covered when the trajectory path flies through the cell. This consideration is based on the field of view of the UAV that flies at a height of 1.2m as stated in Section 3.1.2. Coverage percentage is based on the count of explored cell under the trajectory line:

$$\text{Coverage} = \frac{\text{cells visited per trajectory}}{\text{total number of cells}} \text{ x } 100\% \qquad (3)$$

Time taken to complete coverage path planning for each pattern is taken directly from the ROS software, ROS elapsed, which can be reset when needed and considers the actual translation and rotation capacity of Hector to record time. The position of Hector UAV goes through translation and rotational instruction. As an instance, the following instructions generate a motor driving combination, for linear and rotational motions: *self.robot_velocity.linear.x* *and* *self.robot_velocity.angular.z*. The simulation results for each coverage path planning pattern will be compared to determine the most optimum coverage path candidate among the three. To maneuver the drone the positive or negative values are assigned to these commands, which can make the UAV to either move back-and-forth or turn left-and-right.

5.2. *Raster-Scan Exploration*

Raster-scan pattern was generated through a sequence of back-and-forth motion instructions whose simulation visualization is presented in Fig. 8. The area covered through raster-scan technique for this scenario is roughly 90 percent. Raster-scan was required to make sixteen turns in total for the coverage. The time taken for the drone to complete the coverage in this scenario is 157.48 seconds.

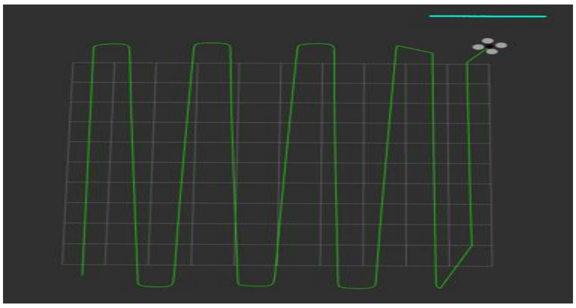


Fig. 8. Visualization of Raster-Scan Exploration of Hector quadcopter where red dots depict unexplored regions

5.2.1. *Brief analysis*

The impact of limitations from the UAV is also visible as every trajectory path is not parallel from the previous path. This is because the drone rotates according to the data fed from the Hokuyo laser range finder. The drone will continuously rotate if the drone gets the reading from Hokuyo Laser. If the data is no longer fed to the drone, the drone will start to move straight.

5.3. *Expanding Spiral Search*

Following the algorithm presented previously in Fig. 6, the result of the expanding spiral search is shown in Fig. 9. It took the spiral pattern sixteen turns in total for the coverage to complete the path coverage. As for the time, it took 130.68 seconds.

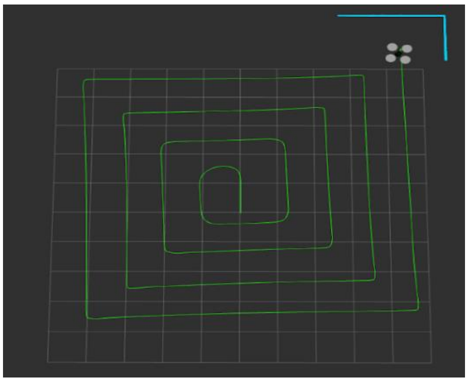


Fig. 9. Visualization of Expanding Spiral Search by Hector quadcopter, red dots representing unexplored regions

5.3.1. *Brief analysis*

First, the UAV movement is not exactly a squared-spiral pattern even if the boundary used is exactly square. This is because of limits in precision for exactly executing a 90-degree or 1.5708 in radian sharp turn. This limitation is occurring because of the data with floating numbers obtained from the yaw value whose increment is not linear in nature. This is almost preventable because an acceleration is needed by the aerial vehicle to execute any translation motion from halt. For instance, when the break condition for the drone to stop rotate at yaw > 1.5708. the data of yaw value obtained may exceed to about 1.65 rad. This makes the difference between the yaw obtained and the yaw desired is 0.08 rad or 4.5 degrees which is enough to create an impact on trajectory that includes multiple turns.

5.4. *Zigzag Coverage*

In the zigzag pattern, the straight movement is equated with the other patterns which is the self.move.linear.x = 1.0 to balance the velocity of the drone along the simulation. A slight difference was made to the command used for the rotation movement to generate a wider angle between paths. The result of the zigzag pattern is as portrayed as Fig. 10. The percentage of coverage maximization performed by the Zigzag Pattern is 73% and made twelve turns and took the least time, 121.66 seconds.

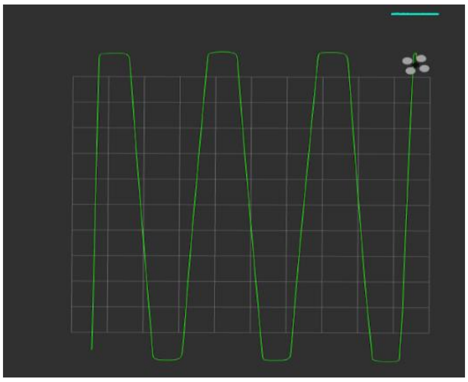


Fig. 10. Visualization of zigzag pattern search by Hector quadcopter, red dots are unexplored cells

5.4.1. *Brief analysis*

It can be observed that the vertices that were generated by the drone are not sharp. The reason for this behavior is due to the fact that the linear.x is not set to 0.0 during performing rotation maneuvers. The value of linear.x = 0.0 has been assessed during testing the algorithm and resulting in the drone hitting the boundary and crashing. This means, the Hector would need to hover till linear.x is zero before taking a rotation and that is a substantial waste of flight time with twelve turns.

5.5. *Comparative Analysis*

A trade-off has been observed, Table 1 among zigzag coverage and expanding spiral search, to maximize coverage area on the cost of a higher number of turns that possibly affect the coverage pattern due to the limiting maneuverability of Hector quadrotor UAV.

Table 1. Summary of results & Comparative Statements

Coverage Scheme / Algorithm	Time taken (s)	Coverage %age	Number of turns
Raster-Scan Coverage	157.48	90%	16
Expanding Spiral	130.68	90%	16
Zigzag Pattern	121.66	73%	12
Comparative Statements			
Raster-Scan Coverage	One of the most frequently technique of exploration took highest time with 16 turns and covered 90%. area.		
Expanding Spiral	Expanding spiral took lesser time than Raster-Scan with 16 turns and 90% area coverage.		
Zigzag Pattern	Zigzag took least number of turns, 12, and least time, however, 73% area was covered.		

6. Conclusion

It is portrayed that for each coverage path planning method, it is impossible for the path generated by the drone to be exactly the same as the model provided. This is mainly because of limitations from the Robot or Aerial Vehicle structure, sensors, physics of motion and data precision. These factors directly impact the UAV’s behavior in terms of maneuverability and mobility which further affects the performance in coverage path planning applications. Consequently, a higher number of turns may lead to a higher deviation from the theoretical model, alongside the desire to reduce exploration time.

This study handpicked three different strategies in a simple exploration scenario with Hector quadcopter in ROS, to prove how critical and essential it is to consider hardware limitations while developing CPP models. Though the expanding spiral exploration took less time to reach maximum coverage area, 90% and with sixteen turns (same as raster-scan search), a trade-off was observed. The zigzag coverage took the least time and least number of turns with 73% of coverage area. This means, if it is needed to minimize the impact of hardware limitations, the zigzag pattern would suit better, alongside, offering a higher flight duration.

Acknowledgment

The study is supported by “IIRG003(b)-19IISS”, University of Malaya, Malaysia.

References

1. T. M. Cabreira, L. B. Brisolara, and R. F. Paulo, "Survey on Coverage Path Planning with Unmanned Aerial Vehicles," (in English), Drones-Basel, Article vol. 3, no. 1, p. 38, Mar 2019, Art no. 4, doi: 10.3390/drones3010004.
2. J. Meyer. "Hector Quadrotor License." https://github.com/tu-darmstadt-ros-pkg/hector_quadrotor/blob/kinetic-devel/LICENSE.txt (accessed Dec, 2022).
3. G. Öst, "Search path generation with UAV applications using approximate convex decomposition," ed, 2012.
4. A. Xu, C. Viriyasuthee, and I. Rekleitis, "Optimal complete terrain coverage using an unmanned aerial vehicle," in 2011 IEEE International conference on robotics and automation, 2011: IEEE, pp. 2513-2519.
5. O. Artemenko, O. J. Dominic, O. Andryeyev, and A. Mitschele-Thiel, "Energy-aware trajectory planning for the localization of mobile devices using an unmanned aerial vehicle," in 2016 25th international conference on computer communication and networks (ICCCN), 2016: IEEE, pp. 1-9.
6. D. Li, X. Wang, and T. Sun, "Energy‐optimal coverage path planning on topographic map for environment survey with unmanned aerial vehicles," Electronics Letters, vol. 52, no. 9, pp. 699-701, 2016.
7. M. Torres, D. A. Pelta, J. L. Verdegay, and J. C. Torres, "Coverage path planning with unmanned aerial vehicles for 3D terrain reconstruction," Expert Systems with Applications, vol. 55, pp. 441-451, 2016.
8. "melodic - ROS Wiki." Ros.org. http://wiki.ros.org/melodic (accessed 17 January 2022, 2022).
9. OSRF. "Gazebo." Gazebosim.org. http://gazebosim.org/ (accessed 17 January 2022, 2022).
10. J. Meyer, A. Sendobry, S. Kohlbrecher, U. Klingauf, and O. von Stryk, "Comprehensive Simulation of Quadrotor UAVs Using ROS and Gazebo," Berlin, Heidelberg, 2012: Springer Berlin Heidelberg, in Simulation, Modeling, and Programming for Autonomous Robots, pp. 400-411.
11. Z. Song, H. Zhang, X. Zhang, and F. Zhang, "Unmanned Aerial Vehicle Coverage Path Planning Algorithm Based on Cellular Automata," in 2019 15th International Conference on Computational Intelligence and Security (CIS), 13-16 Dec. 2019 2019, pp. 123-126, doi: 10.1109/CIS.2019.00034.

Authors Introduction

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.

Orcid:https://orcid.org/0000-0002-1363-3769

Dr. Norrima Binti Mokhtar

Dr. Norrima is working as a Senior Lecturer at Department of Electrical Engineering, UM, Malaysia. Her interests include Human Machine Interaction, Machine Learning & Robotics, Digital Image Processing

Researcher ID Link:

http://www.researcherid.com/rid/B-9395-2010

Noraisyah Binti Mohamed Shah

Dr. Noraisyah is working as a Senior Lecturer at Department of Electrical Engineering, UM, Malaysia. Her interests include Bio Medical Engineering & Signal Processing.

Researcher ID Link:

http://www.researcherid.com/rid/B-9338-2010

Mahidzal Bin Dahari

Dr. Mahidzal is working as an Associate Professor in Department of Electrical Engineering, UM, Malaysia. His research interest includes instrumentation, automation, robotics, and control strategy especially related to manufacturing technology, industrial robots, and oil & gas applications

Researcher ID Link:

http://www.researcherid.com/rid/B-5401-2010

Muhammad Syazni Ikmal Bin Ramlee

He received his bachelor's degree from the Department of Electrical Engineering, University of Malaya, Malaysia in 2022. He is currently working as Fiber Engineer in Maxis Broadband Sdn. Bhd.

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.

Masahiro Iwahashi

Professor Masahiro Iwahashi is serving at the department of Information, Telecommunication and Control Systems Group in Nagaoka University of Technology, Japan. His research themes include audio / video compression and communication, signal processing, lossless coding, and pattern recognition / classification in image processing applications. Profile and further information:

https://souran.nagaokaut.ac.jp/view?l=en&u=53

Blood Vessels Segmentation in Eye Fundus Images Using Image Processing Algorithms

Obaid Al-quraan
Department of Biomedical Systems and Medical Informatics Engineering, Yarmouk University, Irbid, Jordan

Hiam Alquran
Department of Biomedical Systems and Medical Informatics Engineering, Yarmouk University, Irbid, Jordan

Mohammed Alsalatie
The Institute of Biomedical Technology, King Hussein Medical Center, Royal Jordanian Medical Service, Amman 11855, Jordan

Wan Azani Mustafa
Faculty of Electrical Engineering Technology, University of Malaysia Perlis, UniCITI Alam Campus, Sungai Chuchuh, Perlis, Malaysia

Wan Khairunizam
Faculty of Electrical Engineering Technology, UniCITI Alam Campus,
Universiti Malaysia Perlis, 02100 Padang Besar, Perlis, Malaysia

E-mail: obaidah.quraan@gmail.com, heyam.q@yu.edu.jo, mhmdsliti312@gmail.com ,
wanazani@unimap.edu.my

Abstract

The retinal blood vessels have a huge impact on the diagnosis of eye diseases in addition to other systematic diseases in the human body. In this paper, we presented an automated segmentation method to extract retinal blood vessels, starting with preprocessing, then passing the image into segmentation stage using Bradley technique, and lastly, morphological operations. The proposed method was assessed and tested on STARE dataset, followed by comparing the auto-segmented images to the manually segmented ones. The comparison results Accuracy, Sensitivity, and Specificity were 94.63%, 95.02%, and 80.73% respectively.

Keywords: List four to six keywords which characterize the article.

1. Introduction

Fundus imaging technique has recently played a significant role in detecting and diagnosing many eye diseases such as Diabetic Retinopathy (DR) and Glaucoma. According to World Health Organization (WHO), DR and Glaucoma are considered among the leading causes of blindness and vision impairments in the world [1]. Their symptoms – in addition to other diseases that affect the vision- appear in the retina, where the blood vessels are located to supply eye parts with Oxygen and nutrients [2]. Thus, the segmentation of retinal vasculature and investigating the characteristics of it are used as a guide to physicians and specialists to identify, assess and diagnose the abnormalities linked with ocular diseases and other complications like cardiovascular system diseases [2], [3]. However, physicians may lead to misdiagnosis due to the manual methods of segmentation, and a large number of images needs to be interpreted [4], [5] specifically in developing countries, where there is always a shortage in specialists accompanied with less access to healthcare services [6]. Therefore, automated blood vessel segmentation will have more accuracy and high access to many medical specialists such as optometrists, ophthalmologists, and orthoptists remotely.

2. Literature view

During the last few years, many authors have contributed to the segmentation of retinal blood vessels. They presented various new methods to help the medical field in the stage of diagnosis. In this section, we show the most recent related works; like C. Zhou et al. [7] proposed an algorithm using a line detector to extract the large ones and applied the Hidden Markov model (HMM) to detect vessel centerlines that include thin ones. The method is tested on two databases (DRIVE and STARE) with an accuracy of 0.9475 and 0.9535, respectively. A. Ooi et al. [8] proposed an interactive blood vessel segmentation from retinal fundus image based on Canny edge detection. In another study, K. Upadhyay et al. [9] presented a rule-based retinal blood vessel segmentation algorithm that implements two multi-scale approaches, local directional-wavelet transform, and global curvelet transform, together in a novel manner for vessel enhancement and segmentation. The algorithm is tested on four public databases, with an average accuracy of 0.957. F. Tian et al. [10] proposed an improved Frangi Hessian model. They constructed it by introducing the scale equivalence factor and eigenvector direction angle of the Hessian matrix into the traditional Frangi filtering algorithm to enhance blood vessels of the global enhanced image. The improved algorithm was performed on the public DRIVE and STARE datasets. Accuracy, sensitivity, and specificity of retinal images in DRIVE and STARE are 95.54%, 69.42%, and 98.02% and 94.92%, 70.19%, and 97.71%, respectively. In a different paper, S. Swathi et al. [11] proposed a technique that creates a tri map consequently by using area highlights of veins. They apply various leveled picture tangling strategies to separate the vessel parts in obscure regions. In a paper by R. Kushol et al. [12], an efficient retinal blood vessel segmentation approach was proposed by constructing a 4-D feature vector for the output of Bendlet transform, which can capture directional information much more efficiently than the traditional wavelets. The approach is performed on two image datasets (DRIVE and STARE) with an average accuracy for vessel segmentation achieved of approximately 95%. Finally, O. Ramos-Soto et al. [13] proposed a method consisting of three stages, the main processing stage, consists of two configurations, the first to extract thick vessels through the new optimized top-hat, homomorphic filtering, and median filter. The proposed method was evaluated using two datasets (DRIVE and STARE) with an accuracy of 0.9667 and 0.9580, respectively. In this study, we present an automated segmentation method of retinal blood vessels from fundus images using the Bradley segmentation method accompanied by an image enhancement process and morphological operations.

3. Method and Materials

The method is divided into 3 main stages: In the first stage, the fundus image was preprocessed and enhanced to create a mask. In the second stage, a green channel was extracted to apply the histogram equalization and Bradley method on the masked image. Finally, blood vessels were extracted followed by performing some morphological operations. The proposed method is shown in Figure 1.

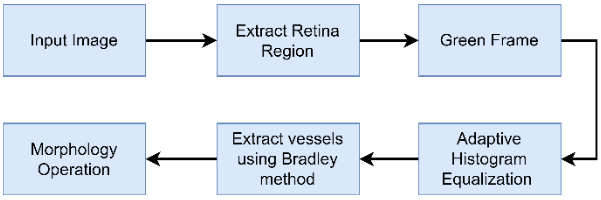

Input Image
Extract Retina Region
Green Frame
Adaptive Histogram Equalization
Extract vessels using Bradley method
Morphology Operation

Figure. 1. The flow chart of proposed method

3.1. *Dataset description*

The presented method was tested and assessed on The STARE (Structured Analysis of the Retina) database, which was initiated in 1975 by Michael Goldblum, M.D., at the University of California, San Diego. The total set consists of around 400 raw fundus images, 20 of them were manually segmented as ground truth images [14].

3.2. *Preprocessing and Enhancement*

In this stage, the fundus image is in RGB color space, it was transformed into HSV (Hue, Saturation, Value) color space. The mask is created by extracting the Hue plane and then binarizing it. We chose to do this step first to separate the foreground from the noisy background and to eliminate the noise. The colored image was multiplied by the mask and then converted into a gray level by obtaining the green plane. In the gray level in the fundus image, the vessels have dark pixels, whereas the surrounding regions have bright pixels. We utilized contrast-limited adaptive histogram equalization to increase the contrast between dark regions (vessels) and bright regions (non-vessels). This vital process enhanced the contrast and adjusted intensities to make vessels darker relative to the neighbor pixels. Moreover, CLAHE is different from traditional Histogram equalization by computing multiple histograms and redistributing the intensity values over the whole image. In this stage, the fundus image is in RGB (Red, Green, and Blue) color space, it was transformed into HSV.

3.3 *Segmentation*

Retinal blood vessel segmentation is the core stage in our study, Bradley method is used here to detect and extract the blood vessels. Bradley's algorithm calculates the average brightness of the neighbor pixels within a window of a specific size. Then, it adjusts pixels, whose intensities are less than a certain percent of the average brightness, to become (0). The resulting image shows that the blood vessels combined with tiny pieces of noise appeared in black while the rest image parts were in white [15].

3. ***Post processing***

In the last stage, the morphological opening operation was performed to remove the tiny objects that surround the vessels. After that, the image was passed through the morphological process to improve the shape of the extracted vessels. Morphological closing was selected to connect between pixels and restore the shape of branches.

4. Results and Discussion

In our work, the proposed method was performed on STARE dataset as following, the colored fundus image was converted into HSV color space, and the Hue channel was extracted. Figure. 2 shows the conversion.

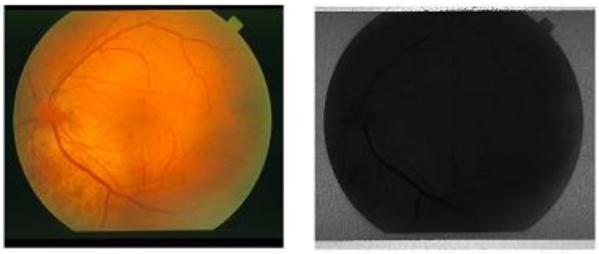


Figure. 2. (a) original image (b) Hue channel

Hue channel was binarized and improved by a morphological opening to create the mask, after that the original image was multiplied by the mask as shown in Figure 3.

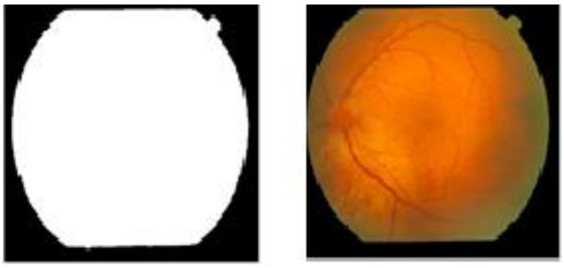


Figure. 3. (a) Improved mask (b) masked colored image

The masked RGB mage requires green channel to be extracted and then to perform contrast-limited adaptive histogram equalization (CLAHE) on it to amplify the contrast difference between the dark and the bright regions. Figure. 4 illustrates the effect of CLAHE on visualizing the blood vessels.

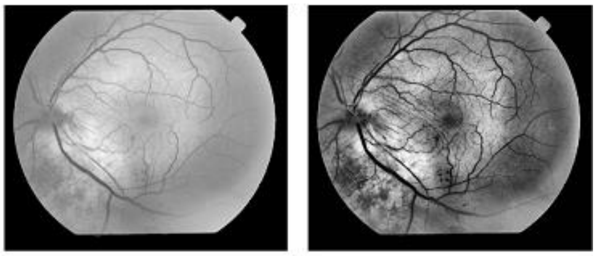


Figure 4. (a) Green channel of masked image (b) using CLAHE

Bradley method transformed the blood vessels into black pixels, meanwhile the surrounding regions were transformed unto white pixels as shown in Figure 5.

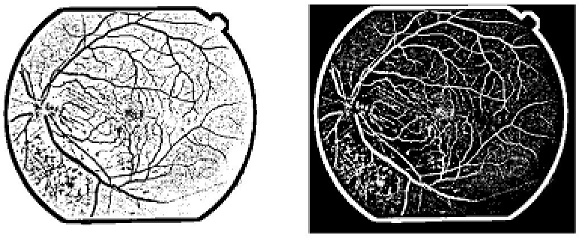


Figure. 5. (a) the image resulted from Bradley method (b) the complement of it.

Morphological operations, following then area opening, are performed on the Bradley image to give the extracted retina vasculature in its last shape. Figure 6 shows the extracted blood vessels.

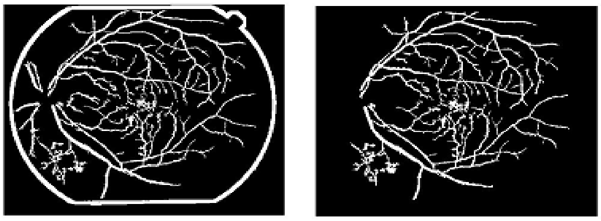


Figure. 6. (a) applying area opening (b) morphological closing

The proposed segmentation method is tested and evaluated on the STARE dataset. The dataset contains 20 manually segmented images that can validate this method by comparing these images to the extracted blood vessels resulting from our work. Accuracy (ACC), Sensitivity (Se), and Specificity (Sp) are the main measures considered here to assess the resultant images. Access refers to the proportion of accurately classified pixels in total pixels of the fundus retinal image. Se and Sp refer to the respect of vascular and nonvascular pixels that are recognized accurately in the segmentation result, respectively. From the comparison process, the performance of blood vessels segmentation using the proposed method. From the comparison process, the performance of blood vessel segmentation using the proposed method achieved an accuracy of 94.63%, a sensitivity of 95.02%, and a specificity of 80.73%. Consequently, the proposed method showed prominent results, especially in ACC and Se, relative to the results mentioned in the literature view. Specifically, our work tried to make the balance between maintaining the small details of the blood vessels and reducing the noise as possible as we can. Furthermore, the proposed method used a simple approach to processing the fundus images in an attempt to avoid the undesired complexities in the processing stages. However, there were some problems appeared in some images in the dataset. For example, some images have excessive illumination that left a spotlight on the vessels, which, in turn, will make it more difficult to deal with these images. Figure 7 shows the comparison between our proposed method and manual segmented.

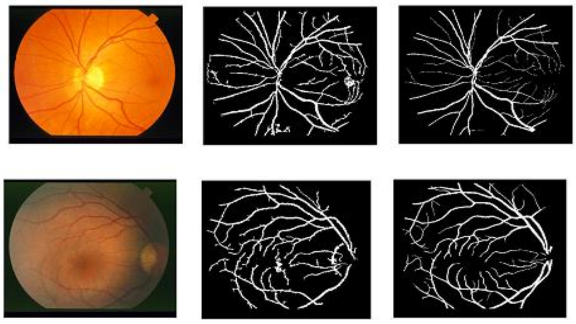


Figure 7. Original images (b) auto segmented images (c) manual segmented images.

5. Conclusion

The segmentation of retinal blood vessels has an increasing contribution to the ophthalmology field. In this paper, an automated method of segmentation is proposed to extract retinal blood vessels. We illustrated the main stages of the method, and the effect was shown on the fundus images. The proposed method was tested on the publicly available STARE dataset, and the performance results were 94.63% of Accuracy, 95.02% of Sensitivity, and 80.73% of Specificity. However, we faced some problems related to image quality. Therefore, the future study may work on solving these problems, as well as we may work on improving the performance of the method on other datasets.

References

1. "Blindness and vision impairment." https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment (accessed Nov. 03, 2022)
2. M. D. Abràmoff, S. Member, M. K. Garvin, and M. Sonka, "Retinal Imaging and Image Analysis," vol. 3, pp. 169–208, 2010.
3. C. G. Owen et al., "Retinal arteriolar tortuosity and cardiovascular risk factors in a multi-ethnic population study of 10-year-old children; the Child Heart and Health Study in England (CHASE)," Arterioscler. Thromb. Vasc. Biol., vol. 31, no. 8, pp. 1933–1938, Aug. 2011
4. R. Khanduzi and A. K. Sangaiah, "A fast genetic algorithm for a critical protection problem in biomedical supply chain networks," Appl. Soft Comput., vol. 75, pp. 162–179, Feb. 2019
5. S. Pirbhulal, O. W. Samuel, W. Wu, A. K. Sangaiah, and G. Li, "A joint resource-aware and medical data security framework for wearable healthcare systems," Futur. Gener. Comput. Syst., vol. 95, pp. 382–391, Jun. 2019
6. D. H. Peters, A. Garg, G. Bloom, D. G. Walker, W. R. Brieger, and M. Hafizur Rahman, "Poverty and access to health care in developing countries," Ann. N. Y. Acad. Sci., vol. 1136, pp. 161–171, 2008
7. C. Zhou, X. Zhang, and H. Chen, "A new robust method for blood vessel segmentation in retinal fundus images based on weighted line detector and hidden Markov model," *Comput. Methods Programs Biomed.*, vol. 187, p. 105231, Apr. 2020, doi: 10.1016/J.CMPB.2019.105231.
8. A. Z. H. Ooi *et al.*, "Interactive Blood Vessel Segmentation from Retinal Fundus Image Based on Canny Edge Detector," *Sensors 2021, Vol. 21, Page 6380*, vol. 21, no. 19, p. 6380, Sep. 2021, doi: 10.3390/S21196380.
9. K. Upadhyay, M. Agrawal, and P. Vashist, "Unsupervised multiscale retinal blood vessel segmentation using fundus images," *IET Image Process.*, vol. 14, no. 11, pp. 2616–2625, Sep. 2020, doi: 10.1049/IET-IPR.2019.0969.
10. F. Tian, Y. Li, J. Wang, and W. Chen, "Blood Vessel Segmentation of Fundus Retinal Images Based on Improved Frangi and Mathematical Morphology," *Comput. Math. Methods Med.*, vol. 2021, 2021, doi: 10.1155/2021/4761517.
11. S. Swathi, S. Sushma, C. Devi Supraja, V. Bindusree, L. Babitha, and V. Vijay, "A hierarchical image matting model for blood vessel segmentation in retinal images," *Int. J. Syst. Assur. Eng. Manag. 2021 133*, vol. 13, no. 3, pp. 1093–1101, Sep. 2021, doi: 10.1007/S13198-021-01397-0.
12. Kushol R, Kabir MH, Abdullah-Al-Wadud M, Islam MS. Retinal blood vessel segmentation from fundus image using an efficient multiscale directional representation technique Bendlets. Math Biosci Eng. 2020 Nov 6;17(6):7751-7771. doi: 10.3934/mbe.2020394. PMID: 33378918.
13. O. Ramos-Soto et al., "An efficient retinal blood vessel segmentation in eye fundus images by using optimized

top-hat and homomorphic filtering," Comput. Methods Programs Biomed., vol. 201, p. 105949, Apr. 2021

14. https://cecas.clemson.edu/~ahoover/stare/
15. Derek Bradley & Gerhard Roth (2007) Adaptive Thresholding using the Integral Image, Journal of Graphics Tools, 12:2, 13-21, DOI: 10.1080/2151237X.2007.10129236

Authors Introduction

Obaida Al-quraan

He obtained his Bsc degree in Biomedical Engineering from Yarmouk University in Jordan. He is currently a trainee in the Directorate of Biomedical Engineering – Royal Scientific Society in Jordan. His research interests include Image and Signal Processing, Artificial Intelligence, and Medical Imaging.

Hiam H Alquran

Hiam H Alquran, Associate Prof at Department of Biomedical Systems and Informatics Engineering, Yarmouk University , Jordan. Alquran received her PhD. (2014) degree in Biomedical and Biotechnology Engineering from Massachusetts Lowell University USA .M.Sc. degree (2008) in Automation Engineering from Yarmouk University. B.S.c in Biomedical Engineering form JUST -Jordan (2005). Research Interest in Medical Image Processing, Digital Signal Processing, Pattern Recognition and Deep Learning.

Mohammed Alsalatie

He received his B.S. degree in biomedical engineering from Yarmouk University in Jordan. he is a lecturer at The Institute of Biomedical Technology, Royal Jordanian Medical Service, Amman, Jordan. His research interests include Image and Signal Processing, and Artificial intelligenc.

Wan Azani Mustafa

He obtained his PhD in Mechatronic engineering from University Malaysia Perlis. He is currently in the Faculty of Electrical Engineering Technology, Universiti Malaysia perlis, as a Senior Lecturer. His research interests include Image and Signal Processing, Artificial intelligence, Medical Imaging and Robotic.

Wan Khairunizam

Khairunizam WAN 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 an Assoc. Prof. at School Of Mechatronic Engineering, University Malaysia Perlis. 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.

Automated Diagnosis of Eye Fundus Images

Ala'a Zyout
Department of Biomedical Systems and Informatics Engineering, Yarmouk University, Irbid, Jordan

Hiam Alquran
Department of Biomedical Systems and Informatics Engineering, Yarmouk University, Irbid, Jordan

Wan Azani Mustafa
Faculty of Electrical Engineering & Technology, Campus Pauh Putra, Universiti Malaysia Perlis, Arau 02000, Perlis, Malaysia
Advanced Computing, Centre of Excellence (CoE), Universiti Malaysia Perlis (UniMAP), Arau 02000, Perlis, Malaysia,

Mohammed Alsalatie
The Institute of Biomedical Technology, King Hussein Medical Center, Royal Jordanian Medical Service, Amman 11855, Jordan

Alaa Al-Badarneh
Department of Biomedical Systems and Informatics Engineering, Yarmouk University, Irbid, Jordan

Wan Khairunizam
Faculty of Electrical Engineering & Technology, Universiti Malaysia Perlis, 02100 Padang Besar, Perlis, Malaysia
E-Mail: alzuet@yu.edu.jo, heyam.q@yu.edu.jo, wanazani@unimap.edu.my, mhmdsliti312@gmail.com, alaa_aaa@yu.edu.jo, khairunizam@unimap.edu.my

Abstract

Eye disease is a severe health problem. Advanced stages of the disease may lead to vision loss. Early detection may limit the development of the severity and enhance the chance of treatment. Eye disease comes from various factors such as diabetes, increasing pressure in the eye (Glaucoma), and age-related macular degeneration. Ocular fundus 2D images are one of the most common tools used to diagnose the lining of tissue eyes. Huge data availability, increasing cases, and heavy responsibility in the health sector encourage seeking new diagnosis techniques to enhance accuracy and reduce false positive and false negative diagnoses. Computer-aided diagnosis (CAD) is the state-art-technology. This paper proposes a CAD system that combines image processing techniques and artificial intelligence. The proposed method used the green channel of fundus eye images to extract the most representative features by the trained convolutional neural network to classify five eye diseases of fundus images. The build CAD system exploits deep learning and support vector machine classifier to achieve a highly accurate model of 98% for five types of eye diseases.

Keywords: Fundus images, deep learning, support vector machine.

1. Introduction

A fundus image is a 2D projection of the fundus that was made using a monocular camera, which can be acquired in a non-invasive and cost-effective manner, which makes them more suitable for large-scale screening than other eye scans including optical coherence tomography images and angiographs [1]. Clinically, the usage of fundus images for early eye disease detection is crucial. Deep learning is becoming more and more common in related applications due to its strong performance [2].

One of the available of the deep convolutional neural network called ResNet is reviewed by Esfahani et al. and used for classification of eye fundus Images into two groups of normal and diabetic images and simulation result has achieved 85% accuracy and 86% sensitivity [3]. While Raghavendra et al. trained eighteen layers of convolutional neural networks (CNNs) to extract powerful features to categorize into normal and glaucoma with accuracy of 98.13%. Burlina et al. [4] investigate the suitability of applying image features calculated from pre-trained deep neural networks on the Age-related macular generation identification which show achievement with 92% to 95% accuracy. Choi et al. [5] used fundus images and multi-categorical deep learning algorithms to automatically identify 10 categories of retinal disorders with 30.5% accuracy. In addition, the transfer learning incorporated with ensemble classifier of enhancement the multi-categorical classification performance with accuracy of 36.7%. The multi-category classifier demonstrated accuracy of 72.8% when three integrated normal, background diabetic retinopathy (BDR), and dry age-related macular degeneration (AMD) were taken into account. While using integrated normal, dry AMD, and wet AMD, achieved 77.2% accuracy, three other integrated normal, BDR, and proliferative diabetic retinopathy (PDR) demonstrated accuracy of 80.8%. Additionally, using the same model structure, 5 disease categories normal, background DR, PDR, dry AMD, and wet AMD achieved a maximal accuracy of 59.1% [6].

To have a greater knowledge of the several categories used to classify fundus images, Chea et al. [7] provided deep neural networks and efficient image reprocessing methods, such as shrinking the region of interest, iso-luminance plane contrast-limited adaptive histogram equalization, and data augmentation, to classify the three most prevalent eye diseases, DR, glaucoma (GLC), and AMD, which achieved peak and average accuracies of 91.16% and 85.79%, respectively.

2. Materials and Methods

The set of steps followed in this work are shown in Figure 1. The proposed methodology is divided into four main parts: 1) data augmentation, 2) Image preprocessing, 3) Convolution Neural Network (CNN), and then 4) Support Vector Machine (SVM).

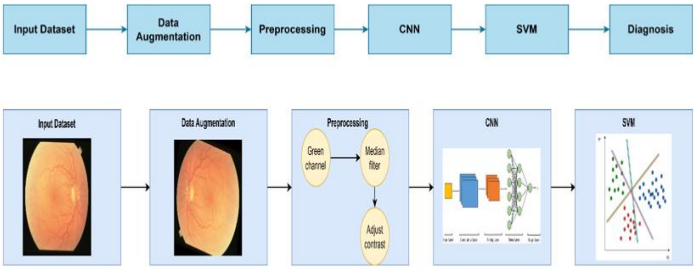
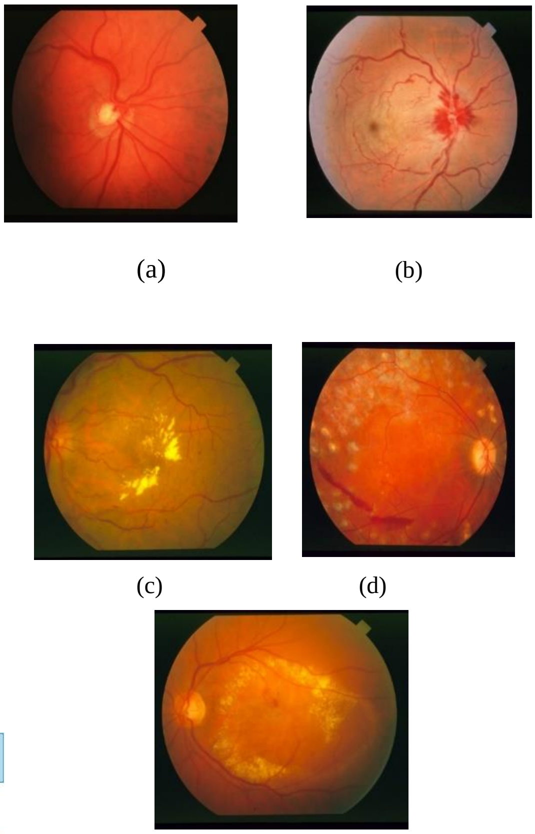


Figure 1: The proposed method architecture

2.1. *Image Database*

This paper utilizes full raw images from the STARE database (structured analysis of retina) project, which is carried out in 1975 in California [8]. Almost 400 images are acquired and diagnosed for 13 types of eye diseases. Only five types of classes are selected to build an automated diagnosis system in this paper. The aim behind this selection is the intended system aims to diagnose a single kind of disease per image, not multi diseases, while some cases are diagnosed for multiclass. The classes are normal, Central Retinal Vein Occlusion (CRVO), Background Diabetic Retinopathy (BDR), Proliferative Diabetic Retinopathy (PDR), and Choroidal Neovascularization (CNS). One sample for each class is illustrated in Figure 2 (a), (b), (c) (d), and (e), respectively.



(a) (b)

(c) (d)

(e)

Figure 2: Samples of eye fundus images (a) Normal, (b) CRVO, (c) BDR, (d) PDR, and (e) CNS [8].

Table 2(1?). The planning and control components.

Class	Name	Number of images
Class00	Normal	39
Class05	Central Retinal Vein Occlusion (CRVO)	25
Class07	Background Diabetic Retinopathy (BDR)	69
Class08	Proliferative Diabetic Retinopathy (PDR)	23
Class13	Choroidal Neovascularization (CNS)	61

2.2. *Image Augmentation*

Over-fitting may occur when a classifier model is overly complicated in relation to the quantity of training examples. The deep learning community has proposed numerous regularization and data augmentation techniques to overcome this issue, including various data augmentation techniques as cropping, flipping, and translating [9]. The need for large data volume in deep learning helped in using data augmentation to increase the number of samples for the training set. In this paper all classes are augmented to 100 images for each class. The total images for five classes are 500.

2.3. *Image Preprocessing*

Medical images are affected by noise during the image collection process, so medical image analysis requires an image pre-processing stage. At first, the green channel was used in RGB path which gives better results compared to other channels. In order to remove input noise from the image while preserving edges, especially salt and pepper noise, a median filter was used [10]. After that, image adjustment was utilized to enhance the contrast and brightness of the image.

2.4. *Convolution Neural Network*

Convolution Neural Network (CNN) is developed based on multi-layer neural networks, which are a type of deep learning model created specifically for image classification and recognition. For high accuracy, CNN uses automatic feature extraction. It employs specific convolution layers, pooling techniques, and parameter sharing [11]. Modified CNN was devised in this study [12]. The structure started from an input layer of green channel image, while the classification layer terminated the proposed network. The goal of using the modified CNN is to extract the features for use in the SVM part. Figure 3 shows the layout of its layers with distinct mass convolutional layers.

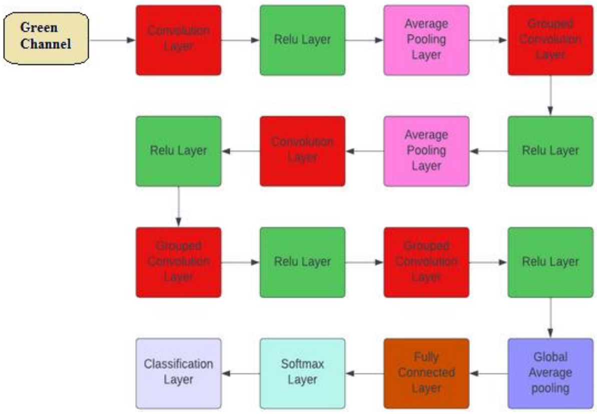


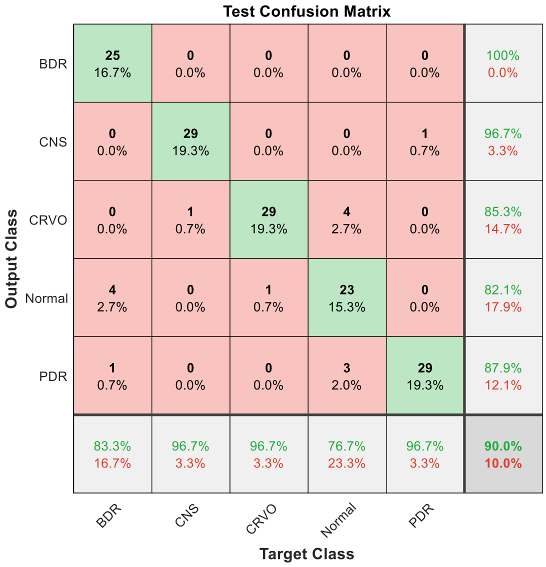
Figure 3: The modified CNN structure [12]

2.5. *Support Vector Machine*

Supportive Vector Machine (SVM) is an educational technique used to classify data entered two supervised classes. The SVM algorithm creates a model that predicts the new instance class using specific training data. Finding the ideal cut-off hyperplane for a data set - which optimizes the distance between the nearest data point and the cut-off hyperplane - is the goal of SVM [13]. In this work, a classification layer of modified CNN output was used as the property input data for an SVM classifier, to achieve the output objective of classifying the five classes of fundus images. The complete input data set was divided as: 70% subsets were used as training data, and the remaining subset was used to test a prediction model.

3. Results and Discussion

The green chancel of each image is selected and passed to the modified CNN. The first scenario is deep learning classification. Figure 4 shows the output of the modified CNN model after transfer learning to be compatible with five classes.

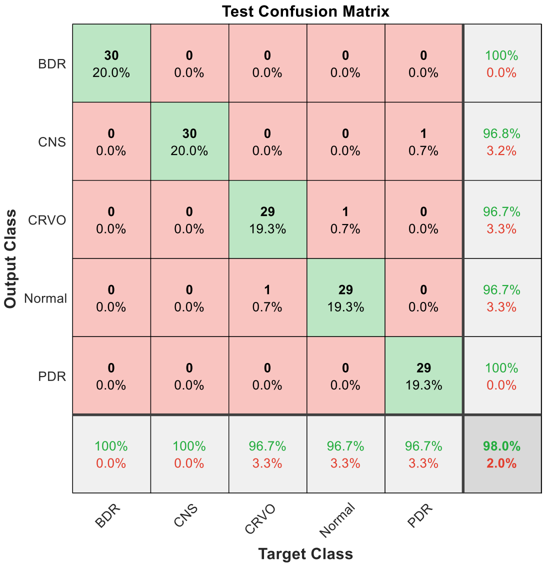


	BDR	CNS	CRVO	Normal	PDR	
BDR	**25** 16.7%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	100% 0.0%
CNS	**0** 0.0%	**29** 19.3%	**0** 0.0%	**0** 0.0%	**1** 0.7%	96.7% 3.3%
CRVO	**0** 0.0%	**1** 0.7%	**29** 19.3%	**4** 2.7%	**0** 0.0%	85.3% 14.7%
Normal	**4** 2.7%	**0** 0.0%	**1** 0.7%	**23** 15.3%	**0** 0.0%	82.1% 17.9%
PDR	**1** 0.7%	**0** 0.0%	**0** 0.0%	**3** 2.0%	**29** 19.3%	87.9% 12.1%
	83.3% 16.7%	96.7% 3.3%	96.7% 3.3%	76.7% 23.3%	96.7% 3.3%	**90.0%** **10.0%**

Figure 4: Confusion matrix using deep learning approach

As clear in Figure 4 twenty-five occurrences are correctly classified for BDR class with a sensitivity of 83.3% and a misclassification rate of 16.7%, besides a precision of 100%. Meanwhile, the CNS has equal results in both recall and positive predictive value with 96.7%. On the other hand, the CRVO obtain 96.7% sensitivity and 85.3 % of precision. The Normal class has the lowest results in terms of sensitivity and precision at 76.7% and 82.1%, respectively. The fifth class is PDR, where 29 cases are classified correctly among 30 test cases with a recall of 96.7% and a precision of 87. 9%. The overall accuracy does not exceed 90% for all classes.

The improvement is obtained by combining deep learning as an automated features extractor and machine learning as classifiers. This method has been widely used in the literature [14]. Therefore, the modified CNN extracts five features. Each one represents its corresponding class. These features pass to multi-class SVM. However, only 70%of them train SVM model and the rest evaluates the model generalization. Figure 5 explains the results of the hybrid model.



	BDR	CNS	CRVO	Normal	PDR	
BDR	**30** 20.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	100% 0.0%
CNS	**0** 0.0%	**30** 20.0%	**0** 0.0%	**0** 0.0%	**1** 0.7%	96.8% 3.2%
CRVO	**0** 0.0%	**0** 0.0%	**29** 19.3%	**1** 0.7%	**0** 0.0%	96.7% 3.3%
Normal	**0** 0.0%	**0** 0.0%	**1** 0.7%	**29** 19.3%	**0** 0.0%	96.7% 3.3%
PDR	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**29** 19.3%	100% 0.0%
	100% 0.0%	100% 0.0%	96.7% 3.3%	96.7% 3.3%	96.7% 3.3%	**98.0%** **2.0%**

Figure 5: Confusion matrix using hybrid (machine learning and deep learning)

As clear in Figure 5 the performance is improved where all cases in BDR and CNS classes are distinguished correctly with a sensitivity of 100%. One case from the PDR class is misclassified as CNS. Therefore, the precision of the CNS class is reduced to 96.8%. For CRVO cases, one occurrence is misclassified as normal. Therefore, the sensitivity is 96.7% and one case from normal is classified as CRVO. The precision is reduced to 96.7%. The sensitivity for normal and PDR is 100%, but the positive predictive value for normal is 96.7%. The overall results are better than deep learning classification scenario. The overall accuracy is 98%. Figure 6 illustrates the comparison between the two scenarios.

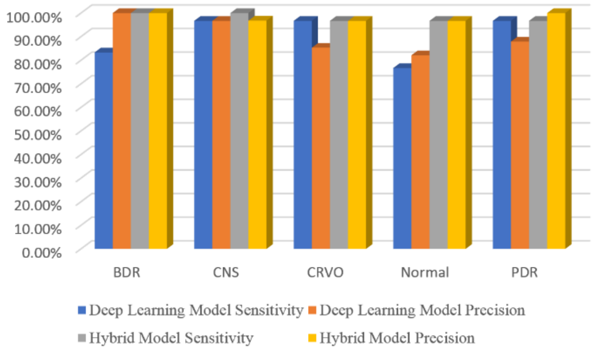


Figure 6: Comparison two approaches in terms sensitivity and precision

As clear from Figure 6 the combination between deep learning and machine learning classifiers yields high performance results in distinguishing five fundus classes. When the proposed approach results are compared with literature, it indicates to prominent computer aided diagnosis for eye diseases.

4. Conclusion

This paper presents a new CAD system for eye fundus images diagnoses for five classes: Normal, Central Retinal Vein Occlusion (CRVO), Background Diabetic Retinopathy (BDR), Proliferative Diabetic Retinopathy (PDR), and Choroidal Neovascularization (CNS) using extraction features from modified pre-trained CNN. The representative features passed for multi-class SVM. The system obtains high accuracy reaching 98% for all five classes and a sensitivity of 100 for BDR, CNS, and normal. Whereas the precision is 100% for BDR, CRVO, Normal, and PDR. The system can be dependable software for eye diagnosis by utilizing a large dataset to be used in rural areas and poor countries.

References

1. V. G. Edupuganti, A. Chawla and A. Kale, "Automatic Optic Disk and Cup Segmentation of Fundus Images Using Deep Learning," 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 2018, pp. 2227-2231, doi: 10.1109/ICIP.2018.8451753.
2. Li T, Bo W, Hu C, Kang H, Liu H, Wang K, Fu H. Applications of deep learning in fundus images: A review. Med Image Anal. 2021 Apr;69:101971. doi: 10.1016/j.media.2021.101971. Epub 2021 Jan 20. PMID: 33524824.
3. Esfahani, M. T., Ghaderi, M., & Kafiyeh, R. J. L. E. J. P. T. (2018). Classification of diabetic and normal fundus images using new deep learning method. Leonardo Electron. J. Pract. Technol, 17(32), 233-248
4. Raghavendra, U., Fujita, H., Bhandary, S. V., Gudigar, A., Tan, J. H., & Acharya, U. R. (2018). Deep convolution neural network for accurate diagnosis of glaucoma using digital fundus images. Information Sciences, 441, 41-49.
5. P. Burlina, D. E. Freund, N. Joshi, Y. Wolfson and N. M. Bressler, "Detection of age-related macular degeneration via deep learning," 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), Prague, Czech Republic, 2016, pp. 184-188, doi: 10.1109/ISBI.2016.7493240.
6. Choi JY, Yoo TK, Seo JG, Kwak J, Um TT, Rim TH. Multi-categorical deep learning neural network to classify retinal images: A pilot study employing small database. PLoS One. 2017 Nov 2;12(11):e0187336. doi: 10.1371/journal.pone.0187336. PMID: 29095872; PMCID: PMC5667846.
7. N. Chea and Y. Nam, "Classification of fundus images based on deep learning for detecting eye diseases," Computers, Materials & Continua, vol. 67, no.1, pp. 411–426, 2021.
8. Boyd, J. (1996). STARE software documentation: Disk—Optic disk locator. Vis. Comput. Lab., Dept. Elect. Comput. Eng., Univ. California, San Diego, CA.
9. Zhong, Zhun, Liang Zheng, Zhedong Zheng, Shaozi Li, and Yi Yang. "Camstyle: A novel data augmentation method for person re-identification." *IEEE Transactions on Image Processing* 28, no. 3 (2018): 1176-1190.
10. Badarneh, Alaa, Isam Abu-Qasmeih, Mwaffaq Otoom, and Mohammad A. Alzubaidi. "Semi-automated spine and intervertebral disk detection and segmentation from whole spine MR images." *Informatics in Medicine Unlocked* 27 (2021): 100810.
11. Li, Lingyun. "Application of deep learning in image recognition." In *Journal of Physics: Conference Series*, vol. 1693, no. 1, p. 012128. IOP Publishing, 2020.
12. Alquran, H.; Alsalatie, M.; Mustafa, W.A.; Abdi, R.A.; Ismail, A.R. Cervical Net: A Novel Cervical Cancer Classification Using Feature Fusion. Bioengineering 2022, 9, 578.

13. H. Alquran et al., "The melanoma skin cancer detection and classification using support vector machine," 2017 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT), Aqaba, Jordan, 2017, pp. 1-5, doi: 10.1109/AEECT.2017.8257738.
14. H. Alquran, W. Azani Mustafa, I. Abu Qasmieh, Y. Mohd Yacob, M. Alsalatie et al., "Cervical cancer classification using combined machine learning and deep learning approach," Computers, Materials & Continua, vol. 72, no.3, pp. 5117–5134, 2022.

Authors Introduction

Ala'a Zyout

She received her B.S degree in Biomedical Systems Engineering from Yarmouk University in Jordan. Since 2013, she has been working as a laboratory instructor with the Department of Biomedical Systems and Informatics Engineering, Hijjawi Faculty for Engineering Technology, Yarmouk University, Jordan

Hiam H Alquran,

She is Associate Prof at Department of Biomedical Systems and Informatics Engineering, Yarmouk University , Jordan. Alquran received her PhD. (2014) degree in Biomedical and Biotechnology Engineering from Massachusetts Lowell University USA .M.Sc. degree (2008) in Automation Engineering from Yarmouk University. B.S.c in Biomedical Engineering form JUST -Jordan (2005). Research Interest in Medical Image Processing, Digital Signal Processing, Pattern

Mohammed Alsalatie

He received his B.S. degree in biomedical engineering from Yarmouk University in Jordan. he is a lecturer at The Institute of Biomedical Technology, Royal Jordanian Medical Service, Amman, Jordan. His research interests include Image and Signal Processing, and Artificial intelligen

Wan Azani Mustafa

He obtained his PhD in Mechatronic engineering from University Malaysia Perlis. He is currently in the Faculty of Electrical Engineering Technology, Universiti Malaysia perlis, as a Senior Lecturer. His research interests include Image and Signal Processing, Artificial inte ligence, Medical Imaging and Robotic.l

.

AlaAa Ahmad Badarneh

Alaa Badarneh currently works at the Department of Biomedical Systems and Informatics Engineering, Yarmouk University, Jordan. She obtained her MSc Computer Engineering from Yarmouk University, Jordan. Her research interests include image processing and artificial intelligence.

.

Wan Khairunizam

Khairunizam WAN 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 an Assoc. Prof. at School Of Mechatronic Engineering, University Malaysia Perlis. He is member of Board of Engineer and Institute of Engineer, Malaysia. His research interest is in Human-Computer Interaction (), Intelligent Transportation System, Artificial Intelligence and Robotics.

The Dam Gate Cybersecurity Testbed

Chen-Yu Lee
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

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

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
E-mail: cylee@cans.ee.ncku.edu.tw, ihliu@cans.ee.ncku.edu.tw, mwchang@cans.ee.ncku.edu.tw, jsli@mail.ncku.edu.tw[*]
www.ncku.edu.tw

Abstract

The testbeds are very important for cybersecurity research on critical infrastructure. In today's drastic climate change, the dam gate control system is a very important part of the critical infrastructure for people's livelihood. In traditional research, because the real control system cannot be used directly, most of the research can only be carried out in a simulation way. The research based on simulation alone lacks practical value due to too many assumptions. This research was supported by the Water Resources Agency, Ministry of Economic Affairs and National Science and Technology Council in Taiwan. The gate control cybersecurity testbed was built with a blueprint of the real world.

Keywords: Cyber Security, ICS Security, PLC, Dam Gate Testbed

1. Introduction

In the era of rapid development of modern network communication and industrial control, large-scale key infrastructures have also been built by various countries, aiming to activate the industrial economy, develop an all-around industrial environment that benefits the country, and gradually improve the industry to meet the needs of the people.

There are many elements in the Industrial Control System (ICS)[1] such as the Industrial Internet of Things or the Supervisory Control and Acquisition System (SCADA), etc. Among them, the Programmable Logic Controller (PLC) is a key operating component. It can not only communication between the device's signals, and can be directly connected to the computer for signal transmission.

And because PLC is concerned with the physical behavior of the infrastructure, it will also face the information security crisis of cyber-attacks[2]. Russian hackers launched a data destruction attack on a Ukrainian power plant on April 8, 2022, and related companies cooperated with Ukrainian authorities to successfully prevent this attack[3].

In order to maintain the network security of critical infrastructure, this paper designs a set of testbeds that simulate water resource basins and connects PLC and other physical devices to build a set of physical testbeds for the implementation of new technologies for key infrastructure.

2. Background

In this section, we mainly discuss the components based on the critical infrastructure testbed. In traditional critical infrastructure, there are always problems, whether it is production efficiency or information security attack and defense[4], but in practice, it is impossible to directly implement new research or technology to critical infrastructure, because if there is a problem with its operation, it will cause serious problems. The country suffered great harm. Therefore, based on various considerations, most research will first be established and analyzed on a testbed, and then implemented in critical infrastructure.

2.1. *Programmable logic controller (PLC)*

PLC can control commands through memory access, and can also use various modules for customized functions. In recent years, the PLC in the industrial environment has gradually developed into a microcomputer. Whether it is digital and analog output and input, or directly building a human-machine interface (HMI), PLC has played a pivotal role in the industrial environment.

2.2. *Modbus/TCP*

Modbus is currently a communication protocol widely used in PLCs and has also been extended to become the communication protocol standard of the entire industrial environment. And because of the convenience, the current industrial environment gradually relies on Ethernet as a connection, so Modbus TCP is often used in industrial environments. However, with the advantages of such convenience and flexibility, Modbus is also vulnerable to unauthorized writing and packet analysis because it is transmitted in plain text, so it will be subject to behaviors such as penetration attacks.

2.3. *Critical infrastructure testbed*

The establishment of the testbed can cooperate with the supervisory control and data acquisition system (SCADA) [5] and can be classified by many aspects, mainly divided into three categories: information technology, communication, and operation technology, and there will also be different settings for each research. Key infrastructure testbed such as water conservancy and power generation. The establishment of the testbed is mainly aimed at the fact that if various tests are directly carried out on key infrastructures, it may cause them to be shut down or damaged, and the country will suffer huge losses. Therefore, it is very important to set up a testbed that meets its own needs and conduct tests on it.

3. Design of Dam Gate Testbed

Due to the cybersecurity issues of critical infrastructure, this paper designs a testbed for simulating the dam gate system. Because each gate is driven by the PLC, and the PLC has registers to access related instructions and data. Therefore, this information can express various values of the current industrial environment such as water pressure

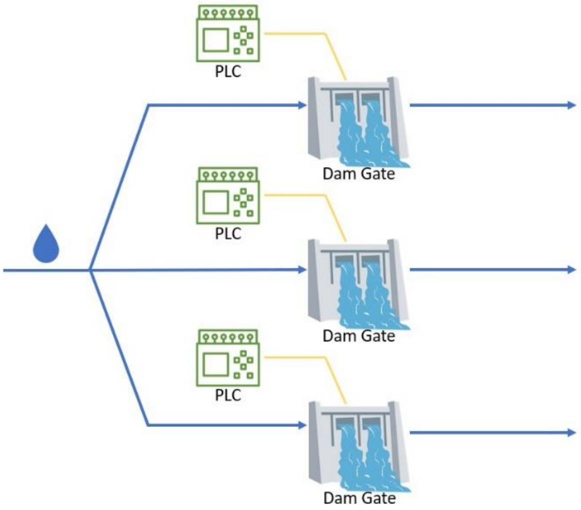


Fig. 1. Dam Gate Testbed Framework

and current. Taking Modbus/TCP as an example, a testbed for simulated dam gates in the computer is designed, and the gates are all connected and transmitted by physical PLC. In order to facilitate users to conduct research related to industrial control.

3.1. *System architecture*

First of all, we designed the framework in Fig. 1. according to the actual interaction between PLC and computer and controlled specific gates by sending relevant packet information through PLC. In addition, we have also designed the relevant main control panel to transmit digital signals to the PLC through the switch, which can be more suitable, and the output signal of the PLC can also be controlled by the program without connecting the main control panel to maintain the overall operation flexibility. Finally, we also design a virtual gate system based on packet sending and receiving, and the establishment of the testbed will be described in detail in 3.2.

3.2. *Virtual dam gate testbed*

In this section, we establish a set of virtual dam gate testbed based on Modbus/TCP in Fig. 2. By connecting the computer and the PLC and setting the relevant virtual gate status, you can observe the changes in various values:

- Packet flow
- Gate state change
- Water flow information

It is hoped that this virtual testbed can carry out simulation research such as PLC attack and defense and gate behavior first. In Section 4, we design a set of physical dam gate testbed based on the virtual testbed.

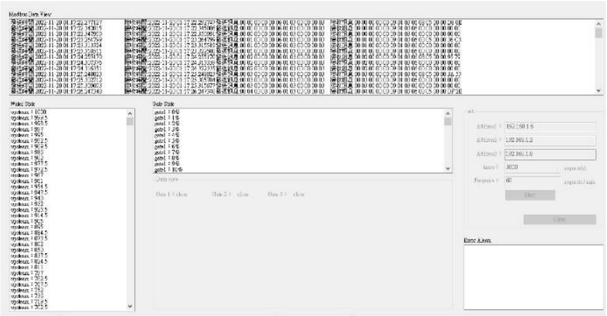


Fig. 2. Virtual Dam Gate Testbed

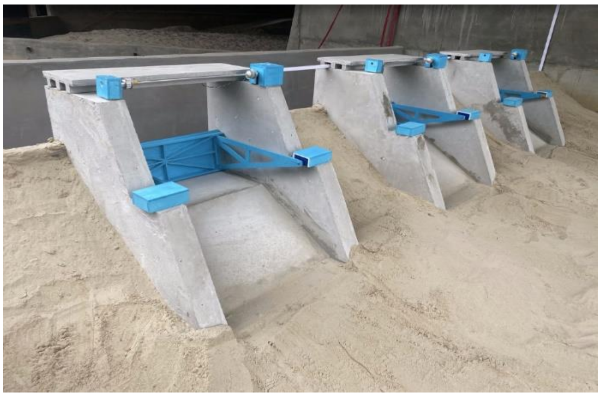


Fig. 3. Physical Dam Gate Testbed

4. Experiment

The team built a set of physical dam gate testbed as shown in Fig. 3. For the sake of reality, the physical gates and terrain heights were also designed to make it fit the practical technology. Moreover, computer equipment is designed for personnel to operate on site, which not only allows on-site personnel to monitor in real-time but also allows operations such as information security offensive and defensive drills or novel technology research.

This research can not only be applied to the research and application of current personnel but it is also expected that the technology and expertise can be further extended to critical infrastructure to improve people's well-being.

5. Conclusion

At present, the development of critical infrastructure is quite rapid, and all countries attach great importance to the development of this facility. However, due to its huge interest in critical infrastructure, it is subject to cyber-attacks. Therefore, the testbed for simulating critical infrastructure is gradually gaining attention.

In order to protect the safety and operability of key infrastructure, this paper designs a virtual dam gate testbed based on PLC and Modbus/TCP protocol. By integrating the physical dam gate system into a virtual testbed, related settings and operations can be completed in a lightweight and flexible manner.

This paper also further realizes the dam gate testbed. In this testbed, research such as PLC packet sending and receiving and gate control, etc., is research with considerable research potential. It is also expected that the research in this paper can be further applied to large-scale critical infrastructure to promote the development of the people and the country.

Acknowledgements

This work was supported by the Water Resources Agency (WRA) under the Ministry of Economic Affairs (MOEA) and the National Science and Technology Council (NSTC) in Taiwan under contract numbers 111-2218-E-006-010-MBK.

References

1. W. Knowles, D. Prince, D. Hutchison, Jules Ferdinand Pagna Disso, K. Jones, “A survey of cyber security management in industrial control systems”, International Journal of Critical Infrastructure Protection, vol. 9, pp. 52-80, 2015.
2. A. Ghaleb, S. Zhioua, and A. Almulhem, “On PLC network security”, International Journal of Critical Infrastructure Protection, vol. 22, pp. 62-69, 2018
3. ESET Research, “Industroyer2: Industroyer reloaded This ICS-capable malware targets a Ukrainian energy company”, https://www.welivesecurity.com/2022/04/12/industroyer2-industroyer-reloaded/
4. J. Frauenschläger, J. Mottok, “ Security-Gateway for SCADA-Systems in Critical Infrastructures”, 2022 International Conference on Applied Electronics (AE), 2022.
5. J. Jarmakiewicz, K. Maślanka, K. Parobczak, “ Development of Cyber Security Testbed for Critical Infrastructure”, 2015 International Conference on Military Communications and Information Systems (ICMCIS), 2015.

Authors Introduction

Mr. Chen-Yu Lee

He was born in Taipei, Taiwan in 1998. 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 Communications, Navigation and Control Engineering, National Taiwan Ocean University, Taiwan in 2021. His interests are Cyber-Security , PLC and ICS Security.

Dr. I-Hsien Liu

He is a research fellow in the Taiwan Information Security Center @ National Cheng Kung University (TWISC@NCKU) and Department of Electrical Engineering, National Cheng Kung University, Taiwan. He obtained his PhD in 2015 in Computer and Communication Engineering from the National Cheng Kung University. His interests are Cyber-Security, Wireless Network, Group Communication and Reliable Transmission.

Dr. 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.

Dr. 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 wired and wireless network protocol design, network security, and network management. He is currently involved in funded research projects dealing with optical network, VANET, Cloud security and resource allocation, and IP QoS architectures. He is the director of Taiwan Information Security Center @ National Cheng Kung University. He serves on the editorial boards of the International Journal of Communication Systems.

Domain Name Infringement in Taiwan

Shih-Chin Lin
Department of Financial Law, Ming Chuan University,
No 250, Zhong Shan N. Rd., Sec. 5
Taipei 111, Taiwan

E-mail: lsc929@mail.mcu.edu.tw
web.mcu.edu.tw

Abstract

When a desired domain name has been registered by someone else or is similar to a trademark, proceeding with registration could infringe upon an existing domain name or a trademark. This is so-called domain name Infringement. This study thoroughly examined the research reports published by the TWNIC and the Intellectual Property Office (IPO) of Ministry of Economic Affairs (MoEA) or National Police Agency (NPA) of Ministry of the Interior (MOI) to gain an understanding of domain name".tw" Infringement in Taiwan.

Keywords: Domain Name, Domain Name Infringement, Cybersquatting, Typosquatting

1. Introduction

A domain name is a text version of an Internet protocol address and be used to identify Internet resources and obtain information. Compared with an Internet protocol address, which consists of a string of numbers, a domain name is represented by a set of phrases that are meaningful and easy to remember. Therefore, a domain name could be abused as a tool of trademark infringement or cybercrime. Here, it is about domain name infringement. His cases can be divided into the following two categories[1]: cybersquatting and typosquatting. Cybersquatting occurs when someone registers a domain name, which is the same to a trademark that is protected, with the intent to sell the name, prevent the trademark holder from gaining access to the name, or divert traffic. Typosquatting, similar to cybersquatting, is where the domain name registrant registers a variant of a famous trademark (e.g., "hogoboss").

Aiming to gain an understanding of situation of domain name ".tw" Infringement in Taiwan, this study will thoroughly examine the research reports published by Taiwan Network Information Center (TWNIC) and relevant authorities such as the Intellectual Property Office (IPO) of Ministry of Economic Affairs(MoEA) or National Police Agency (NPA) of Ministry of the Interior (MOI).

2. TWNIC's Taiwan Internet Report

As the organization that oversees domain name ".tw" management, the TWNIC releases the Internet Report every year. The latest report is available on its official website was issued in 2022[2]. In this report, a broad array of topics ranging from the Internet usage survey, web application service, the digital literacy and the digital divides are addressed. However, the report fails to describe current domain name disputes in Taiwan. Following a careful review of the Internet Report issues released before 2022[3], we discovered a lack of investigation into domain name infringement both domestically and abroad.

3. IPO Statistics and Annual Reports

The Taiwan IPO is the official agency responsible for trademarks and other Intellectual Property rights. The IPO releases trademark data, including application filings and registration, on a monthly basis. However, as of now, it does not provide domain name trademark infringement statistics[4]. The IPO also publishes an annual report[5], which comprises the following seven sections: up-to-date intellectual property filings, reviews and types of services, intellectual property laws, the development of service convenience in an electronically mediated environment, the application of types of industrial properties, international exchange and cross-strait cooperation, and the enforcement of intellectual property rights. No investigations have been conducted exploring domain name trademark infringement in Taiwan.

4. The NPA's Statistics List

As the highest law enforcement agency, the NPA maintains a statistics list[6] that is updated every 2 weeks[7]. The statistics list provides a comprehensive picture of cybercrime[8]. The list released by the NPA in the 20th week of 2022 indicated a yearly average of 12,000 to 15,000 cases of cybercrime in the period 2017–2021. A total of 14,997 cybercrime cases were reported in 2017, with cybercrime reaching a record high at this time. Major cybercrimes could be classified into 4 crime types, namely fraud, offenses against reputation and credit, offenses against computer users, and intellectual property infringement in Fig. 1; these cybercrime types accounted for 72%–77% of cybercrime in the period 2017–2021[9].

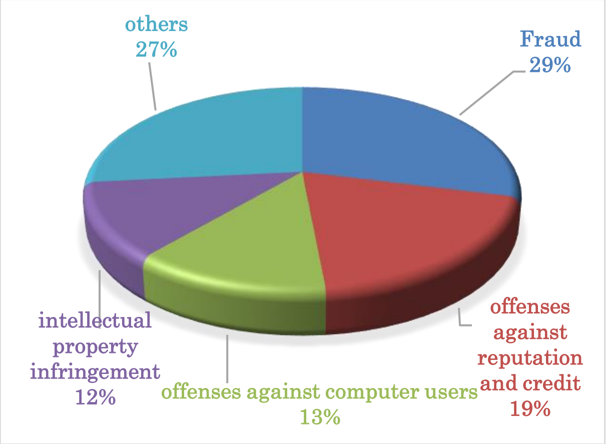


Fig. 1. Cybercrime types from 2017 to 2021

In the 48th week of 2021, the NPA published an updated statistics list titled Police Investigating Intellectual Property Crime for the Period January to October 2021. Intellectual property crime refers to any criminal act that involves manufacturing, selling, or trafficking counterfeit or pirated goods or the criminal theft of trade secrets[10]. A total of 4,946 intellectual property crime cases were reported in 2016, which was the highest rate for the period 2016–2020. Regarding criminal offenses, the statistics list only stated that "…in terms of IP infringement, the online infringement of intellectual property rights through over-the-top media services, digital cloud platforms, and online technology has risen, whereas cases of unauthorized documents, Digital Versatile Discs (DVDs), or other physical media have been declining over the years. A large proportion of cybercrime was related to BitTorrent programs, online image theft, and software piracy" [11].

With no mention of cybersquatting or typosquatting in the two aforementioned statistics lists released by the NPA, whether these two categories of domain name infringement didn't occur in Taiwan or whether the NPA hasn't paid sufficient attention to their emergence remains unclear. Regardless of the lack of data, at least, the cybercrime has attracted attention from concerned authorities.

5. Conclusion

From this study, we have learned that domain name ".tw" infringement has not yet gained its due attention in Taiwan. Thereby, we make the following recommendations in order to fight against the domain name ".tw" infringement. First, as the organization that oversees domain name ".tw" management, the TWNIC is advised to create a website for monitoring the use of the Taiwan country code's top-level domains (ccTLDs) ".tw". When any suspicious domain name activity is detected, the TWNIC must suspend the contract with the domain name owner and report the suspicious activity to investigative agencies for follow-up. Second, we urge police departments, which are the first line in cybercrime control, to publish statistics lists about domain-name-related crime. Police departments are also encouraged to work with schools, businesses, and government agencies to increase public awareness of domain name infringement and to implement a series of measures aimed at preventing and combating the abuse of domain name ".tw" in the future.

Acknowledgements

This work was supported by the National Science and Technology Council (NSTC) in Taiwan under contract numbers 111-2218-E-006-010-MBK.

References

1. Lin Shih-Chin, Lee Li-Chuan, Lu Ying-Chieh, Preventing Domain Name Infringement: An Introduction to Approaches Adopted by the United Nations, the European Union, and Taiwan, NTUT Journal of Intellectual Property Law and Management (SCOPUS), Vol. 11, No 2,; Study on legislative measures related to online IPR infringements, EUIPO, September 2018, p. 21.
2. Taiwan Internet Report 2022, TWNIC, July 2022.
3. Taiwan Internet Report from 2002 to 2021: https://twnic.tw/stat_n.php (accessed on October 2, 2022).
4. https://topic.tipo.gov.tw/trademarks-tw/np-581 201.html (accessed on October 2, 2022).
5. https://www.tipo.gov.tw/tw/cp-177-483190-adbf7-.html (accessed on October 2, 2022).
6. https://www.npa.gov.tw/ch/app/data/list?module=wg057&id=2218 (accessed on October 2, 2022).
7. This statistics list used to be released on a weekly basis before April 2021.
8. According to the definition provided by the NPA, cybercrime refers to "any crimes that are committed with a computer," NPA Statistics List (week 20, 2022), May 18, 2022, p. 3. Downloaded from: https://www.moi.gov.tw/cl.aspx?n=15288 (accessed on Octobe 2, 2022).
9. Supra [8], p. 2-3.
10. The Statistics List Released by the NPA (week 48, 2021), December 1, 2021, p. 1. Downloaded from: https://www.npa.gov.tw/ch/app/data/doc?module=wg057&detailNo=915389421547687936&type=s (accessed on October 2, 2022).
11. Supra [10], p. 2.

Authors Introduction

Dr. Shih-Chin Lin

He is a assistant professor in the Department of Financial Law, Ming Chuan University (MCU), Taiwan, and also a member in Taiwan Information Security Center @ National Cheng Kung University (TWISC@NCKU). He graduated from ChengChi University (NCCU), Taiwan, with B.S. in 1999 and Master degree in Criminal Law in 2003. He obtained second Master degree in Procedural Law in 2010 and his PhD in Juvenile Criminal Law in 2017, from the Aix-Marseille University (AMU), France. His research interests focus on the Criminal Justice for Juvenile Offenders, Cybercrim**e,** Cyber-Security and Rights of Children.

Device's Operation Tracking using Blockchain in Industrial Control System

Chien-Hsin 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

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

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
E-mail: {chwu, ihliu,}@cans.ee.ncku.edu.tw, jsli@mail.ncku.edu.tw, chufenli@gmail.com*
www.ncku.edu.tw, www.nfu.edu.tw

Abstract

Many producing, monitoring and controlling needs are met by using programmable logic controllers. But there is no effective mechanism to audit PLC behavior. So this research designed a mechanism based on Blockchain for the purpose of effectively recording the commands and response actions received by the PLC. Due to the characteristics of the blockchain, the integrity of the data is also guaranteed..

Keywords: Cyber Security, Blockchain, PLC, ICS Security

1. Introduction

A programmable logic controller (PLC) is composed of an I/O module, that is, an input-output module, sensors, and actuators. In a PLC control system, the sensor can detect the status of an on-site switching signal (such as a photoelectric switch, a material position switch, etc.) or an analog signal (such as on-site temperature, on-site pressure, etc.). The I/O module transfers the signal from the sensor to the PLC, which is processed by the CPU. The processed result will be converted into a control signal and sent by the output module to the actuator, which performs operations on the controlled object.

Due to the PLC's core advantages of being both simple and easy, durable and trusted, dependability is a very important factor when the machine may result in thousands to millions of dollars in losses. Control engineers and technical personnel need to know that they can rely on the PLC and thus perform simple troubleshooting quickly when an error occurs.

---

[*] Corresponding author's E-mail: jsli@mail.ncku.edu.tw

PLC was originally only used for automation control and development; its application scenario was extremely close and could hardly be shared with any third party outside the industry network equipment; however, with the rapid development of the Internet and the Internet of Things, and the emergence of intelligent hardware, industrial PLC has become more accessible to the public in recent years. If the instructions were recorded, it would be possible to know when the problems occurred and how to improve them. Therefore, this paper will use blockchain to establish a decentralized anti-tamper system to record PLC instructions, so that if the device fails, the record in the blockchain can be checked for the first time and the original instructions can be obtained.

2. Background

In this chapter, we mainly discuss the current attacks and risks of the ICS (Industrial Control System). At the 2016 Black Hat European Security Conference, Ali Abbasi, a graduate student, and Majid Hashemi, a Quarkslab R&D engineer, proposed that a malicious attacker can destroy and manipulate physical processes managed by a programmable logic controller (PLC) without being detected.[1]

2.1. *Ethernet*

PLC and Ethernet communication is based on the traditional Ethernet communication mechanism, using Ethernet and TCP/IP protocol as the basis for communication, in any case to provide absolute support for TCP/IP communication. In order to meet the real-time requirements of automation, the real-time communication channel is optimized based on PLC B network layer, which reduces the time occupied by communication and improves the performance of automatic data refresh. In the process of communication, there have been attacks on PLC.[2]

2.2. *PLC's operating mode*

The attacker modifies the operating state of the OT device. To obtain permission, PLC has many operating modes that can control the user's state and the API access of the controller, as well as the choice of physical mode. The attacker may try to modify the operating state of the PLC through various means. Some devices provide application programming interfaces (APIs) to facilitate information transfer between developers or machines. Attackers also have the opportunity to execute specific functions or malicious attack instructions through APIs.[3]

2.3. *Malicious PLC attack*

One technique uses data that is not necessarily part of the normal static or offline project files to weaponize PLCS and enable code execution during project connection or upload. Through this medium of attack, the target is not a PLC, such as the notorious Stuxnet[4] malware that secretly changes PLC logic to cause physical damage. Instead, they hope to use the PLC as a fulcrum to attack the engineers who program and diagnose it and gain deeper access to the OT network. Notably, all of the vulnerabilities they found were in the engineering workstation software, not in the PLC hardware. In most cases, they add, the vulnerabilities exist because the software completely trusts the data from the PLC without having to perform extensive security checks.

2.4. *Proof of work*

Here we use POW's proof-of-work algorithm. While processing transaction data, each node continuously performs hash calculation and obtains a hash value less than the network target value, which becomes the nonce golden number. The network target value is what we call the difficulty value, which will also be adjusted continuously with the operation of the whole blockchain system. When a miner in the whole network hashes the nonce, he will publish his packaged block. After receiving the block verification block, other nodes will agree that this block has been connected to the blockchain and continue to carry out the next block packaging and hash calculation.[5]

3. System Architecture

This section provides an illustration of the proposed method and general framework. Some adjustments have been made to the traditional network configuration and verification mechanism to make the blockchain running in PLC more compatible.

3.1. *Network model*

We designed the architecture shown in Fig.1. to implement block chain on PLC execution. When PLC

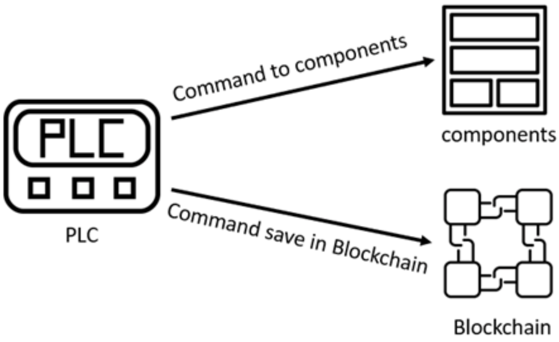


Fig. 1. System architecture.

sends out instructions, the instructions will be connected in the block chain at the same time.

The model describes the operation of a PLC chain. When the instruction is completed, it is simultaneously transmitted to the block chain program, and the instruction is converted into input information in the block chain. After function conversion, the attached value will be generated into a nonce variable. We put the contents of the Modbus packet in the field "nonce" for transmission on the blockchain. Combined with the unchangeable characteristics of blockchain, information in PLC can be effectively managed and the transaction process can be recorded. If the transaction log is sealed in a block, the log contents are not easily changed.

3.2. *Blockchain-based transaction process*

During the chaining process, there are four basic elements: the hash value of the previous block, the transaction information, the nonce variable, and the hash of the current block. The set up is shown in Fig.2.

Step 1: PLC outgoing instructions, including gate switch and so on.
Step 2: When the instructions are sent out, they are also sent to the blockchain to share information.
Step 3: Overload the hash code function in the block, wrapping the information and the time it was obtained as a Nonce variable.
Step 4: Set the difficulty to get the hash value.
Step 5: Hash through the blockchain

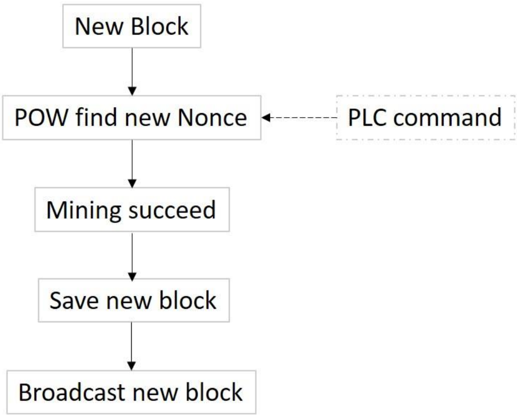


Fig. 2.Transaction Process

4. Conclusion

Due to the evolution of time and technology, the attacks on ICS are becoming more and more diversified, and ICS network security has become a hot research topic today.

To ensure the reliability of ICS packet information during transmission, this paper uses the immutable characteristics of an emerging technology called blockchain to record the instructions of PLC packet transmission. If the system is attacked in the future, it can find the untampered records to see where the problem occurred.

Acknowledgment

This work was supported by the Water Resources Agency (WRA) under the Ministry of Economic Affairs (MOEA) and the National Science and Technology Council (NSTC) in Taiwan under contract numbers 111-2218-E-006-010-MBK.

References

1. Abbasi A., & Hashemi M., "Ghost in the PLC: Designing an Undetectable Programmable Logic Controller Rootkit via Pin Control Attack", Black Hat Europe 2016(pp. 1-35).
2. Wikipedia contributors. (2022, November 4). Ethernet. In Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=Ethernet&oldid=1120015826
3. ATT&CK for ICS - Execution(2).(2021) https://ithelp.ithome.com.tw/m/articles/10275404
4. Wikipedia contributors. (2022, November 14). Stuxnet. In Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=Stuxnet&oldid=1121937222

5. Ayokomi L, & Sonya H, "Consensus Mechanism in Enterprise Blockchain", 2019 IEEE International Conference on Intelligence and Security Informatics (ISI)

Authors Introduction

Ms. Chien-Hsin Wu

She was born in Tainan, Taiwan in 1999. She is acquiring the master's degree in Department of Electrical Engineering/Institute of Computer and Communication Engineering, National Cheng Kung University in Taiwan. She received her B.S. degree from the Department of Communication Engineering, National Taipei University, Taiwan in 2021. Her interests are Cyber Security.

Dr. I-Hsien Liu

He is a research fellow in the Taiwan Information Security Center @ National Cheng Kung University (TWISC@NCKU) and Department of Electrical Engineering, National Cheng Kung University, Taiwan. He obtained his PhD in 2015 in Computer and Communication Engineering from the National Cheng Kung University. His interests are Cyber-Security, Wireless Network, Group Communication and Reliable Transmission.

Dr. 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 wired and wireless network protocol design, network security, and network management. He is currently involved in funded research projects dealing with optical network, VANET, Cloud security and resource allocation, and IP QoS architectures. He is the director of Taiwan Information Security Center @ National Cheng Kung University. He serves on the editorial boards of the International Journal of Communication Systems.

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 Journal of Operational Research, Journal of System and Software, International Journal of Information and Management Sciences, Asia Journal of Management and Humanity Sciences, and others.

Strengthen the Security of the Industrial Control System using SDN Technology

Min-Wei Huang
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

Hsin-Yu Lai
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-Huan Lee
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
E-mail: mwhuang@cans.ee.ncku.edu.tw, ihliu@cans.ee.ncku.edu.tw, hylai@cans.ee.ncku.edu.tw,
*mhlee@cans.ee.ncku.edu.tw, jsli@mail.ncku.edu.tw**
www.ncku.edu.tw

Abstract

In the field of OT, most of the network architectures operated in the way of isolation from internal and external networks. Only firewalls are installed on the external network without any protection measures for the internal network. In this paper, we leverage a Software-defined network (SDN) with an industrial control system (ICS), so controllers can manage the equipment and keep track of each switch and its connection with the programmable logic controller (PLC) in the ICS. Additionally, only the critical flows can be allowed by adding flow entries. So the transmission between the PLC and Human Machine Interface (HMI) can be protected. The transmission quality of the ICS and its availability can be improved.

Keywords: ICS, Cybersecurity, OT security, Software-defined network.

1. Introduction

The network of equipment in the industrial control system gradually operates and develops in the form of the Internet of Things. Operating technicians can access these equipment through the network, but it also means that more and more information is transmitted between these devices through the network, which undoubtedly increases the chance of being attacked. In the current operational technology field, most network structures are operated in the way of internal and external isolation. Only a firewall is connected to the external end, and no protections are taken for the internal network. Once the attacker invaded, the attack will be out of control, and let system weak. The impact even extends to the country's critical infrastructure. In our research, we use a SDN controller to manage the equipment and keep track of each network between the system environment and its connected PLC. And create key flow entry rules to protect the critical flow of communication between in ICS equipment to maintain the transmission quality of industrial network.

2. Background

Most of the critical infrastructures are constructed by industrial control systems. Besides lots of information security issues in the information technology (IT) field, accidents of ICS are also frequent currently. Companies and countries need to think about how to build and use security testbeds to defend against attacks [1], and the establishment of a testbed for water resources will be described in 2.1. And there are also many scholars using centralized management structures in SDN for protecting industrial networks [2] will be described in 2.2

2.1. *Water resources security testbed*

Due to the cost and other various considerations, there is not much large-scale security testbed in the world at present. Most of them are built on power systems [1]. The water resources testbed is much fewer globally. A small-scale testbed was used for simulates dam pumping and distribution in [3]. Moreover, the more well-known SWaT [4] is used for the security research of the ICS for water treatment and for operator training.

2.2. *Software-defined network*

Software-defined network (SDN) architecture is widely used nowadays due to that it can make users get rid of the limitations of hardware in traditional networks. The SDN separates the control plane and the data plane of the traditional switch. Researchers can design the network more flexibly and control network traffic through the centralized management controller to solve the problem caused by a generally distributed network. The data plane has a network topology composed of network devices. The devices at this layer are responsible for packet forwarding. The forwarding rules are determined by the rule table, which can be set and modified by the control layer.

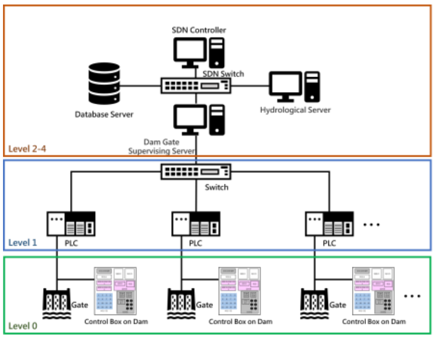


Fig. 1. Water Critical Infrastructure Architecture with SDN

3. System Design

In this section, our architecture and method are described. To enhance the transmission quality of ICS, we combined the SDN technology and the water-critical infrastructure architecture.

3.1. *System architecture*

In the water infrastructure information security testbed built by our team, we replaced the Ethernet switch of the dam gate operating system with another switch that supports the OpenFlow protocol. That made the switch to be the southbound interface protocol in the SDN. The water-based critical infrastructure information security testbed can be combined with SDN. The establishment of the Openflow instance enables the switch to enable the functionality of the Southbound interface protocol with OpenFlow protocol. The critical infrastructure structure for water resources is shown in Fig. 1.

3.2. *Flow entry rules*

Between the controller and the switch, a flow table that matches the environment is set, and flow entries are defined so that the OpenFlow switch can forward messages and perform actions according to the rules of the controller. The priority is set according to the individual flow entry. When the switch executes according to the flow entry in the flow table, the action is

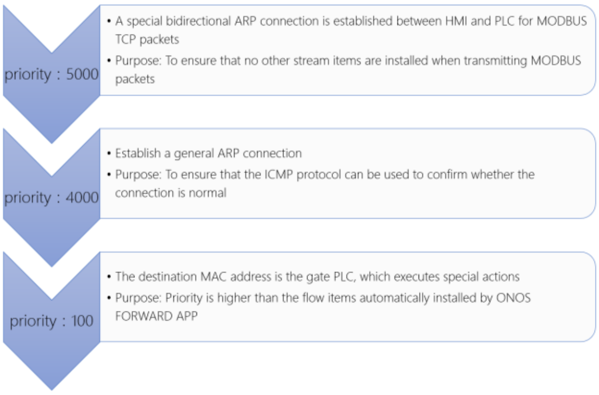
Table 1. Attack Classification and Process.

Classification	Name	Tools
Reconnaissance Attacks	Network Scanning	Nmap
Reconnaissance Attacks	ARP Spoofing	dsniff - arpspoof Wireshark
Command Injection Attacks	MODBUS TCP Read and Write Memory	Modbus TCP test software Wireshark
Denial of Service Attacks	ICMP Flood	hping
Denial of Service Attacks	ICMP Advance Flood	hping

executed according to the set priority, and the message protection of the key flow entry is achieved in the communication of the programmable logic controller. The mechanism is shown in Fig. 2.

4. Experiment and Result

In our experiment, we use Hewlett-Packard 5130 Switch as the original switch and the Openflow-supported switch. To make the connection be end-to-end, we first limited the source MAC, destination MAC address, and the port to Openflow switch. Then, we respectively write the bidirectional flow entry and increases the priority. Besides, flow entry rules between HMI and PLC are set to the highest priority. As for the special settings of flow entries, in this experiment, the packets whose target is the MAC address of the dam gate programmable logic controller and whose priority is higher than that of the general forwarding flow entry are dropped to achieve the effect of whitelisting. That makes sure that only the connection between HMI and PLC can be successful. Additionally, the experiment will be divided into three scenarios: the original switch, the Openflow switch with normal flow rule, and the Openflow switch with key flow entries rules for comparison to do the attack experiment.



- A special bidirectional ARP connection is established between HMI and PLC for MODBUS TCP packets
- Purpose: To ensure that no other stream items are installed when transmitting MODBUS packets

- Establish a general ARP connection
- Purpose: To ensure that the ICMP protocol can be used to confirm whether the connection is normal

- The destination MAC address is the gate PLC, which executes special actions
- Purpose: Priority is higher than the flow items automatically installed by ONOS FORWARD APP

Fig. 2. Flow Entry Rule Sample.

The attack process shown in Table 1. takes the attacker's perspective as the starting point. First, the attacker examines and judges the devices in the network through reconnaissance attacks to identify the manufacturer of the device and open TCP ports. Launch a inject attack through the response of ARP Spoofing, so that the attacker can cut off the communication between the two ends and see the master-slave relationship, and further understand the environment and available equipment. Then, attack PLC by a command injection that is using the MODBUS test software. All the data in the holding register will be changed and operated according to incorrect commands. Finally, a flood attack is carried out, which paralyzes the operation of the equipment on the system. Furthermore, it will destroy the process of the system.

In our research, we take three different switches mentioned earlier to experiment and get the result that the SDN flow entry rules set according to the process in Fig. 2 can make the MAC of gate PLC can under the protection. It prevents the possibility of being attacked by using ARP connections with the gate PLC.

Once we avoid all the above-mentioned reconnaissance attacks and subsequent attacks, the service quality of ICS network transmission of critical water infrastructure facilities can be ensured not be affected.

5. Conclusions

Most of the weaknesses existing in the current ICS environment are due to too much trust in the firewall, resulting in no protections in the intranet. And that weak network will create vulnerability for attackers. However, in the current OT environment, most of the equipment in the ICS is used for many years and has no chance to be replaced due to the cost and difficulties of deployment. Hence, we take use of setting the flow entries rule, prioritizing all devices traffic to let the system communication follow the flow entries. Finally, give critical traffic the highest priority and make a whitelistlike mechanism to protect PLCs for critical infrastructure gates.

Acknowledgments

This work was supported by the National Science and Technology Council (NSTC) in Taiwan under contract numbers 111-2218-E-006-010-MBK and 111-2218-E-006-079- .

References

1. K. Barnes, B. Johnson, "National SCADA Test Bed Substation Automation Evaluation Report", 2009.
2. R. D. Lallo, F. Griscioli, G. Lospoto, H. Mostafaei, M. Pizzonia, M. Rimondini, "Leveraging SDN to monitor critical infrastructure networks in a smarter way", 2017 IFIP/IEEE International Symposium on Integrated Network Management, Lisbon, Portugal, 8-12 May, 2017.
3. L. Faramondi, F. Flammini, S. Guarino, R. Setola, "A Hardware-in-the-Loop Water Distribution Testbed Dataset for Cyber-Physical Security Testing", IEEE Access, vol. 9, pp. 122385-122396, 2021.
4. A. P. Mathur, 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 Apr., 2016.

Authors Introduction

Ms. Min-Wei Huang

She is acquiring the master's degree in Department of Electrical Engineering / Institute of Computer and Communication Engineering, National Cheng Kung University in Taiwan. She received her B.S. degree from the Department of Communications, Navigation and Control Engineering, National Ocean University, Taiwan in 2021. Her interests are Cyber-Security and Software-Defined Network.

Dr. I-Hsien Liu

He is a research fellow in the Taiwan Information Security Center @ National Cheng Kung University (TWISC@NCKU) and Department of Electrical Engineering, National Cheng Kung University, Taiwan. He obtained his PhD in 2015 in Computer and Communication Engineering from the National Cheng Kung University. His interests are Cyber-Security, Wireless Network, Group Communication and Reliable Transmission.

Mr. Hsin-Yu Lai

He got the M.S. degree in National Cheng Kung University in Taiwan. He also received his B.S. degree from the Department of Electrical Engineering, National Chung Cheng University, Taiwan in 2019. His interests are Cyber-Security and Software-Defined Network.

Mr. Meng-Huan Lee

He is studying for his master's degree in Department of Electrical Engineering / Institute of Computer and Communication Engineering, National Cheng Kung University in Taiwan. He graduated from the Department of Communications Engeineering, National Chung Cheng University, Taiwan in 2021. His interests are Cyber-Security and Cellular Network Security.

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 wired and wireless network protocol design, network security, and network management. He is the director of Taiwan Information Security Center @ National Cheng Kung University. He serves on the editorial boards of the International Journal of Communication Systems.

Fake Base Station Threats in 5G Non-Public Networks

Meng-Huan Lee
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.

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.
*E-mail: mhlee@cans.ee.ncku.edu.tw, ihliu@cans.ee.ncku.edu.tw, jsli@mail.ncku.edu.tw**
www.ncku.edu.tw

Abstract

With 5G technology, traditional industrial and business equipment can now be connected wirelessly in a non-public network separated from public mobile services. Benefit from features such as high bandwidth and massive machine-type communications, while being able to control their own private 5G networks. But fake base stations used by law enforcement and hackers may collect private information and cause disruptions in cell services, thus compromising the security. In this research, we will analyze existing attack methods and detection mechanisms. And look at how those threats can affect the devices and operations in 5G non-public network.

Keywords: 5G, Non-public network, Fake base station, Industrial Internet

1. Introduction

Driven by Industry 4.0 and Industrial Internet initiatives, 5G has gained popularity for being a key enabler of such use cases. 5G provides higher bandwidth and more reliable communication compared to previous generations. Allowing more devices, like autonomous robots in a factory, to cover larger areas. These kinds of industrial scenarios are the most common use of 5G NPNs [1]. But more 5G usage may expose more attack surfaces compared to traditional wired or wireless LAN technology.

One of the big concerns regarding the security of mobile networking is in the Radio Access Network (RAN). And fake base stations are the most popular radio-layer attacks, being known for their disruptive capability since the 2G era. With the rise of low-cost Software-Defined Radio (SDR) and open-sourced radio software, the possibility of fake base station attacks is increasing. Although security standards have improved over the years, these risks are still relevant to this day. For companies and organizations to safely deploy 5G in crucial operations such as industrial facilities or utility infrastructure, stakeholders need to understand such attacks. So, in our research, we categorized several ways of fake base station attacks, and identify their threats to 5G NPNs.

---

[*] Corresponding author's E-mail: jsli@mail.ncku.edu.tw

2. Backgrounds

We briefly introduce the basics of mobile networks, with a focus on 5G NPN in an industrial scenario. And the security measures in the 5G System that are related to fake base stations.

2.1. *5G Non-Public Network*

5G Non-Public Networks (also called private mobile networks) are purpose-built, independent networks. In contrast to Public Land Mobile Networks (PLMNs) that offer mobile network services to public subscribers, NPNs are intended for the exclusive use of an enterprise or an organization. According to 3GPP Release-16 [2], NPNs are categorized into SNPNs and PNI-NPNs.

NPNs give organizations control over the quality of their own connectivity although it could be beneficial to have support from a third-party supplier or Mobile Network Operator (MNO) to help configure, optimize, and operationally manage the NPN. An NPN can be isolated from external networks and reside behind corporate firewalls. This is the most common way for companies and organizations to deploy 5G in industrial scenarios.

2.2. *5G System, Authentication and registration*

A 5G system mainly consists of three parts [3]:

- *User Equipment (UE).* The UE (essentially the modem) stores a permanent identifier and permanent key on a Universal Subscriber Identity Module (USIM) card. With these credentials, user and network establish mutual authentication. Three identifiers are important: the permanent identifier SUPI (4G: IMSI), the concealed identifier SUCI, and the temporary identifier 5G-GUTI.
- *Base Stations.* Base Stations create the wireless network. They act as access points for user equipment to attach to the Radio Access Network (RAN), thus connecting to the mobile network.
- *Core Network.* The back-end core network performs all management tasks and traffic routing.

In 4G, the base station and core network are called eNB (Evolved Node B) and EPC (Evolved Packet Core). In 5G system, they are called gNB (Next Generation Node B) and NGC (Next Generation Core). For a UE to register to the mobile network, the Authentication and Key Agreement (AKA) procedure are performed between the UE and the Core via BS. But some user information may still be transmitted in plain text before the AKA finished the authentication, like IMSI in 4G or SUCI in 5G. Because encryption is only activated *after* both parties agree on a session key. As shown in Fig.1. This is one of the important attack vectors for FBS.

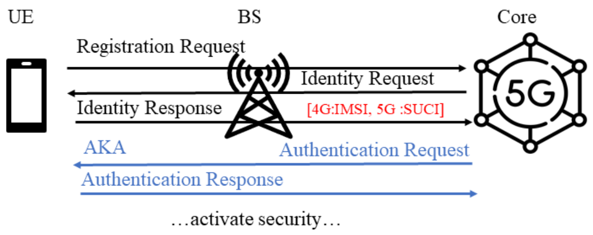


Fig.1. Part of the registration process and the Authentication and Key Agreement (AKA) procedure.

3. Fake base station attacks

Fake base stations are malicious radio devices that disguise themselves as legitimate ones to attract nearby signals. Fig.2 illustrates an FBS attack in NPN, the FBS can try to trick UEs to connect to them and/or listen to the messages broadcasted by legitimate BS to obtain information about the network. They are mainly used for two purposes: 1) to identify or track users, 2) to perform Denial of Services.

User tracking is the most common use of fake base stations. By exploiting the vulnerability in the authentication procedure when the messages aren't yet encrypted (see 2.2), attackers obtain identifiers and track

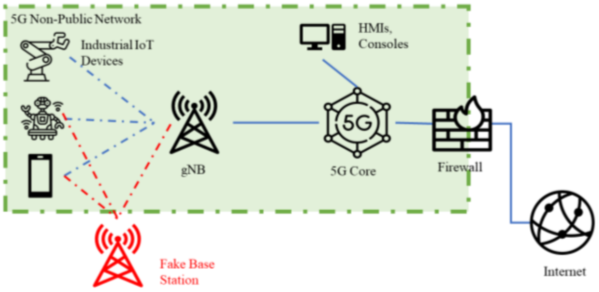


Fig. 2. Fake base station attack in NPN

the target of interest. “SUCI(IMSI)-catcher” is the collective name given to devices used to eavesdrop and track mobile network subscribers [4]. In public networks, the target is usually a person’s phone or personal device. There are several reports of such devices (branded “Stingray”) being used by law enforcement or appearing in cities [4], [5]. Which compromised users’ privacy and leaked their locations. Another use of FBSs is to trick UEs into downgrading their connections to older generations (2G/3G/4G) by sending them reject messages during the registration phase [6].

4. Fake base stations threats in NPN

Here we identify how the above-mentioned attacks can affect the NPN in an industrial scenario. In this scenario, the UEs in the network are mostly industrial devices such as robotic arms, sensors, or mobile robots instead of usual smartphones. So, the goals of the attack and the impact would be different.

4.1. *Impacts of fake base stations attacks in NPN*

In industrial NPN, low-latency and reliable services are key. The threats posed by FBSs in an NPN are mostly similar to public networks, but the impact can be more severe.

- *SUCI(IMSI)-Catchers:*
  Just like public networks, attackers can use FBS to track the movement of a particular UE in an NPN. This is especially dangerous in an industrial setting. As it can be used to track and monitor the activities of personnel or industrial devices. Though tracking a stationary target (like a static robotic arm) may not have much use, tracking Autonomous Mobile Robots could let the attackers learn more about the physical environment like the factories’ layout and route of the robots. And if the attacker can track certain employees via “SUCI-Catching” their company phones, it could pose big security risks.
- *Downgrade or Denial of Services:*
  By downgrading the connection of UEs in an NPN, attackers can exploit the vulnerabilities of older communication standards and redirect them to an unsafe network controlled by the attackers. Or they can force UEs to temporarily lose mobile service, causing higher latencies or unreliable connections. Which leads to disruption in production lines or other operations.

The different types of FBS attacks and their key aspects are shown in Table 1.

4.2. *Countermeasures*

In order for NPN operators to combat these threats, there are some existing measures from public networks that can be used in NPN. One way is to monitor nearby BSs and check the presence of unknown or malicious BSs. So operators can then respond to possible FBSs or warn the users of incoming dangers. This can be done by special apps[7] or network-side detection mechanisms.

By measuring physical parameters like signal strengths, or detecting abnormal behaviors like duplicate requests or registration procedures that are out-of-order[8]. And rate-limit attach requests like some MNOs do in public networks [3].

4.3. *Challenges*

While 3GPP is aware of the FBS attacks and putting effort to remedy the issues in newer standards, lots of mobile services may not keep up with those updates. One example is the previously mentioned downgrade attacks, where it’s still common for 4G and 5G devices to coexist in a network. New security standards may not be

Table 1. Different types of fake base station attacks.

Attack type	**Attack vectors**	**Result**	**Threats to NPNs**
SUCI(IMSI)-Catchers [3], [4], [5]	Collect and track identifiers. Listen to paging messages.	Tracking and locating of specific users in an area. Compromising user privacy.	Keep track of static devices and moving robots. Track important employee’s phones.
Downgrade or DoS [6]	Faking reject messages.	Redirect users to older standards or unsafe networks. Causing UEs to lose connection, lead to DoS.	Unreliable connections, higher latencies. Gaining access to devices via unsafe networks.

implemented for compatibility issues, or simply bypassed by attacking legacy devices.

Though detection-based countermeasures might be a good option, it requires additional apps or mechanisms to be put into UEs or Core. Which are added costs for NPN operators and could add unwanted latencies. Existing methods also generally aims for public network, not NPN in an industrial scenario.

And as S.Park et.al's research shows, detection apps have their own limitations[7], so one should not solely rely on apps.

5. Conclusions

In this reasearch, we analyzed the threats of fake base station attacks in a 5G Non-public network. We identified the two main attack vectors, user tracking and Denial of Services, and examined their effects in an industrial scenario. We also discussed some existing countermeasures and identified the challenges that NPN operators may face. As 5G NPNs starting to be widely deployed, further research is needed to understand the threats and develop more effective countermeasures for such scenarios. We hope to raise awareness and advise operators and organizations to take precautions when configuring their NPNs.

Acknowledgements

This work was supported by the National Science and Technology Council (NSTC) in Taiwan under contract number 111 - 2221 - E - 006 - 079 – .

References

1. A. Aijaz, "Private 5G: The Future of Industrial Wireless," in IEEE Industrial Electronics Magazine, vol. 14, no. 4, pp. 136-145, Dec. 2020, doi: 10.1109/MIE.2020.3004975.
2. ETSI, "System architecture for the 5G System (5GS); Release 16", 3GPP TS 23.501
3. M. Chlosta, D. Rupprecht, C. Pöpper, T. Holz, "5G SUCI-catchers: still catching them all?" the 14th ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec '21). New York, NY, USA, Jun. 28-Jul. 2, 2021.
4. C. Cullen, B. Bureau, "Someone is spying on cellphones in the nation's capital", CBC News, April 2017
5. A. Ramirez, "ICE Records Confirm that Immigration Enforcement Agencies are Using Invasive Cell Phone Surveillance Devices", ACLU, 2020.
6. H. Lin, "LTE REDIRECTION: Forcing Targeted LTE Cell-phone into Unsafe Network", Hack in the Box Security Conference, Amsterdam, Netherlands, May 2016.
7. S. Park, A. Shaik, R. Borgaonkar, A. Martin, Jean-Pierre Seifert, "White-Stingray: Evaluating IMSI Catchers Detection Applications." In Workshop on Offensive Technologies (WOOT). USENIX Association, Aug. 14-15, 2017
8. M. Echeverria, Z. Ahmed, B. Wang, M. Fareed Arif, Syed R. Hussain, O. Chowdhury, "PHOENIX: Device-Centric Cellular Network Protocol Monitoring using Runtime Verification", Jan 2021.

Authors Introduction

Mr. Meng-Huan Lee

He is studying for his master's degree in Department of Electrical Engineering / Institute of Computer and Communication Engineering, National Cheng Kung University. He graduated from the Department of Communications Engeineering, National Chung Cheng University, Taiwan in 2021. His interests are Cyber-Security and Cellular Network Security.

Dr. I-Hsien Liu

He is a research fellow in the Taiwan Information Security Center @ National Cheng Kung University (TWISC@NCKU) and Department of Electrical Engineering, National Cheng Kung University, Taiwan. He obtained his PhD in 2015 in Computer and Communication Engineering from the National Cheng Kung University. His interests are Cyber-Security, Wireless Network, Group Communication and Reliable Transmission.

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 wired and wireless network protocol design, network security, and network management. He is the director of Taiwan Information Security Center @ National Cheng Kung University. He serves on the editorial boards of the International Journal of Communication Systems.

Cyber-Physical Security Testbed for River Basin Gate Control System

Meng-Wei Chang
Department of Electrical Engineering / Institute of Computer and Communication Engineering,
National Cheng Kung University,
No.1 Daxue 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 Daxue 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 Daxue Rd., East Dist., Tainan City, 701401, Taiwan
E-mail: mwchang@cans.ee.ncku.edu.tw, ihlu@cans.ee.ncku.edu.tw, jsli@mail.ncku.edu.tw
www.ncku.edu.tw

Abstract

Due to the flourishing development of critical infrastructures in recent years, increasing importance has been attached to the security of the Cyber-Physical System (CPS) of the infrastructures. Machine learning technology nowadays is evolving rapidly, and is widely implemented in detecting or preventing such attacks. As a result, this research constructs a Testbed to collect relevant data sets to support machine learning requirements, such as training models and analyzing attacks, etc.

Keywords: Testbed, CPS, Critical infrastructure, Dataset, Machine learning.

1. Introduction

Critical infrastructures have improved quality of our life during years of development. After the Industry 4.0 concept been introduced in 2011 [1], a secured CPS [2] has been the main goal of various fields of industry. A CPS are integration of computation, networking, and physical process. With CPS, we can supervise both the physical process and the network traffic of the system, even improve the performance and resource allocation of the system.

Nevertheless, there are many vulnerabilities that exist in the CPS of the infrastructure which may put people in great danger. In the case of dam facilities, there are failures and attack events happening to the dams every year. Some of these failures happened due to the anomaly inflow which can be caused by extreme weather, such as the Loas Dam collapse [3] and the Sandford Dam failure [4] in 2018 were both caused by heavy rains; on the other hand, the cyberattack toward the Bowman Avenue Dam in 2013 [5] had revealed the potential crisis that hackers could do to the system.

Thankfully, the maturity of machine learning and neural networks brings different kinds of detection models to prevent such threats. For instance, J. Goh et al. [6] take advantage of the Recurrent Neural Network (RNN) to train models with datasets to detect cyberattacks. Another example is the anomaly detection for water treatment by J. Inoue et al. [7]

---

[*] Corresponding author’s E-mail: jsli@mail.ncku.edu.tw

Though the two aforementioned methods both focus on water treatment, the security of dam CPS can no longer be ignored after the tragedies happened around the world. As the result, a testbed that contains both physical and network aspects of data in a dam scenario is needed and became the main goal of this study.

2. Research Background

Our study refers to the concepts of the Industrial Control System (ICS) [8] and ISA-95 [9] framework. Building our system by following the structures that are currently running in the industry makes our data more convincing. Hence, we are taking a brief look at these frameworks before diving into our testbed.

2.1. *Industrial Control System*

An ICS is a set of devices, systems, and networks that operate or automate industrial processes. ICSs usually include some core components, such as Human Machine Interfaces (HMI) [10] and Supervisory Control and Data Acquisition (SCADA) [11] systems that monitor and lead the whole operation; some Master/Remote Terminal Units (MTU/RTU) that send commands; some Programmable Logic Controllers (PLC) [12] that execute the commands by controlling the physical devices; and a Data Historian which record all historical log data of the operation. Despite there are many different ICSs for particular use cases, all of them are managed to control, monitor, and merge Informational Technology (IT) and Operational Technology (OT) aspects of the system.

2.2. *ISA-95*

ISA-95 [9] is an international standard from the International Society of Automation (ISA) that defines the interface between enterprise systems and ICSs. This testbed will be covering the bottom four levels in the five levels of the standard, which includes signals, PLCs, HMIs, and the database.

When collecting the data of our dam testbed, it is important to comply with the proper time scale at each level since time-related features of the data are crucial elements for detecting attacks or anomaly status of the dam system.

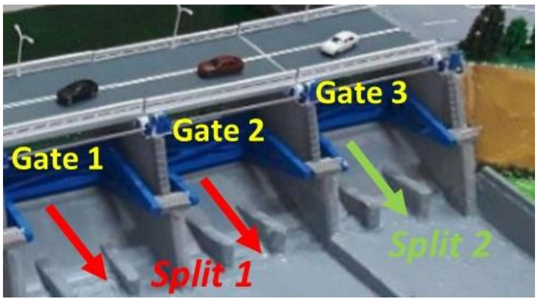


Fig. 1. The three gateways dam scenario of the testbed.

3. Testbed Architecture

The goal of our testbed is to make the data we collect similar to an actual dam CPS, so we build our testbed imitating a retired dam in Taiwan. The details of the architecture will be described in this section.

3.1. *Dam Environment*

The architecture of the testbed is shown in Fig.1, we simulate three gateways of the dam with PLCs, which control the water discharge of the dam.

Water level of the dam is affected by the upstream inflow, the rainfall, and the downstream water discharge. The downstream of the dam will be divided into two splits: *Split 1* controlled by the two gateways representing the main discharge; and *Split 2* is defined as a regular outflow considering some dams contain gateways arranged for the intake of other purposes, such as domestic water supply, agricultural water supply, etc.

3.2. *CPS framework*

As shown in Fig.2, the HMI will not only continuously ask the connected PLCs for the status of the gateways, but also be able to send commands to PLCs through Modbus TCP [13] packets and transmit the log data to the connected SQL server to record the operations.

The control panel is an entity for on-site dam control. After all, in a real dam operation, workers usually interact with the facilities by pressing buttons of control panels on the spot due to the security concerns.

On the other hand, the actual DI/DO status of the gateways will be simulated and recorded, the reason behind this design is that Man-in-the-middle (MITM)

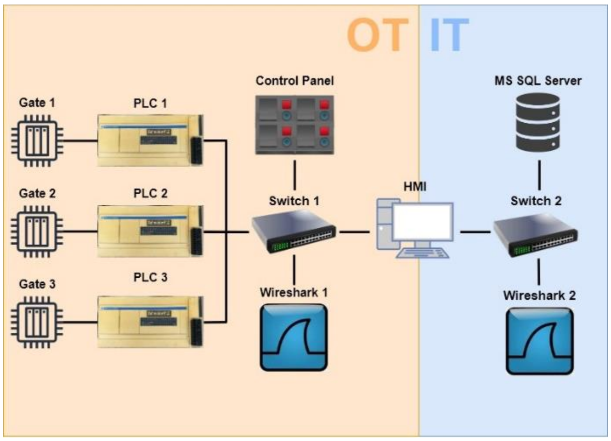


Fig. 2. The CPS framework of the testbed, including IT and OT aspects of the system.

attacks [14] can happen between PLCs and HMI, which makes the SQL server receive fake data.

Finally, the transmitted packets within OT and IT networks will also be recorded separately through Wireshark [15] hosts.

The operations of the three gates and HMI will follow the historical data of a retired dam system as input, the input data contain normal operations of the dam and some non-normal events such as abnormally high rainfall, abnormal gateway operation, abnormally upstream inflow increase, etc. We are also planning to add more attack situations on the testbed to increase the variety of the generated dataset.

3.3. *Data Structure*

Similar to the most used water treatment dataset SWaT [16], the structure of the data will be split into two parts: physical and network.

The physical part of data comes in Comma-Separated Values (.csv) files, which contain a few attributes of each status of dam facilities, such as the open degree of gateways, the inflow values, water level, etc., and the timestamps.

For the network part, all log data will be collected in the form of Packet Capture (.pcap) files through the Wireshark application.

4. Testbed Implementation

To implement the aforementioned testbed architecture in Fig.2, we utilize the devices of a retire dam system, including the Schneider TWDLCAE40DRF PLCs [17] and the HMI and database in the form of Windows 7 virtual machines on a Windows 10 PC. We have also arranged two extra Windows 10 PCs for the Wireshark packets recording, the Cisco IE 4000 [18] for *Switch 1*, and the HP 1810-8G [19] for *Switch 2*; as for the gates of the dam, the Arduino MEGA 2560 boards [20] are utilized to present their DI/DO values.

5. Conclusions and Future Works

In this study, a testbed CPS that covers both OT and IT aspects of a dam environment is established. With all the components of the retired dam system, the testbed is highly associated with the real dam infrastructures.

In the future, we are heading to fully publish our dataset based on historical data, and working on generating our own input data with the features of historical data, or even adding more physical and cyber attack scenarios to the testbed.

After all, the testbed must be continuously improved in terms of reliability and variety to make a greater contribution to the security of critical infrastructures.

Acknowledgements

This work was supported by the Water Resources Agency (WRA) under the Ministry of Economic Affairs (MOEA), the National Science and Technology Council (NSTC) in Taiwan, under contract numbers NSTC 111-2218-E-006-010-MBK.

References

1. Lasi, H., Fettke, P., Kemper, HG. *et al.* “Industry 4.0”, Business and Information Systems Engineering, Vol. 6, pp. 151-152, 2015.
2. E. A. Lee, “CPS foundations”, Design Automation Conference, pp. 737-742, 2010.
3. BBC NEWS, “Loas dam collapse: Many feared dead as floods hit villages”, https://www.bbc.com/news/world-asia-44935495
4. WECT, “Sanford Dam fails due to rising waters from Florence”, https://www.wect.com/story/39099087/sanford-dam-fails-due-to-rising-waters-from-florence/
5. GARY COHEN, “Throwback Attack: How the modest Bowman Avenue Dam became the target of Iranian hackers”, https://www.industrialcybersecuritypulse.com/facilities/throwback-attack-how-the-modest-bowman-avenue-dam-became-the-target-of-iranian-hackers/

6. J. Goh, S. Adupu, M. Tan and Z. S. Lee, "Anomaly Detection in Cyber Physical Systems Using Recurrent Neural Networks", 2017 IEEE 18th International Symposium on High Assurance Systems Engineering (HASE), pp. 140-145, 2017.
7. J. Inoue, Y. Yamagata, Y. Chen, C. M. Poskitt and J. Sun "Anomaly Detecion for a Water treatment System Using Unsupervised Machine Learning", 2017 IEEE International Conference on Data Mining Workshops (ICDMW), pp. 1058-1065, 2017.
8. K. Stouffer, S. Lightman, V. Pillitteri, M. Abrams and A. Hahn "Guide to Industrial Control Systems (ICS) Security", NIST Special Publication 800-82, R2, 2014.
9. ISA, "ISA95, Enterprise-Control System Integration", https://www.isa.org/standards-and-publications/isa-standards/isa-standards-committees/isa95
10. 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), pp. 289-296, 2018.
11. A. Daneels and W. Salter, "WHAT IS SCADA?", International Conference on Accelerator and Large Experimental Physics Control System, Trieste, Italy, 1999.
12. E. R. Alphonsus, M. O. Abdullah "A review on the applications of programmable logic controllers (PLCs)", Renewable and Sustainable Rnergy Reviews, Vol. 60, pp. 1185-1205, 2016.
13. Andy Swales, "Open Modbus/TCP Specification", Schneider Electric, Vol. 29, pp.3-19, 1999.
14. Avijit Mallik, "Man-in-the-middle-attack: Understanding in simple words", Cyberspace: Jurnal Pendidikan Teknologi Informasi, Vol. 2, pp.109-134, 2018.
15. Gerald Combs, "Wireshark", https://www.wireshark.org
16. iTrust, "Secure Water Treatment (SWaT) Dataset", https://www.isa.org/standards-and-publications/isa-standards/isa-standards-committees/isa95
17. Schneider Electic, "TWDLCAE40DRF", https://www.se.com/ww/en/product/TWDLCAE40DRF/compact-plc-base-twido-100-240-v-ac-supply-24-i-24-v-dc-16-o/
18. Cisco, "Industrial Ethernet 4000 Series Switches", https://www.cisco.com/c/en/us/products/collateral/switches/industrial-ethernet-4000-series-switches/datasheet-c78-733058.html
19. HPE, "HPE OfficeConnect 1810 Switch Series", https://support.hpe.com/hpesc/public/docDisplay?docId=emr_na-c02500478
20. Arduino, "Arduino MEGA 2560 Rev3", https://store.arduino.cc/products/arduino-mega-2560-rev3

Authors Introduction

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.

Dr. I-Hsien Liu

He is a research fellow in the Taiwan Information Security Center @ National Cheng Kung University (TWISC@NCKU) and Department of Electrical Engineering, National Cheng Kung University, Taiwan. He obtained his PhD in 2015 in Computer and Communication Engineering from the National Cheng Kung University. His interests are Cyber-Security, Wireless Network, Group Communication and Reliable Transmission.

Prof. Jung-Shain 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 wired and wireless network protocol design, network security, and network management. He is the director of Taiwan Information Security Center @ National Cheng Kung University. He serves on the editorial boards of the International Journal of Communication Systems.

Generation of Arbitrarily-Oriented Ripple Images Using Smoothing Filter with Translated Window

Shogo Noma
Department of Information Systems, University of Nagasaki, Japan

Toru Hiraoka
Department of Information Systems, University of Nagasaki, Japan
E-mail: hiraoka@sun.ac.jp
https://sun.ac.jp/researchinfo/hiraoka/

Abstract

A non-photorealistic rendering method for automatically generating ripple images from photographic images using region-division smoothing filter has been proposed. Ripple patterns are composed of continuous lines with fluctuations, and ripple images are expressed by superimposing ripple patterns on photographic images. To generate ripple images that give different visual effects, this paper develops a method for generating ripple patterns with a texture different from the conventional method. The proposed method is executed by an iterative calculation using smoothing filter with the translated window. In the proposed method, the orientation of ripple patterns can be arbitrarily controlled by changing the amount of translation of the window used in smoothing filter. To verify the effectiveness of the proposed method, an experiment using various photographic images was conducted. Additionally, an experiment to visually examine how ripple patterns generated by changing the values of the parameters in the proposed method change.

Keywords: Non-photorealistic rendering, Ripple pattern, Smoothing filter, Translation of window, Arbitrary orientation

1. Introduction

In contrast to traditional computer graphics which focuses on photorealism, non-photorealistic rendering[1][2] which focuses on enabling different styles of expression in digital art is attracting attention. Non-photorealistic rendering often converts photographic images, videos, and three-dimensional data into art styles that imitate paintings, drawings, and cartoons. Recently, many studies have been conducted on non-photorealistic rendering to convert to unprecedented art styles such as cell-like images[3], moire-like images[4], and ripple images[5]. Cell-like patterns are composed of cell membrane and cell nucleus, and then cell-like images are overlaid with cell-like patterns on photographic images. Moire-like images are non-realistic images obtained by overlaying moire patterns on photographic images. Ripple images are non-realistic images expressing photographic images by ripple patterns, and ripple patterns are generated along the edges of photographic images.

This paper focuses on ripple images and proposes a method for generating ripple patterns with a texture different from the conventional method. The proposed method is executed by an iterative calculation using smoothing filter with the translated window. A feature of the proposed method is that the orientation of ripple patterns can be arbitrarily controlled by changing the amount of translation of the window. In the conventional method, for example, when ripple patterns in the orientation rotated by $\pi/4$ radian are generated, ripple patterns are expressed as being slightly bent, but the proposed method does not do so. An experiment using several photographic images show that the proposed method can generate ripple patterns on the entire image. Additionally, an experiment with changing the amount of R6translation of the window show that the proposed method can control the orientation of ripple patterns.

2. Proposed Method

The proposed method is largely executed in three steps. Step 1 is to apply smoothing filter. As a result, the finally generated ripple patterns can be expressed smoothly and clearly. Step 2 is to apply smoothing filter with the translated window. Step 3 is to apply inverse filter[6]. Inverse filter restores the image converted in Step 2 to the original image. By repeating the processes of Steps 1 and 2, restoration errors are accumulated, and ripple images are generated. A flow chart of the proposed method is shown in Fig. 1.

Details of the steps in Fig. 1 are explained below.

Step 0: The input pixel values on coordinates (i, j) of a gray-scale photographic image are defined as $o_{i,j}$. The pixel values $o_{i,j}$ have value of U gradation from 0 to $U-1$. Ripple patterns in the orientation of θ radian are generated.

Step 1: Smoothing filter is applied to the pixel values $o_{i,j}$, and the smoothed pixel values are defined as $f_{i,j}$.

$$f_{i,j} = \frac{\sum_{k=-W}^{W}\sum_{l=-W}^{W} o_{i+k,j+l}}{(2W+1)^2} \quad (1)$$

where k and l are the positions in the window, and W is the window size.

Step 2: Smoothing filter with the translated window is apply to the pixel values $f_{i,j}^{(t-1)}$, and the smoothed pixel values are defined as $g_{i,j}^{(t-1)}$, where t is the iteration number and $f_{i,j}^{(0)} = f_{i,j}$.

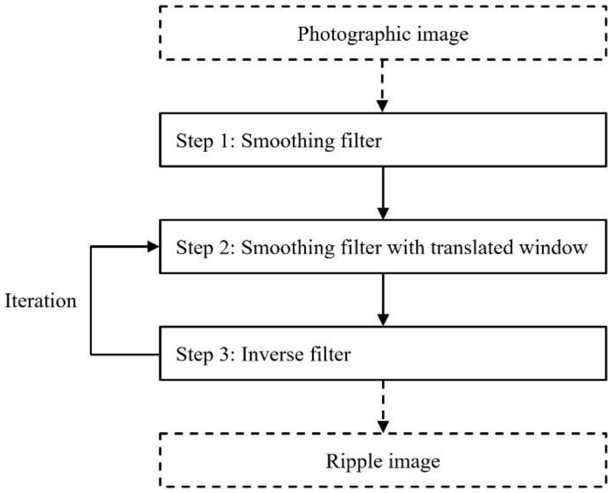


Fig.1. Flow chart of the proposed method

$$g_{i,j}^{(t)} = \frac{\sum_{k=-W}^{W}\sum_{l=-W}^{W} f_{i,j}^{(t-1)} o_{i+k+M,j+l+N}}{(2W+1)^2} \quad (2)$$

$$M = \begin{cases} round(\sqrt{2}sin(\theta)) & (t\%2=0) \\ round(\sqrt{2}sin(\theta+\pi)) & (t\%2=1) \end{cases} \quad (3)$$

$$N = \begin{cases} round(\sqrt{2}cos(\theta)) & (t\%2=0) \\ round(\sqrt{2}cos(\theta+\pi)) & (t\%2=1) \end{cases} \quad (4)$$

where % is a modulo operation and round is a function that rounds to an integer.

Step 3: Inverse filter is applied to the pixel values $g_{i,j}^{(t)}$, and the pixel values applied inverse filter are defined as $f_{i,j}^{(t)}$.

$$f_{i,j}^{(t)} = f_{i,j}^{(t-1)} - g_{i,j}^{(t)} + f_{i,j} \quad (5)$$

If $f_{i,j}^{(t)}$ is smaller than 0, then $f_{i,j}^{(t)}$ must be set to 0, and if $f_{i,j}^{(t)}$ is greater than $U-1$, then $f_{i,j}^{(t)}$ must be set to $U-1$.

Steps 1 and 2 are repeated T times. The image composed of the pixel values $f_{i,j}^{(T)}$ is a ripple image.

3. Experiments

Two experiments were conducted. The first experiment visually confirmed ripple patterns generated by changing the values of the parameters in the proposed method using Woman image shown in Fig. 2. The second experiment applied the proposed method to four photographic images shown in Fig. 3. All photographic images used in the experiments were 512 * 512 pixels and 256 gradations.

3.1. *Experiment with changing parameters*

Ripple images generated by changing the iteration number T were confirmed visually. The iteration number T was set to 5, 10, 20, and 30. The parameters W and θ were set to 2 and $\pi/4$, respectively. The results of the

Fig.2. Woman image

Fig.3. Various photographic images

experiment are shown in Fig. 4. As the value of T was larger, ripple patterns became clearer and were expressed finely.

Ripple images generated by changing the window size W were confirmed visually. The window size W was set to 1, 2, 3, and 4. The parameters T and θ were set to 30 and $\pi/4$, respectively. The results of the experiment are shown in Fig. 5. As the value of W was larger, the interval of ripple patterns became wider.

Ripple images generated by changing the orientation θ were confirmed visually. The orientation θ was set to 0, $\pi/6$, $\pi/6$, $\pi/4$, $\pi/3$, $\pi/2$, $2\pi/3$, $3\pi/4$, and

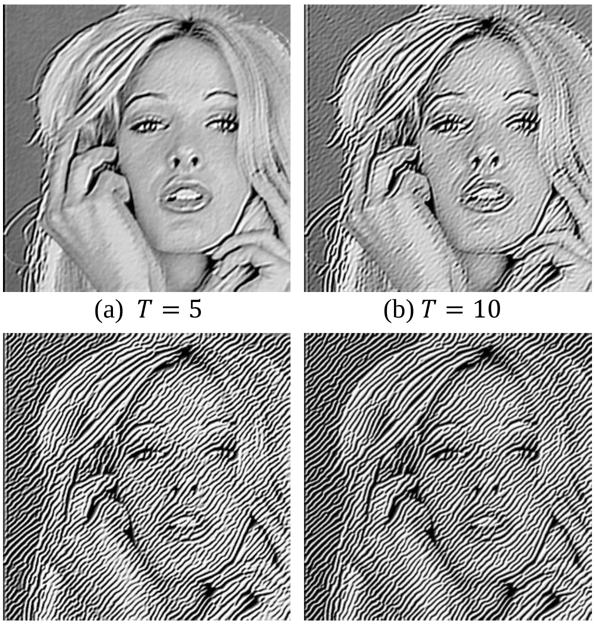

(a) $T = 5$ (b) $T = 10$
(c) $T = 20$ (d) $T = 30$

Fig.4. Ripple images in the case of $T = 5, 10, 20, 30$

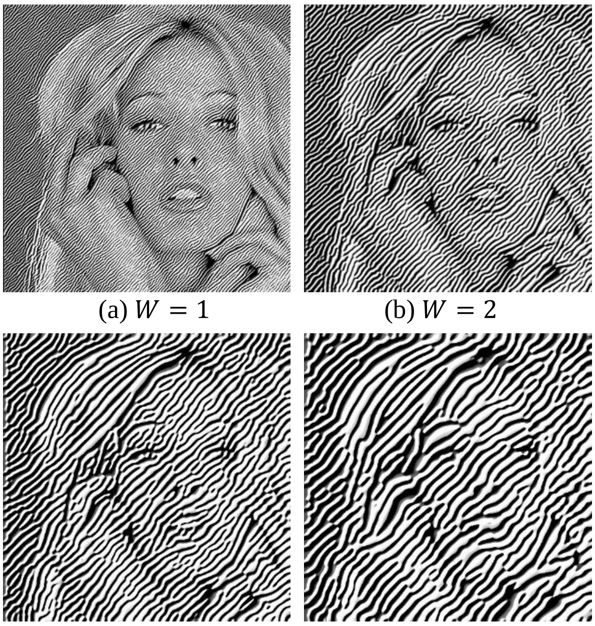

(a) $W = 1$ (b) $W = 2$
(c) $W = 3$ (d) $W = 4$

Fig.5. Ripple images in the case of $W = 1, 2, 3, 4$

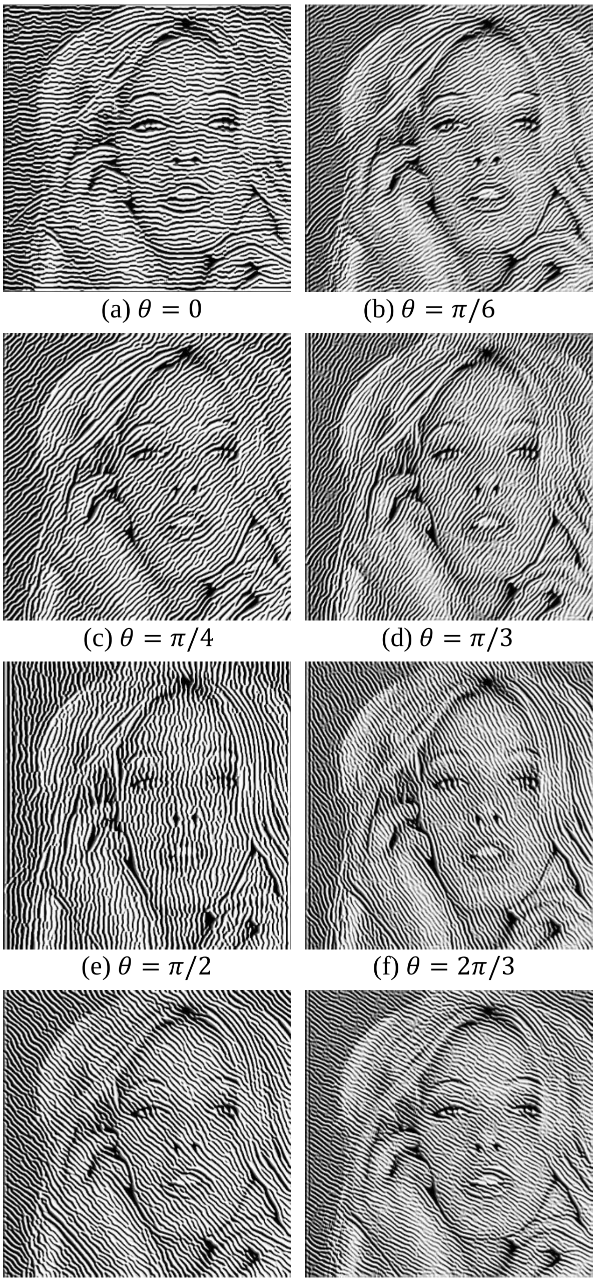

(a) $\theta = 0$ (b) $\theta = \pi/6$
(c) $\theta = \pi/4$ (d) $\theta = \pi/3$
(e) $\theta = \pi/2$ (f) $\theta = 2\pi/3$
(g) $\theta = 3\pi/4$ (h) $\theta = 5\pi/6$

Fig.6. Ripple images in the case of $\theta = 0, \pi/6, \pi/4, \pi/3, \pi/2, 2\pi/3, 3\pi/4, 5\pi/6$

$5\pi/6$. The parameters T and W were set to 30 and 2, respectively. The results of the experiment are shown in Fig. 6. For example, ripple patterns occur in the horizontal orientation when θ was 0, and ripple patterns occur in the vertical orientation when θ was $\pi/2$. By changing the value of θ, the orientation of ripple patterns could be controlled.

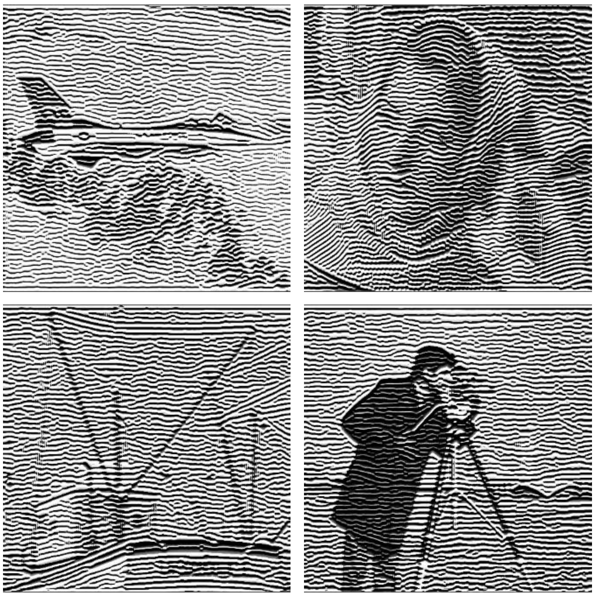


Fig.7. Various ripple images in the case of $\theta = 0$

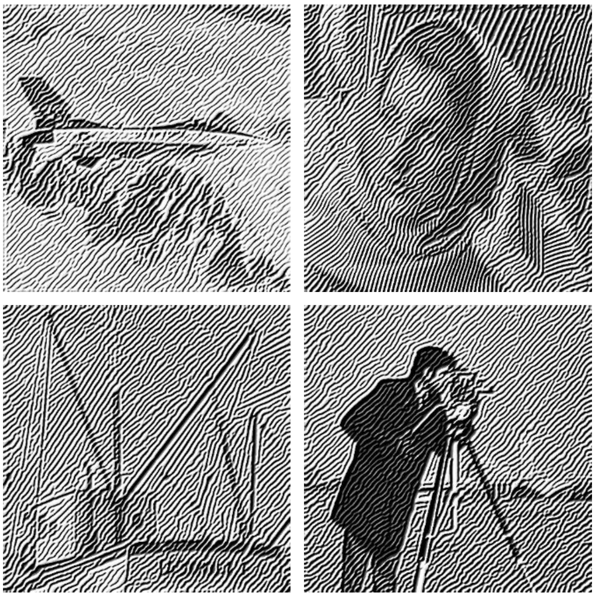


Fig.8. Various ripple images in the case of $\theta = \pi/4$

3.2. *Experiment using various photographic images*

The proposed method was applied to four photographic images shown in Fig. 3. The parameters T and W were respectively set to 30 and 2, and the orientation θ was set to 0, $\pi/4$, and $\pi/2$. The results of the experiment are shown in Fig. 7 to Fig. 9. The proposed method could generate ripple patterns on the entire image even in various photographic images. Additionally, the proposed method could control the orientation of ripple patterns in various photographic images.

4. Conclusion

This paper focused on ripple images by non-photorealistic rendering and proposed a method for generating ripple patterns with a texture different from the conventional method. The proposed method was executed by an iterative calculation using smoothing filter with the translated window. Experiments with various photographic images have shown that the proposed method can generate ripple patterns on the entire image and control the orientation of ripple patterns.

A subject for future study is to expand the proposed method for application to color photographic images, videos, and three-dimensional data.

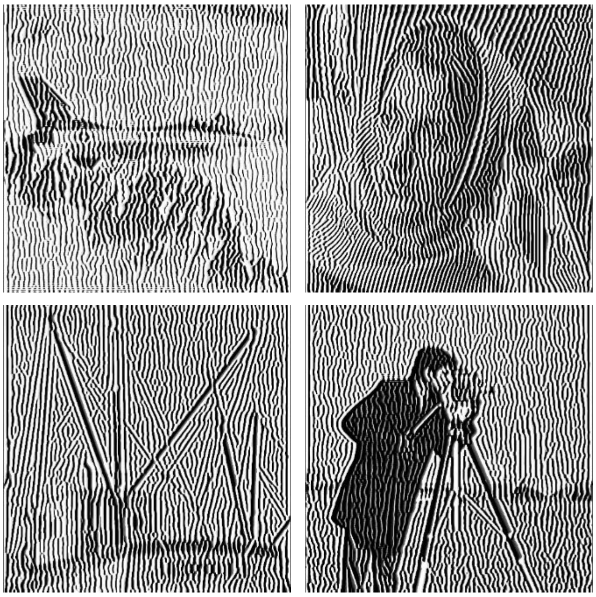


Fig.9. Various ripple images in the case of $\theta = \pi/2$

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number JP19K12664.

References

1. P. Haeberli. Paint by numbers: abstract image representations. ACM SIGGRAPH Computer Graphics, 1990, 24(4), 1990: 207-214.
2. J. Lansdown and S. Schofield. Expressive rendering: a review of nonphotorealistic techniques. IEEE Computer Graphics and Applications, 1995, 15(3): 29-37.
3. T. Hiraoka, M. Hirota, K. Inoue and K. Urahama. Generating cell-like color images by inverse iris filter. ICIC Express Letters, 2017, 11(2): 399-404.
4. T. Hiraoka. Generation of moire-like images smoothing stepwise changes using non-local bilateral filter. ICIC Express Letters, 2021, 15(8): 829-835.
5. T. Hiraoka. Generation of arbitrarily-oriented ripple images using circular-sector-type smoothing filter and inverse filter. Journal of Robotics, Networking and Artificial Life, 2020, 6(4): 213-216.
6. Z. Yu and K. Urahama. Iterative method for inverse nonlinear image processing. IEICE Transactions on Fundamentals, 2014, E97-A(2): 719-721.

Authors Introduction

Mr. Shogo Noma

He is a student in the Department of Information Systems, University of Nagasaki. His research interests inculede IoT 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.

Generation of Moire-Like Videos from RGB-D Videos Window

Sho Enomoto
Department of Information Systems, University of Nagasaki, Japan

Toru Hiraoka
Department of Information Systems, University of Nagasaki, Japan
E-mail: hiraoka@sun.ac.jp
https://sun.ac.jp/researchinfo/hiraoka/

Abstract

A non-photorealistic rendering method has been proposed to generate moire-like images from photographic images using bilateral filter and unsharp mask. Extensions to the conventional method have also been proposed to generate moire-like videos from videos or to generate moire-like images from RGB-D images. In this paper, a method is proposed to generate moire-like videos from RGB-D videos. Flickering is a problem in NPR videos, but the proposed method can suppress flicker. Through experiments using an RGB-D video taken by the authors, the flicker of moire-like videos generated by the proposed method was evaluated visually and quantitatively.

Keywords: Non-photorealistic rendering, Moire, Depth, Video, Bilateral filter, Unsharp mask

1. Introduction

Many non-photorealistic rendering (NPR) method using image processing have been proposed [1][2]. One of the NPRs has proposed to automatically generate moire-like images from photographic images [3]. Moire-like patterns are generated according to the change of the edges and the shading in the photographic images, and moire-like images are a type of op art expressed by superimposing moire-like patterns on photographic images. Moire-like images are generated by iterative calculations using bilateral filter[4] and unsharp mask. Extensions to the conventional method[3] have also been proposed to generate moire-like videos from videos[5] or to generate moire-like images from RGB-D images[6]. The conventional methods[5][6] added time and depth terms to the exponent part in the equation of bilateral filter, respectively. Moire-like videos[5] can improve the visual effect by moving moire-like patterns, and moire-like images[6] can generate more distorted moire-like patterns than moire-like images[3].

This paper proposes an NPR method to automatically generate moire-like videos from RGB-D videos. As far as the authors investigate, there is no research on NPR using RGB-D videos. The proposed method adds both time and depth terms to the exponent part in the equation of bilateral filter. Flickering is generally a problem in NPR videos[7], and when the conventional method[5] is applied to each frame of RGB-D videos to generate moire-like videos, flicker occurs in the same way. Hereinafter, moire-like videos generated by applying the conventional method[7] to each frame of RGB-D videos will be referred to as conventional moire-like videos. To confirm that the proposed method can generate moire-like videos with suppressed flicker, experiments were conducted using RGB-D videos taken by the authors.

2. Proposed Method

The proposed method is executed in two steps: Step 1 is the process using bilateral filter embedded the time and depth, and Step 2 is the process using unsharp mask and bilateral filter embedded the time. A flow chart of the proposed method is shown in Fig. 1.

Details of the steps in Fig. 1 are explained below.
Step 0: The input pixel values (R, G, B) and the depths

©The 2023 International Conference on Artificial Life and Robotics (ICAROB2023), Feb. 9 to 12, on line, Oita, Japan

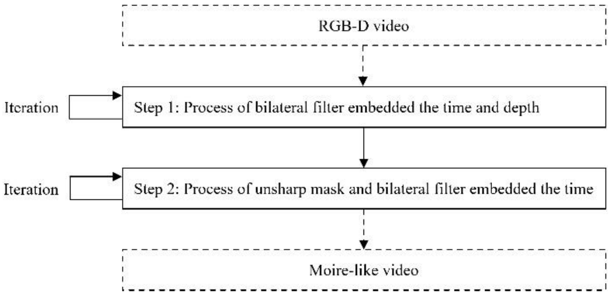

RGB-D video
Iteration
Step 1: Process of bilateral filter embedded the time and depth
Iteration
Step 2: Process of unsharp mask and bilateral filter embedded the time
Moire-like video

Fig.1. Flow chart of the proposed method

for spatial coordinates (i,j) in k-th frame of an RGB-D video are defined as $f_{R,i,j,k}$, $f_{G,i,j,k}$, $f_{B,i,j,k}$, and $f_{D,i,j,k}$, respectively, where $i=1,2,\cdots,I$, $j=1,2,\cdots,J$, and $k=1,2,\cdots,K$. The pixel values $f_{R,i,j,k}$, $f_{G,i,j,k}$, $f_{B,i,j,k}$, and $f_{D,i,j,k}$ have value of U gradation from 0 to $U-1$. The depths $f_{D,i,j,k}$ are linearly transformed so that the minimum distance becomes 0 and the maximum distance becomes $U-1$.

Step 1: The pixel values of the image at the t-th iteration number are defined as $f^{(t)}_{R,i,j,k}$, $f^{(t)}_{G,i,j,k}$, and $f^{(t)}_{B,i,j,k}$, where $f^{(0)}_{R,i,j,k}=f_{R,i,j,k}$, $f^{(0)}_{G,i,j,k}=f_{G,i,j,k}$, and $f^{(0)}_{B,i,j,k}=f_{B,i,j,k}$. The output pixel values $f^{(t)}_{R,i,j,k}$, $f^{(t)}_{G,i,j,k}$, and $f^{(t)}_{B,i,j,k}$ in bilateral filter embedded the time and depth are calculated by the following equations.

$$f^{(t)}_{R,i,j,k}=\frac{\sum_{l=i-W}^{i+W}\sum_{m=j-W}^{j+W}\sum_{n=k-O}^{k+O}E^{(t-1)}_{1,R,i,j,k,l,m,n}f^{(t-1)}_{R,l,m,n}}{\sum_{l=i-W}^{i+W}\sum_{m=j-W}^{j+W}\sum_{n=k-O}^{k+O}E^{(t-1)}_{1,R,i,j,k,l,m,n}} \tag{1}$$

$$E^{(t-1)}_{1,R,i,j,k,l,m,n}=e^{-\alpha((i-l)^2+(j-m)^2)}e^{-\beta(f^{(t-1)}_{R,i,j,k}-f^{(t-1)}_{R,l,m,n})^2}e^{-\gamma(f_{D,i,j,k}-f_{D,l,m,n})^2}e^{-\delta(k-n)^2} \tag{2}$$

$$f^{(t)}_{G,i,j,k}=\frac{\sum_{l=i-W}^{i+W}\sum_{m=j-W}^{j+W}\sum_{n=k-O}^{k+O}E^{(t-1)}_{1,G,i,j,k,l,m,n}f^{(t-1)}_{G,l,m,n}}{\sum_{l=i-W}^{i+W}\sum_{m=j-W}^{j+W}\sum_{n=k-O}^{k+O}E^{(t-1)}_{1,G,i,j,k,l,m,n}} \tag{3}$$

$$E^{(t-1)}_{1,G,i,j,k,l,m,n}=e^{-\alpha((i-l)^2+(j-m)^2)}e^{-\beta(f^{(t-1)}_{G,i,j,k}-f^{(t-1)}_{G,l,m,n})^2}e^{-\gamma(f_{D,i,j,k}-f_{D,l,m,n})^2}e^{-\delta(k-n)^2} \tag{4}$$

$$f^{(t)}_{B,i,j,k}=\frac{\sum_{l=i-W}^{i+W}\sum_{m=j-W}^{j+W}\sum_{n=k-O}^{k+O}E^{(t-1)}_{1,B,i,j,k,l,m,n}f^{(t-1)}_{B,l,m,n}}{\sum_{l=i-W}^{i+W}\sum_{m=j-W}^{j+W}\sum_{n=k-O}^{k+O}E^{(t-1)}_{1,B,i,j,k,l,m,n}} \tag{5}$$

$$E^{(t-1)}_{1,B,i,j,k,l,m,n}=e^{-\alpha((i-l)^2+(j-m)^2)}e^{-\beta(f^{(t-1)}_{B,i,j,k}-f^{(t-1)}_{B,l,m,n})^2}e^{-\gamma(f_{D,i,j,k}-f_{D,l,m,n})^2}e^{-\delta(k-n)^2} \tag{6}$$

where W is the window size, O is the number of the forward and backward frames used in the calculation, α, β, γ, and δ are positive constants, l and m are the positions in the window, and n is the position in forward and backward frames. In Equations (2), (4), and (6), the depth term $-\gamma(f_{D,i,j,k}-f_{D,l,m,n})^2$ and the time term $-\delta(k-n)^2$ are added to the coefficient of exponential function in the equation of bilateral filter. The larger the value of γ, the more different moire-like patterns from the conventional method[3] are generated. The smaller the value of δ, the more affected by the forward and backward frames. When the value of the frame number O is 0, the conventional moire-like videos are generated. The process of Step 1 is repeated T_1 times.

Step 2: The pixel values $f^{(T_1)}_{R,i,j,k}$, $f^{(T_1)}_{G,i,j,k}$, and $f^{(T_1)}_{B,i,j,k}$ are defined as $g^{(0)}_{R,i,j,k}$, $g^{(0)}_{G,i,j,k}$, and $g^{(0)}_{B,i,j,k}$, respectively. The output pixel values $g^{(t)}_{R,i,j,k}$, $g^{(t)}_{G,i,j,k}$, and $g^{(t)}_{B,i,j,k}$ in unsharp mask using bilateral filter embedded the time are calculated by the following equations.

$$g^{(t)}_{R,i,j,k}=2g^{(t-1)}_{R,i,j,k}-\frac{\sum_{l=i-W}^{i+W}\sum_{m=j-W}^{j+W}\sum_{n=k-O}^{k+O}E^{(t-1)}_{2,R,i,j,k,l,m,n}g^{(t-1)}_{R,l,m,n}}{\sum_{l=i-W}^{i+W}\sum_{m=j-W}^{j+W}\sum_{n=k-O}^{k+O}E^{(t-1)}_{2,R,i,j,k,l,m,n}} \tag{7}$$

$$E^{(t-1)}_{2,R,i,j,k,l,m,n}=e^{-\alpha((i-l)^2+(j-m)^2)}e^{-\beta(g^{(t-1)}_{R,i,j,k}-g^{(t-1)}_{R,l,m,n})^2}e^{-\delta(k-n)^2} \tag{8}$$

$$g^{(t)}_{G,i,j,k}=2g^{(t-1)}_{G,i,j,k}-\frac{\sum_{l=i-W}^{i+W}\sum_{m=j-W}^{j+W}\sum_{n=k-O}^{k+O}E^{(t-1)}_{2,G,i,j,k,l,m,n}g^{(t-1)}_{G,l,m,n}}{\sum_{l=i-W}^{i+W}\sum_{m=j-W}^{j+W}\sum_{n=k-O}^{k+O}E^{(t-1)}_{2,G,i,j,k,l,m,n}} \tag{9}$$

$$E^{(t-1)}_{2,G,i,j,k,l,m,n}=e^{-\alpha((i-l)^2+(j-m)^2)}e^{-\beta(g^{(t-1)}_{G,i,j,k}-g^{(t-1)}_{G,l,m,n})^2}e^{-\delta(k-n)^2} \tag{10}$$

$$g^{(t)}_{B,i,j,k}=2g^{(t-1)}_{B,i,j,k}-\frac{\sum_{l=i-W}^{i+W}\sum_{m=j-W}^{j+W}\sum_{n=k-O}^{k+O}E^{(t-1)}_{2,B,i,j,k,l,m,n}g^{(t-1)}_{B,l,m,n}}{\sum_{l=i-W}^{i+W}\sum_{m=j-W}^{j+W}\sum_{n=k-O}^{k+O}E^{(t-1)}_{2,B,i,j,k,l,m,n}} \tag{11}$$

$$E^{(t-1)}_{2,B,i,j,k,l,m,n}=e^{-\alpha((i-l)^2+(j-m)^2)}e^{-\beta(g^{(t-1)}_{B,i,j,k}-g^{(t-1)}_{B,l,m,n})^2}e^{-\delta(k-n)^2} \tag{12}$$

In case $g^{(t)}_{R,i,j,k}$, $g^{(t)}_{G,i,j,k}$, and $g^{(t)}_{B,i,j,k}$ are less than 0, then $g^{(t)}_{R,i,j,k}$, $g^{(t)}_{G,i,j,k}$, and $g^{(t)}_{B,i,j,k}$ must be set to 0, respectively. In case $g^{(t)}_{R,i,j,k}$, $g^{(t)}_{G,i,j,k}$, and $g^{(t)}_{B,i,j,k}$ are greater than $U-1$, then $g^{(t)}_{R,i,j,k}$, $g^{(t)}_{G,i,j,k}$, and $g^{(t)}_{B,i,j,k}$ must be set to $U-1$, respectively. The process of Step 2 is repeated T_2 times, and a video composed of the pixel values $g^{(T_2)}_{R,i,j,k}$, $g^{(T_2)}_{G,i,j,k}$, and $g^{(T_2)}_{B,i,j,k}$ is the moire-like video.

3. Experiments

Experiments were conducted using an RGB-D video which consists of 440 frames, 49 frames / second, 320 * 180 pixels, and 256 gradations. The RGB-D video was shot using ZED stereo camera to capture a scene of a man moving indoor. The farthest distance from the camera was 5.987 meters. The RGB and depth images at the 1th, 40th, 80th, 120th, 160th, 200th, 240th, 280th, 320th, 360th, 400th, and 440th frames of the RGB-D videos are shown in Fig. 2 and Fig. 3, respectively. In reference to the literatures[3][5][6], the values of the parameters α, β, γ, δ, W, O, T_1, and T_2 used in all experiments were set to 0.01, 0.01, 0.001, 0.1, 10, 3, 10, and 20, respectively. Hereinafter, the moire-like video generated by the proposed method under the above conditions is referred to as a proposed moire-like video, and the moire-like video generated by applying the conventional method[6] to each frame of the RGB-D video under the above conditions is referred to as a conventional moire-like video.

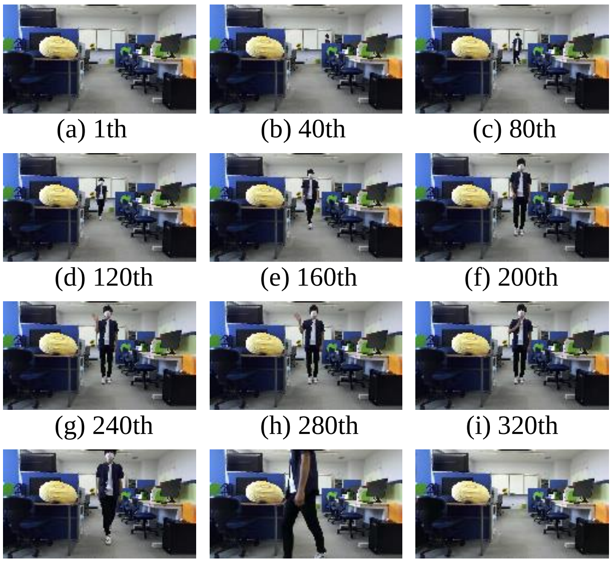

(a) 1th (b) 40th (c) 80th
(d) 120th (e) 160th (f) 200th
(g) 240th (h) 280th (i) 320th
(j) 360th (k) 400th (l) 440th

Fig.2. RGB images

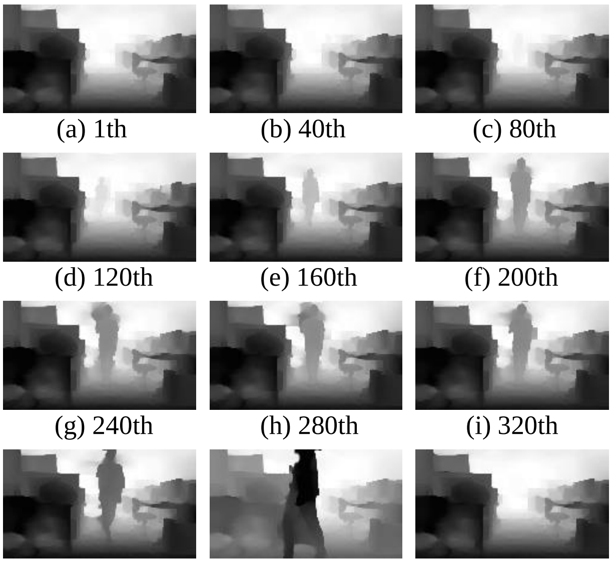

(a) 1th (b) 40th (c) 80th
(d) 120th (e) 160th (f) 200th
(g) 240th (h) 280th (i) 320th
(j) 360th (k) 400th (l) 440th

Fig.3. Depth images

The proposed and conventional moire-like videos were visually evaluated. The conventional moire-like video had many flicker. On the other hand, in the proposed video, flicker was suppressed and the movement of moire-like patterns was smooth. As an example, the proposed and conventional moire-like images of the 300th and 301th adjacent frames are shown in Fig. 4 and Fig. 5, respectively. From Fig. 4 and Fig. 5, the proposed moire-like patterns changed less between the two adjacent frames than the conventional moire-like patterns.

The proposed and conventional moire-like videos were quantitatively evaluated. The averages of the absolute values of the differences between the pixel values between adjacent frames (hereinafter, frame difference averages) were calculated, and the average of the frame difference averages of all frames (hereinafter, all frame difference average) was calculated. The all frame difference average P is calculated by the following equations.

$$P = \frac{\sum_{k=1}^{K-1}\sum_{i=1}^{I}\sum_{j=1}^{J} E_{3,i,j,k}}{3(K-1)} \tag{13}$$

$$E_{3,i,j,k} = |g_{R,i,j,k}^{(T_2)} - g_{R,i,j,k+1}^{(T_2)}| + |g_{G,i,j,k}^{(T_2)} - g_{G,i,j,k+1}^{(T_2)}| + |g_{G,i,j,k}^{(T_2)} - g_{G,i,j,k+1}^{(T_2)}| \tag{14}$$

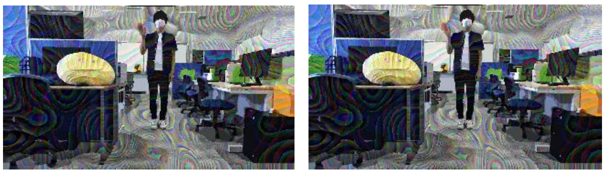

(a) 300th (b) 301th

Fig.4. Proposed moire-like images

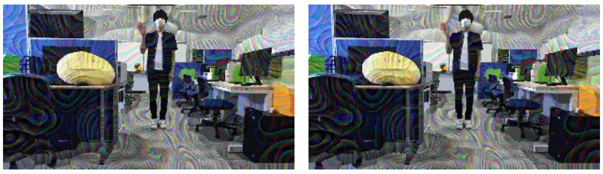

(a) 300th (b) 301th

Fig.5. Conventional moire-like images

Table 1. All frame difference averages

Proposed moire-like video	Conventional moire-like video
485216.922	935160.140

It is judged that the smaller the all frame difference average, the less the flicker of moire-like videos. The all frame difference averages of the proposed and conventional moire-like videos are shown in Table 1. The all frame difference averages of the proposed and conventional moire-like videos are 485216.922 and 935160.140, respectively. The proposed moire-like video had a smaller all frame difference average than the conventional moire-like video. Thus, the proposed moire-like video had less flicker than the conventional moire-like video.

4. Conclusion

This paper proposed an NPR method to automatically generate moire-like videos from RGB-D videos. The proposed method added both time and depth terms to the exponent part in the equation of bilateral filter in the conventional method[3]. Through experiments using an RGB-D video taken by the authors, the appearance of moire-like videos generated by the proposed method was evaluated visually and quantitatively. As a result of the experiments, it was found that the proposed method can suppress flicker.

The future task is to apply the proposed method to more RGB-D videos, although this paper applied the proposed method to one type of RGB-D video.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number JP19K12664.

References

1. P. Haeberli. Paint by numbers: abstract image representations. ACM SIGGRAPH Computer Graphics, 1990, 24(4), 1990: 207-214.
2. J. Lansdown and S. Schofield. Expressive rendering: a review of nonphotorealistic techniques. IEEE Computer Graphics and Applications, 1995, 15(3): 29-37.
3. T. Hiraoka and K. Urahama. Generation of moire-picture-like color images by bilateral filter. IEICE Transactions on Information and Systems, 2013, E96-D(8): 1862-1866.
4. S. Paris, P. Kornprobst, J. Tumblin, and F. Durand. Bilateral filtering: theory and applications. Foundations and Trends in Computer Graphics and Vision, 2008, 4(2): 1-73.
5. T. Hiraoka and N. Ataka. Generating moire-like animation by bilateral filter. The Japanese Journal of the Institute of Industrial Applications Engineers, 2018, 6(1): 17-22.
6. R. Takaki and T. Hiraoka. Generation of moire-like images from RGB-D images. ICIC Express Letters, 2021, 15(1): 37-42.
7. B. J. Meier. Painterly rendering for animation. Proceedings of SIGGRAPH 96, 1996: 477-484.

Authors Introduction

Mr. Sho Enomoto

He is a student in the Department of Information Systems, University of Nagasaki. His research interests inculede machine learning 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.

A Proposal of Shoulder-surfing Attack Countermeasure Method with Improved Usability

Yoshihiro Kita
Faculty of Information Systems,
University of Nagasaki,
1-1-1 Manabino, Nagayo, Nishisonogi , Nagasaki 851-2195, Japan

Shingo Nakamura
Faculty of Information Systems,
University of Nagasaki,
1-1-1 Manabino, Nagayo, Nishisonogi , Nagasaki 851-2195, Japan

E-mail: kita@sun.ac.jp

Abstract

Shoulder-surfing attacks are one of the most familiar password exploitation attacks. It is vulnerable to be attacked while unlocking a smartphone screen. The smartphone users need to take countermeasures against that attack. The fingerprint-based screen unlock system has become the norm, but it is not safe, as there has been increased in the user's fingerprint theft. The existing methods to prevent the shoulder-surfing attacks are effective against such attacks, but many of them are complicated to operate, and difficult to use. In this paper, we propose the prevent method for surfing attacks that it is easy to use. The tool's user operates the lower buttons, moves characters to on the trump's marks in specified advance. The user can input as likely as the password. On the other hand, the attacker does not understand the input characters only shown these buttons has been pushed.

Keywords: shoulder-surfing attack, unlocking a smartphone screen, security and usability.

1. Introduction

Recently, mobile devices had the screen lock using the authentication methods, e.g., PINs and patterns, for the prevention of unlawful using by others. These authentication methods are vulnerable to attack by peeking (shoulder-surfing attack).

These numbers/patterns as the passwords are leaked to others who peek mobile device's screen. The authentication methods [1][2][3][4] have the resistance to shoulder-surfing attack. However, these methods have complicated user interface.

In this paper, we propose the shoulder-surfing attack countermeasure method that is improved usability. In this method, numbers, characters, symbols, and colors are arranged on each of the two layers in reference the background pattern slide authentication [3], and the user selects the combination in which they overlap to be authenticated the user. We confirm the usability of our proposal method, an experiment was performed in which several combination patterns were prepared, and the subjects were asked to choose the one with the best usability.

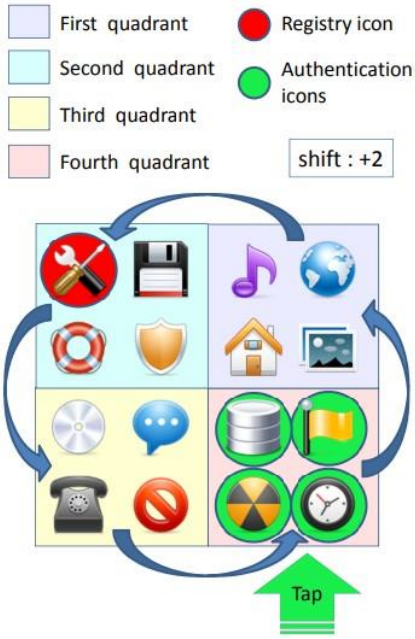
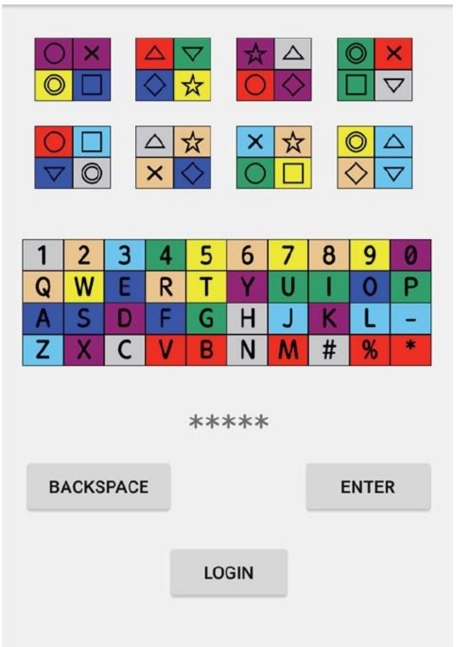


Fig.1. A sample of the secret tap with double shift (STDS) method [2]



1	2	3	4	5	6	7	8	9	0
Q	W	E	R	T	Y	U	I	O	P
A	S	D	F	G	H	J	K	L	-
Z	X	C	V	B	N	M	#	%	*

Fig.2. A sample of the background pattern slide authentication [3]

2. Related Works

2.1. *Secret Tap with Double Shift (STDS)*

The secret tap with double shift (STDS) method [2] is an unlock system for smart phone's display as shown as Fig.1. This method places 16 randomly selected icons in the display area, which is a 4×4 square. The user selects the authentication icon from the 16 icons and taps the selected icon. The user repeats this operation for a predetermined number of registry icons. If all selected icons are correct authentication icons, the authentication is successful, and the smart phone's display is unlocked.

The user needs to memorize the 4 icons and shifts as PIN and needs to calculate the place of tap icons from these. These are a heavy burden for the user. These operations have to be made simpler to reduce the burden of the user.

2.2. *Background Pattern Slide Authentication*

The background pattern slide authentication [3] is an unlock system for smart phone's display as shown as Fig.2. First, the user selects a square from upper grids. Next, changes a target character's color to same color as selected square, and enter the target characters as PIN.

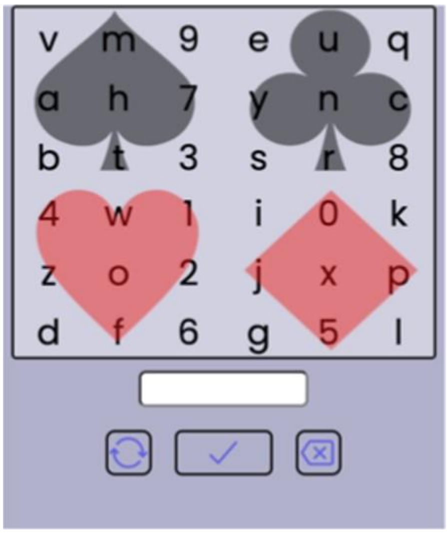
In this method, the target character is not directly touched during PIN enter, multiple target characters on the same background color, are listed as PIN candidates. The attacker cannot be identified PIN uniquely.

However, this method is less usability due to the small size of each grid and complicated coloring. The user needs nervous operation to use this method.

2.3. *CCC (Circle Chameleon Cursor)*

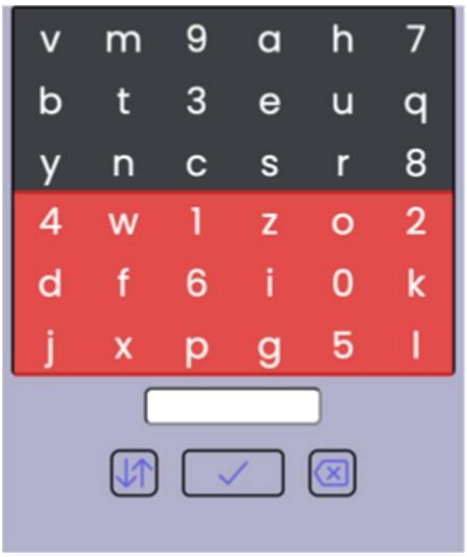
CCC (Circle Chameleon Cursor) [4] is the unlock system using background pattern and phone's vibration. This method is based on fakePointer [1], consists of 3 parts; PIN indicator, the number input dial, and PIN enter button. The numbers are placed on the number input dial.

First, PIN indicator rotates on the number input dial, when PIN indicator comes to a specified position, the smartphone vibrates. The user feels this vibration and recognizes the position as a cursor for PIN enter. Next, the user operates the number input dial, adjust the number as PIN to this position, and tap the PIN enter button.



v	m	9	e	u	q
a	h	7	y	n	c
b	t	3	s	r	8
4	w	1	i	0	k
z	o	2	j	x	p
d	f	6	g	5	l

Fig.3. The background pattern (1) of our proposal method



v	m	9	a	h	7
b	t	3	e	u	q
y	n	c	s	r	8
4	w	1	z	o	2
d	f	6	i	0	k
j	x	p	g	5	l

Fig.4. The background pattern (2) of our proposal method

This method takes time to unlock the smartphone because the user must detect the position for PIN enter by the smartphone’s vibration.

3. Our proposal method for improving usability

In this paper, we propose the shoulder-surfing attack countermeasure method that is improved usability. We consider the following to keep the security for against shoulder-surfing attack.

- PIN is formed by choosing 4 characters from 36 characters that includes numbers and English alphabets.
- The 36 characters are displayed randomity.
- When enter a PIN, multiple characters are candidates for the PIN, so that a character is not uniquely defined it.
- When enter a PIN, the user not directly touch the character as the PIN.

The other hand, we consider the following to improve usability.

- The minimize as much information as possible for the user must be to memorize.
- The user’s operation and display are simpler as possible.

We propose 3 patterns that satisfy these conditions as shown as Fig.3, Fig.4, and Fig.5.

Fig.3 shows pattern (1) which uses trump’s marks at background pattern. This is divided into 4 areas according to the mark, and each area contains 9 characters. The user moves and enter the mark for each PIN so that each mark of area containing a PIN, is all same. In other words, the smartphone is unlocked if all entered marks are same, i.e., 4 marks are entered for 4-digits PIN. The marks are moved clockwise by tapped lower-left button. A mark of the area containing a PIN is entered by tapped lower-center button. If the user is necessary to cancel an entered mark, taps the lower-right button.

Fig.4 shows pattern (2) which colored black and red at background pattern. This is divided into 2 areas, and each area contains 18 characters. The operation and algorithm are the same as pattern (1). The user enter each color of the area containing a PIN. If all entered colors are same, the smartphone is unlocked. Each color is replaced by tapped lower-left button.

Fig.5 shows patten (3) which uses trump’s marks at background patten as same as patten (1). This is divided into 9 areas according to the mark, and each area contains 4 characters. Each mark is rotated clockwise around the center mark by tapped lower-left button. The center mark is displayed a mark “Spade”, but the center area is always

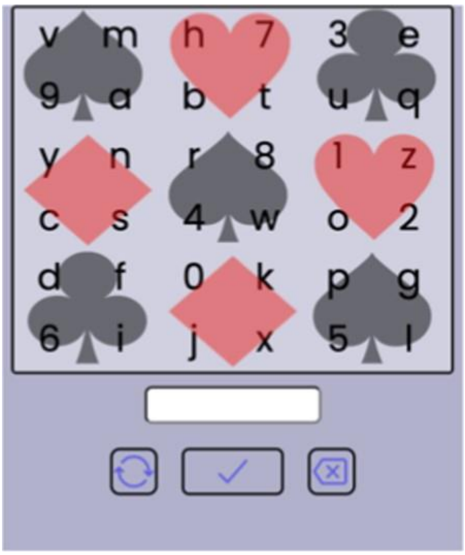


Fig.5. The background pattern (3) of our proposal method

entered, it means “Wildcard”. (Currently using a different mark that is not a trump’s mark.)

4. Evaluation of usability

We conducted the trial experiment to confirm the usability of these patten. The subjects are the 10 students belong to University of Nagasaki.

First, each subject register PIN using their devices. Next, they try to unlock each pattern of our proposal method at 10 times. Finally, they scored each pattern on 4 points scale as usability score as follows.

- Excellent (score 4)
- Good (score 3)
- Fair (score 2)
- Poor (score 1)

Table.1 shows the results of experiment for usability of our method. Pattern (2) had the shortest unlock time and highest usability score of all patterns, although there was 1 mistouch.

The subject’s most common opinion was:
“Pattern (2) is easy to operate, view, and understand.”
Therefore, pattern (2) is the highest usability of patterns.

However, we consider that pattern (2) is not secure against the accidental unlocking. Table.2 shows the probability of accidental unlocking of each patten. Pattern (1) and (3) have a probability of less than 1%, while Pattern (2) has a high probability as 6.25%. In usually, safety and usability are trade-off, so high usability means less safety. Too high level of usability can be dangerous. If the probability of accidental unlocking is kept less than 1%, pattern (1) is suitable high usability method for against shoulder-surfing attack.

5. Conclusion

Table.1. The results of experiment for usability

Patterns	Unlock average time (sec)	Times of mistouch (total)	Usability score (average)
Pattern (1)	16.75	0	3.0
Pattern (2)	11.81	1	3.5
Pattern (3)	15.62	0	2.1

Table.2. The probability of accidental unlocking of each pattern

Patterns	The probability of accidental unlocking
Pattern (1)	(9/36) ^ 4 = 0.39%
Pattern (2)	(18/36) ^ 4 = 6.25%
Pattern (3)	(8/36) ^ 4 = 0.24%

In this paper, we proposed the shoulder-surfing attack countermeasure method that is improved usability. The proposal methods were defined 3 patterns by the different of background patten.

The experiment results shows that pattern (2) is the highest usability method of all patterns, but we led to conclusion that pattern (1) is suitable high usability method for against shoulder-surfing attack by safety factors.

Our future works are as follows,

- Conducting additional experiments to confirm the accidental unlocking by subjects
- Conducting additional experiments to confirm our proposal methods resistance to shoulder-surfing attack.
- Discuss the other safety considerations for smartphone's unlocking

References

[1] T. Takada, "fakePointer: An Authentication Scheme for a Better Security Against a Peeping Attack by a Video Camera", Proceedings of the 2nd International Conference on Mobile Ubiquitous Computing, Systems, Service and Technologies (UBICOMM2008), 2008.

[2] Y. Kita, F. Sugai, M. Park, and N. Okazaki, "Proposal and its Evaluation of a Shoulder-Surfing Attack Resistant Authentication Method: Secret Tap with Double Shift", International Journal of Cyber-Security and Digital Forensics (IJCSDF), Vol.2, No.1, pp.48-55, 2014.

[3] M. Tanaka and H. Inaba, "Proposal of Improved Background Pattern Slide Authentication against Shoulder Surfing in Consideration of Convenience", Journal of Information Processing Society of Japan, Vol.58, No.9, pp.1513-1522, 2017 (in Japanese).

[4] M. Ishiduka and T. Takada, "CCC: Repeated Observation Attack Resilient PIN Authentication System Using Vibration", Journal of Information Processing Society of Japan, Vol.56, No.9, pp.1877-1888, 2015 (in Japanese).

Authors Introduction

Dr. Yoshihiro Kita

He received his Doctor's degree from the Department of Engineering, University of Miyazaki, Japan in 2011. He belongs to University of Nagasaki, Japan. His research area is biometrics, mobile security, and software testing.

Mr. Shingo Nakamura

He belongs to the Faculty of Information Systems, University of Nagasaki, in Japan. He researchs the security and usability when unlock a smart phone's display. He interested in mobile security.

User-movement Estimation in Social Media Sites Based on Seq2Seq Model

Masaharu Hirota
Department of Information Science, Okayama University of Science
1-1 Ridaicho, Kita-ku, Okayama-shi, 700-0005, Japan

E-mail: hirota@ous.ac.jp

Abstract

Many tourists upload content about tourist attractions to social media sites. The sequence of location information annotated on the content represents the user's movement. This movement information is an important source that can be used for recommendations and advertisements. However, users do not consistently post content about all the places they visit on social media. sites. This study aims to develop a method to estimate the location of users' movements. Our method uses a Seq2Seq model, which learns the reconstruction of users' movement trajectories.

Keywords: Sequence-to-sequence model, Human mobility, Trajectory, Twitter,

1. Introduction

In recent years, with the proliferation of social media sites, location-based services, and GPS-equipped mobile devices, many movement trajectories have been generated, such as people's movements, check-ins, vehicle trajectories, and surveillance camera data. These trajectories are essential information for applications such as arrival time prediction [1], traffic volume prediction [2], personalized location recommendation [3], and trajectory similarity measurement [4]. The movement trajectory dataset's quality significantly impacts these applications' performance. This is because low sample rate trajectories increase the uncertainty of the movement information as much of the user's movement is lost. Therefore, ensuring many movement trajectories with high sample rates is essential for these applications.

In addition, social media sites such as Twitter [5] have large amounts of location-annotated content posted by people. Therefore, the sequence of such content represents a user's movement and is used because of the applications mentioned above. However, to use a user's content sequence as a movement trajectory on Twitter, it is necessary to consider the fact that users do not post tweets at all destinations. In other words, there is a significant difference between the true sequence representing the actual movement of a user and the sequence represented by a tweet, which is the user's movement of a tweet. As a result, sequences obtained from social media, such as tweets, are missing information on many moves and have low sample rates. This paper addresses pseudo-estimation of user movement by performing trajectory recovery on sequences of user movements on Twitter.

Trajectory recovery increases the movement trajectory's sample rate by recovering the trajectory's lost points. For example, trajectory recovery methods have been proposed based on the Sequence-to-Sequence (Seq2Seq) [6] model and mainly focused on increasing the sample rate of vehicle and pedestrian movement trajectories. The movement trajectories in the datasets used in those studies often have a sample rate of a few seconds to a few minutes. On the other hand, tweet sequences' sample rate is from a few minutes to a few hours. In this study, we treat sequences of tweets with lower sample rates than those movement trajectories.

In this study, user movement is estimated by increasing the sample rate of tweet sequences using the attention-based Seq2Seq model. The user movement estimated in this study is the movement between sub-areas, where an area is divided into a grid. The model has an encoder that converts the sequence into a vector and a decoder that creates the sequence from the vector. The sequence of tweets is input to the encoder, and the output of the decoder is the estimated movement of the tweet.

The remainder of this paper is organized as follows. Section 2 describes work related to this topic. Section 3 presents our method for estimating user-movements using the Seq2Seq model. Section 4 describes the evaluation result of our method. Section 5 concludes the paper and future work.

2. Related Work

Studies have been conducted to recover low-sample-rate motion trajectories to generate high-sample-rate motion trajectories. Chen et al. propose an algorithm to discover the most popular route from a transfer network based on the popular indicators in a breadth-first manner [7]. Hoteit et al. evaluated cell phone data recovery for movement trajectory using various recovery methods considering mobility parameters [8]. The experiments showed that the effective methods of human movement trajectory recovery differ according to information such as the mode of movement and the user's activity zone.

Over the last few years, there has been an increasing number of studies on trajectory recovery using deep learning. In [9], Wang et al. proposed subseq2Seq, which uses Kalman filter calibration for motion trajectory recovery and considers both temporal and spatial attention. Xia et al. proposed a new model for trajectory recovery based on the neural network using attention [10]. Ren et al. proposed a Seq2Seq-based model for movement trajectory recovery because it uses information on road segments and trajectories [11]. The model implements map matching and recovery to a high sample rate in an end-to-end manner. Sun et al. proposed a model for trajectory recovery using a Graph Neural Network for graphs constructed from movement trajectories [12]. The model employs an attention layer which considers the multilevel periodicity and shifting periodicity of human mobility, respectively.

Most trajectory recovery studies deal with motion trajectory recovery with sample rates of several seconds to a few minutes, as in datasets such as Geolife dataset. For example, taxis usually report GPS locations every 2 ~ 6 minutes to reduce the energy consumption of communication [13]. Therefore, the above methods are helpful for movement trajectory recovery for such data sets. We address movement trajectory recovery for Twitter in this study, having an even lower sample rate.

3. Methodology

In this study, we use the Seq2Seq model because of the estimation of the user's movement. The problem of estimating user movement is similar to that of machine translation. For example, the Seq2Seq model used in machine translation can translate an English sentence into a Japanese sentence. This study translates the original sequence of tweets to produce a recovered sequence.

We denote the sequence of tweets annotated with latitude and longitude information by $S = \{(t_1, t_2, \cdots, t_n)\}$, where $t_i = (lat, lng, time)$, $\forall i$, $1 \leq i \leq n$, which denotes the latitude, longitude, and timestamp. Also, t_i is each tweet in S. Next, we sort the sequence S in order of post time by ascending order.

Next, for each tweet t_i in S, we obtain quadtree keys (quadkey) [14]. quadkey is the world map into tiles (i.e., grids) of the same size at different levels of size and uses quadkey for the grid. We get quadkey of a particular level from the latitude and longitude of t_i. Each tweet t_i is then transformed into $t_i = (quadkey, time)$. If two consecutive quadkey in the sequence are the same, we remove the backward value. The result of quadkey sequence is $K = \{k_1, k_2, \cdots, k_m\}$, where k_i is quadkey, $\forall\, i, 1 \leq i \leq m$.

In this study, we implement a model that learns to recover the original sequence K from this partially deleted sequence $\widehat{K}$. We create a sequence $\widehat{K} = \{\widehat{k_1}, \widehat{k_2}, \cdots, \widehat{k_l}\}$, $\forall\, i$, $1 \leq i \leq l$, by randomly deleting keys in key sequence K.

Our model consists of an encoder and a decoder. The embedding layer of the encoder transforms each key $\widehat{k_\iota}$ of a sequence $\widehat{K}$ into a low-dimensional vector. We use a Gated Recurrent Unit (GRU) as the encoder. GRU is a variant of Long Short-term Memory networks, which is capable of learning long-term dependencies for sequential data without performance decay. The decoder consists of a dropout layer, an attention layer, and a GRU layer. We also use cross-entropy as the loss function of our model. The final model output is $\widehat{K}$ transformed into $K' = \{k'_1, k'_2, \cdots, k'_o\}, \forall\, i, 1 \leq i \leq o$.

4. Experimental Evaluation

4.1. *Experimental Setting*

In this experiment, we use tweets, which include latitude and longitude information. The tweets posted in Kyoto, Japan and were obtained from January 1, 2015, to December 31, 2015. Next, we divided the sequence of user tweets into daily segments. We extracted only those sequences with more than five quadkeys. As a result, the number of tweet sequences is 11,050.

We randomly deleted 20% of each key sequence because we wanted to create sequences with low sample rates. We also split the set of key sequences into a training set, a validation set, and a test set in the ratio 0.90 : 0.05 : 0.05. We set the zoom level of the quadkey to 15. A tile at this level has a side of approximately 1,222 meters. The resulting number of quadkey tiles is 386.

4.2. *Evaluation Criterion*

In this experiment, we use four criteria. The first and second are the mean absolute error (MAE), and the root mean square error (RMSE) between predicted values K' and ground truth K. MAE and RMSE are formulated as:

$$\boldsymbol{MAE = \frac{1}{n}\sum_{i=1}^{n}\left|dis\left(k_i, k'_j\right)\right|}$$

$$\boldsymbol{RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} dis(k_i, k'_i)^2}}$$

where $dis(k_i, k'_i)$ is the geodesic distance (meter) between k_i and k'_i. The smaller values of MAE and RMSE indicate the distance between the original sequence and the predicted sequence is close. The third and last is the recall and precision formulated as:

$$Recall = \frac{|K \cap K'|}{|K|}, Precision = \frac{|K \cap K'|}{|K'|}$$

Here, $|K \cap K'|$ is defined by the value of the longest common subsequence of K and K'. The larger recall and precision values indicate that methods predict the tiles more accurately.

4.3. *Evaluation Result*

Table 1. Evaluation result.

MAE	RMSE	Recall	Precision
1674.642	2447.872	0.482	0.686

Table 1 shows the evaluation result. As the MAE and RMSE values indicate, the difference in distance between the predicted results and the ground truth sequence is insignificant. Also, compared to Precision, Recall is small because it sometimes fails to recover tiles contained in the original sequence. However, each evaluation criterion indicates that the proposed method is somewhat effective. Improving these performances is a future challenge.

5. Conclusion

This paper deals with predicting Twitter users' moves using the Seq2Seq model. The model performed trajectory recovery by learning pairs of missing and true movement sequences.

Future work includes the following topics One is to improve the structure of the model. The model used here is a simple Seq2Seq model, which does not adequately represent the data. The second is to improve the information used by the model. The current model only considers the order of positions in learning Twitter sequences, but it could consider the text of tweets and time intervals.

Acknowledgments

This work was supported by JSPS KAKENHI Grant Number JP19K20418.

References

1. H. Zhang, H. Wu, W. Sun, and B. Zheng, “Deeptravel: a neural network based travel time estimation model with auxiliary supervision,” arXiv preprint arXiv:1802.02147, 2018.
2. M. Li, P. Tong, M. Li, Z. Jin, J. Huang, and X.-S. Hua, “Traffic flow prediction with vehicle trajectories,” in Proceedings of the AAAI Conference on Artificial Intelligence, 2021, vol. 35, pp. 294–302.
3. Y. Cui, H. Sun, Y. Zhao, H. Yin, and K. Zheng, “Sequential-knowledge-aware next POI recommendation: A meta-learning approach,” ACM Transactions on Information Systems (TOIS), vol. 40, Art. no. 2, 2021.
4. D. Yao, G. Cong, C. Zhang, and J. Bi, “Computing trajectory similarity in linear time: A generic seed-guided neural metric learning approach,” in 2019 IEEE 35th international conference on data engineering (ICDE), 2019, pp. 1358–1369.
5. Twitter, https://twitter.com

6. I. Sutskever, O. Vinyals, and Q. V. Le, “Sequence to sequence learning with neural networks,” Advances in neural information processing systems, vol. 27, 2014.
7. Z. Chen, H. T. Shen, and X. Zhou, “Discovering popular routes from trajectories,” in 2011 IEEE 27th International Conference on Data Engineering, 2011, pp. 900–911.
8. S. Hoteit, S. Secci, S. Sobolevsky, C. Ratti, and G. Pujolle, “Estimating human trajectories and hotspots through mobile phone data,” Computer Networks, vol. 64, pp. 296–307, 2014.
9. J. Wang, N. Wu, X. Lu, W. X. Zhao, and K. Feng, “Deep trajectory recovery with fine-grained calibration using kalman filter,” IEEE Transactions on Knowledge and Data Engineering, vol. 33, Art. no. 3, 2019.
10. T. Xia et al., “AttnMove: History Enhanced Trajectory Recovery via Attentional Network,” Proceedings of the AAAI Conference on Artificial Intelligence, vol. 35, Art. no. 5, May 2021.
11. H. Ren et al., “MTrajRec: Map-Constrained Trajectory Recovery via Seq2Seq Multi-task Learning,” in Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, 2021, pp. 1410–1419.
12. H. Sun, C. Yang, L. Deng, F. Zhou, F. Huang, and K. Zheng, “PeriodicMove: Shift-Aware Human Mobility Recovery with Graph Neural Network,” in Proceedings of the 30th ACM International Conference on Information & Knowledge Management, 2021, pp. 1734–1743.
13. J. Yuan, Y. Zheng, C. Zhang, X. Xie, and G.-Z. Sun, “An Interactive-Voting Based Map Matching Algorithm,” in 2010 Eleventh International Conference on Mobile Data Management, 2010, pp. 43–52.
14. quadkey, https://learn.microsoft.com/en-us/azure/azure-maps/zoom-levels-and-tile-grids

Authors Introduction

Dr. Masaharu Hirota

He received a Doctor of Informatics degree in 2014 from Shizuoka University. After working for the National Institute of Technology, Oita College, he has worked in the Faculty of Informatics, the Okayama University of Science, since April 2017. His research interests include data engineering, GIS, and social media. He is a member of ACM, DBSJ, and IPSJ.

Interactive Beating Drum Unity Game

Chung-Wen Hung* ,Cheng-Lung Ko
Department of Electrical Engineering, National Yunlin University of Science & Technology, 123 University Road,
Section 3, Douliou, Yunlin 64002, Taiwan, R.O.C.

Wen-Huei Chou
Department of Digital Media Design, National Yunlin University of Science and Technology, 123 University Road,
Section 3, Douliou, Yunlin 64002, Taiwan, R.O.C.
E-mail: m11012073@yuntech.edu.tw, wenhung@yuntech.edu.tw, cris@yuntech.edu.tw

Abstract

Interactive Beating Drum Unity Game is proposed in this paper. The pandemic forced individuals to maintain a prescribed social distance from others, and this way may ignore the individual's mental health and social needs, especially for the elderly. The Unity Engine is used to implement the system to provide a sport and social plane. And Bluetooth Low Energy (BLE) chip is adopted to transmit the beating signal to Unity Engine what following the Bluetooth (BT) protocol. Then the position beat by users will be showed on the user interface tablet. There are two play mode be implemented: single player mode and multiplayer mode, and the latter will include social function. The users are required to drum in short time interval accurately, and the game award points are evaluated for accuracy. The experimental results show the system workable, and the subjects unanimously stated that the operation of the game Easy-to-Use.

Keywords: Interactive Game, Unity, BLE

1. Introduction

While the World Population Prospects, United Nations report, mentioned the percentage of world population aged that over 65 years old will increased from 9% to 16% between 2019 and 2050. That is, population ageing would become a development trend in future years. The geriatric health make it a key factor to evaluate whether they can contribute for family or community. The physiological degeneration、memory loss and poor physical coordination affecting the daily life as a result of the aging process. So, to improve quality of life, postpone senility is the key. To best our knowledge, the first international congress on gerontechnology in August 1991. Trying to improve the elder's daily life and growing elderly population [1]. In recent years, gerontechnology most used as an adjunct to have continued normal life. For example, health monitoring for wearable device, enable basic warning and prevention. Many papers propose that digital games can benefits improving the elder's wellness, enhancing social bond and generating a sense of happiness [2],[3],[4]. In addition, it considered an effective form of exercise [5].

The pandemic has accelerated the development of information technology and changed people's daily life. The technology makes people keep in touch without the region and time restriction. Moreover, it increases digital game and digital learning development opportunities. During the lockdown measures period, in the beginning of the epidemic, people have increased interest in games. Instead of online games or console games, portable gaming console market share has steady increased. It shows the potential how the games create the social relation.

2. System Architecture

The system architecture is showed in Fig.1. The research use Laser engraving drum head, accomplishing the wireless transmission by Bluetooth Low Energy (BLE) and applying Unity on the tablet to establish the user interface. The applicion of sensing device can display the hitting position effectively. First, piezoelectric sensors convert mechanical energy into electrical energy when the drum is beat. The ADC in Bluetooth module receives the voltage generated by piezoelectric sensors and converts this signal to data in binary format. The module will send the position code to Unity by Bluetooth. In the end, Unity decodes the Bluetooth code and displays the position on the user interface. This rhythm game can not only help the elders enhance muscle strength of their upper limb while playing game, but also interactive with others who are impacted by the pandemic on Internet at home.

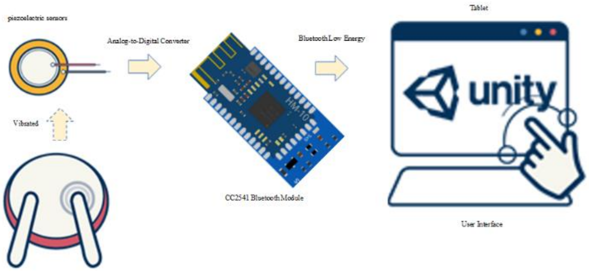


Fig.1 System architecture diagram

3. System Development

3.1 Unity

Unity is a cross-platform game engine that can support Windows, MacOS, Linux, iOS, Android, and etc. Besides games, it can develop variety levels of application[6]. The advantages of Unity including the visual window, the function of cross-platform develop and several source plugins. The method of building engine which depends on users' requirements can be divided into two parts, one is two dimension(2D) and another is three dimension(3D). Unity2D is adopted as our development environment because this paper refers the gameplay of Taiko Drum Master. We implement the unity in Android tablet to receive the signal which is processed from the Bluetooth module. The signal connects the game and decides the corresponding drop point on rhythm beats such as left side, left head, right head, right side, both head and both side. Unity collect and process these information, controlling user interface and responding on the tablet by coding. By receiving and integrating the interactive information, the user interface will complete the drumming interactive application.

3.2 Bluetooth Low Energy (BLE)

Bluetooth Low Energy(BLE) is developed by Bluetooth Special Interest Group (SIG). The performance of BLE are completely compared by [7]. BLE optimizes the transmission of small data and has been widely utilized on various occasions, for example, wearable device, automotive application[8],[9], medical environment and home automation[10],[11],[12],[13]. It uses Generic Attribute Profile (GATT) to send and receive application data. GATT encapsulate the Attribute Protocol (ATT) use service[14], characteristic and property to establish the link, to organize data, to read and to write data. Bluetooth SIG specification for BLE includes several definitions for common applications. However, users can define server and characteristic freely with their requirement of application and data structure. This method is suitable for sensor devices that only need to transmit data status. The user interface of proposed system receives the Analog-to-Digital Converter(ADC) sampling signal from sensor by using the Bluetooth module.

3.3 Piezoelectric sensors

Piezoelectric sensor is made of a piezoelectric material which is a kind of transduction element[15]. We install piezoelectric sensors under the drum head, determining the position of the drum by using the piezoelectric characteristics of component. When there is any external power force to the component, its internal electric dipole moment will be shortened by the pressure and generate voltage in order to resist the change caused by external pressure. If users need to generate electrical energy continuously, they will need to vibrate the sensor. Eq. (1) describe the effect process of piezoelectric that converts mechanical energy into electrical energy. In this paper, we divide the beat position into four, including left side, left head, right head and right side to avoid the vibration effect of beating. Fig.2 shows the method of energy converting of piezoelectric elements.

$$P = d\sigma \quad (1)$$

Where P is the polarization intensity of crystal, d is the piezoelectric constant, σ is the stress.

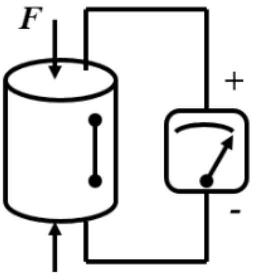


Fig.2 Inverse piezoelectric effect

3.4 Operational Amplifier (OPA)

3.4.1 Voltage follower

Voltage follower is an application of an Operational Amplifier (OPA). We design a high input impedance and a low output impedance for the OPA in the proposed system. Because the gain value is set to 1, output voltage will equal to input voltage. Obviously, the OPA reduce the source impedance without loading effect and reducing sample rate. Finally, we implement a RC filter circuit in the output of piezoelectric sensor to remove the wave noise.

3.5 Analog-to-Digital Converter (ADC)

We use the 12bits ADC in CC2541 Bluetooth module to sample the time domain data. To convert data from continuous-time to discrete-time and to present data as binary coded form are the characteristics of ADC [16]. In our application, we set a threshold range for ADC value. When the converted voltage exceeds the threshold range, Bluetooth will determine where the position you beat and then we will adjust the ADC threshold to avoid wrong measurement because four position affect each other.

4. Result

This paper Unity game engine is used to develop our game content. This software can be used for cross platform development and supports Windows, MacOS, Linux, IOS, Android and other console games. Samsung Galaxy Tab A7 with Unity application is used to be our development platform. The device we proposed in this paper can not only drumming but also interactive with each other on the internet. First, we simplify the gaming operation to reduce the learning burden. Drumming is a low intensity exercise, so it is suitable for elder to increase the strength. Moreover, the behavior of online interaction helps the elder maintain their social connection regardless of time and space.

To combine internet of things technology and sensors, CC2541 Bluetooth module is selected in this paper. The drum design retains the same as general drum, including drum body and sticks. We use the piezoelectric sensors in the drum head to detect the vibration. The sensors are small enough that can hide in the interlayer between the drum head and the bottom to reduce the signal influence while user drumming. When the sensor is pressed by the external force, the piezoelectric element will generate positive output voltage in order to reduce the change caused by the external stress. Fig.3-Fig.6 are the voltage wave form while piezoelectric sensors being pressed. As a result, the voltage will increase with vibration. Also, we applied OPA in order to reduce the loading effect, and filtered the output. At the end, sampling the time domain data by ADC with our Bluetooth module. The Bluetooth module will read the ADC voltage information of the beat, and determine which part of drum be struck. Fig.7-Fig.10 are the RC filtered voltage waveform.

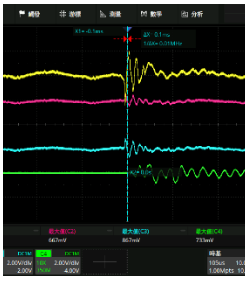


Fig.3 Left side voltage waveform

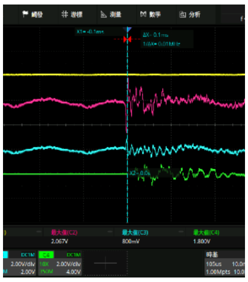


Fig.4 Left head voltage waveform

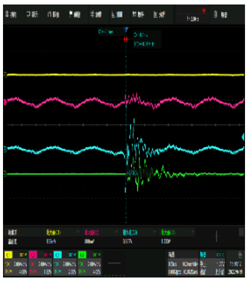


Fig.5 Right face voltage waveform

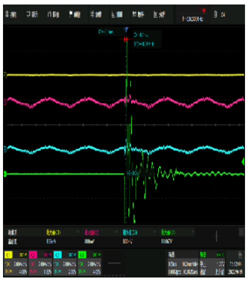


Fig.6 Right side voltage waveform

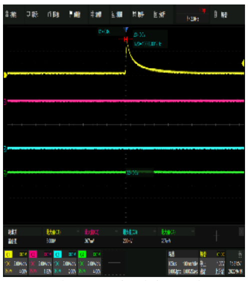


Fig.7 Left side voltage waveform

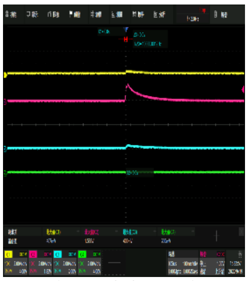


Fig.8 Left face voltage waveform

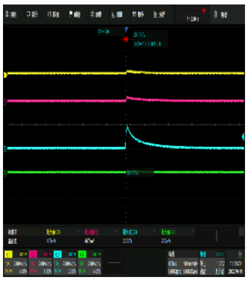


Fig.9 Right face voltage waveform

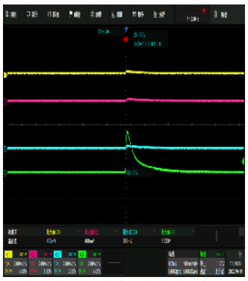


Fig.10 Right side voltage waveform

Different materials of drum head such as traditional taiko drum, Electronic drum with silencer drum head and the drum head are made of laser engraving in Fig.11 has been test in our research. In order to adapt to home environment and consider the accuracy of beating feedback, the drum head made of laser engraving has been chosen. We change the interlayer from foam in Fig.12 to hollow in Fig.13. In Fig.13 interlayer, we lock the screws to decrease the bottom area of contact. The piezoelectric sensors are installed between the drum head and drum bottom in Fig.14 that can identify the position are beat.

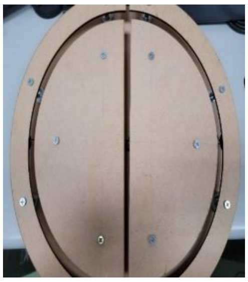


Fig.11 Laser engraving

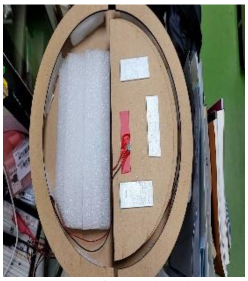


Fig.12 Left head voltage waveform

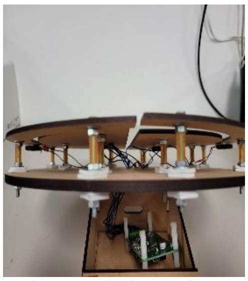


Fig.13 Right face voltage waveform

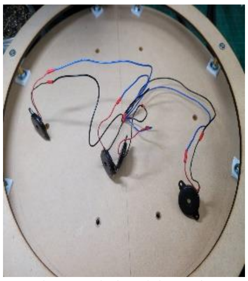


Fig.14 Right side voltage waveform

5. Conclusion

In this paper, a simple operation interface is used to reduce elders' learning burden. Elders can play online game with other players though the connection between Unity and BLE technology even at home. With this simple playing way, elders can not only prevent Sarcopenia but also communicate with others at home.

References

[1] Bouma, Herman; Graafmans,&; Jan A.M. (1992) Gerontechnology. Amsterdam: IOS press.

[2] Ijsselsteijn, W., Nap, H. H., de Kort, Y., & Poels, K. (2007, November). Digital game design for elderly users. In Proceedings of the 2007 conference on Future Play (pp. 17-22).

[3] Theng, Y. L., Dahlan, A. B., Akmal, M. L., & Myint, T. Z. (2009, April). An exploratory study on senior citizens' perceptions of the Nintendo Wii: the case of Singapore. In *Proceedings of the 3rd international convention on rehabilitation engineering & assistive technology* (pp. 1-5).

[4] Gerling, K. M., Schulte, F. P., & Masuch, M. (2011, November). Designing and evaluating digital games for frail elderly persons. In *Proceedings of the 8th international conference on advances in computer entertainment technology* (pp. 1-8).

[5] Pyae, A., Joelsson, T., Saarenpää, T., Mika, L., Kattimeri, C., Pitkäkangas, P., ... & Smed, J. (2017). Lessons Learned from Two Usability Studies of Digital Skiing Game with Elderly People in Finland and Japan. Int. J. Serious Games, 4(4).

[6] P. P. Patil and R. Alvares, "Cross-platform Application Development using Unity Game Engine," *International Journal of Advance Research in Computer Science and Management Studies*, vol. 3, no. 4, pp. 19–27, 2015.

[7] J. Tosi, F. Taffoni, M. Santacatterina, R. Sannino, D. Form ica Performance evaluation of Bluetooth low energy: a systematic review Sensors, 17 (2017), p. 2898,

[8] Lin, J.-R.; Talty, T.; Tonguz, O.K. On the potential of bluetooth low energy technology for vehicular applications. *IEEE Commun. Mag.* **2015**, *53*, 267–275.

[9] Xia, K.; Wang, H.; Wang, N.; Yu, W.; Zhou, T. Design of automobile intelligence control platform based on Bluetooth low energy. In Proceedings of the 2016 IEEE Region 10 Conference (TENCON), Singapore, 22–25 November 2016; pp. 2801–2805.

[10] Sherratt, R.S.; Janko, B.; Hui, T.; Harwin, W.; Diaz-Sanchez, D. Dictionary memory based software architecture for distributed Bluetooth Low Energy host controllers enabling high coverage in consumer residential healthcare environments. In Proceedings of the 2017 IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, USA, 8–10 January 2017; pp. 406–407.

[11] Collotta, M.; Pau, G. A Novel Energy Management Approach for Smart Homes Using Bluetooth Low Energy. *IEEE J. Sel. Areas Commun.* **2015**, *33*, 2988–2996.

[12] Collotta, M.; Pau, G. A Solution Based on Bluetooth Low Energy for Smart Home Energy Management. *Energies* **2015**, *8*, 11916–11938.

[13] Collotta, M.; Pau, G. An Innovative Approach for Forecasting of Energy Requirements to Improve a Smart Home Management System Based on BLE. *IEEE Trans. Green Commun. Netw.* **2017**, *1*, 112–120.

[14] Townsend, K.; Cufi, C.; Wang, C.; Davidson, R. Introduction. In *Getting Started with Bluetooth Low Energy: Tools and Techniques for Low-Power Networking*; Sawyer, B., Loukides, M., Eds.; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2014; pp. 1–14.

[15] Gautschi, Gustav. "Piezoelectric sensors." *Piezoelectric Sensorics*. Springer, Berlin, Heidelberg, 2002. 73-91.

[16] Walden, Robert H. "Analog-to-digital converter technology comparison." *Proceedings of 1994 IEEE GaAs IC Symposium*. IEEE, 1994.

Authors Introduction

Prof. Chung-Wen Hung

Chung-Wen Hung 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.

Mr. Cheng-Lung Ko

He is a graduate student at department of Electrical Engineering, National Yunlin University of Science & Technology.

Wen – Huei Chou

Wen-Huei Chou is a professor at the Digital Media Design department in National Yunlin University of Science and Technology in Taiwan. She holds a BA in Art education in National Hsinchu Education University, an MA in Applied Art (major in Design) in National Chaio Tung University, and a doctorate degree in Design in Swinburne university of Technology. She dedicates herself in the field of design research; especially in the innovation and integration in new media deign domain. Her current research interests include innovative and transdisciplinary design research, social and service design for social practice.

Error Backpropagation Neural Network Based Image Identification for a Foot Massage Machine and Its Mechanism Design

Chun-Chieh Wang
National Yunlin University of Science and Technology, 123 University Road, Section 3, Douliou, Yunlin 64002, Taiwan

E-mail: bigjasonwang@gmail.com
https://www.yuntech.edu.tw/

Abstract

In the past twenty years, many companies have developed different styles of foot massage machines. At present, the common massage products on the market include roller type and pressing type. However, it is very difficult to accurately stimulate all acupuncture points for different sizes of feet. Besides, the massage roller cannot be controlled independently. Therefore, a novel computer vision technology is proposed to identify the foot acupuncture points by error backpropagation neural network (EBNN) in this paper. First, we use cameras to capture the sole of users' soles and execute image preprocessing procedures to segment the region of interest (ROI) of soles. We map foot acupuncture points to foot images to obtain reference massage positions. Second, the YCbCr color space is used to separate the brightness to complete the segmentation of the foot image in the skin detection. Moreover, EBNN is used to train users' soles-image sets to improve the success rate of image segmentation. Finally, to improve the rate of image recognition and user convenience, a foot massage machine was redesigned. Experimental results validate the superiority and practicality of the proposed image identification method for foot massage machines.

Keywords: Error backpropagation neural network (EBNN), Foot massage machines, Image identification

1. Introduction

Dr. Fitzgerald proposed the concept of foot reflexology massage in 1913 [1]. It wasn't until the 1970s that foot massage techniques received the attention of the medical community. Since ancient times, the concept of acupuncture points corresponding to body organs is used in the traditional Chinese medicine therapy. Therefore, to combine the concept of reflexology and acupuncture points, many companies have developed different styles of foot massage machines in the past two decades. The current products on the market contain two kinds of forms of massage, roller-type and push-type.

The rotate-type massage machines can fit different feet, but the rollers are limited in certain massage areas. Moreover, for the push-type massage machines, it is difficult to accurately stimulate all acupuncture points for different foot shapes. Besides, the massage roller cannot be controlled independently. Therefore, for the research of foot massage machines, how to find the corresponding acupuncture points will be the most problem to be overcome.

The foot reflexology map that is most often referred to in the development process of the foot massage machine [2], as shown in Fig. 1. The author hopes to design a foot massage machine that can give corresponding force to different acupuncture points. However, the existing foot massage machines are usually designed for the foot shape of adults. Not only that, but most of the machines cannot accurately press all the acupuncture points.

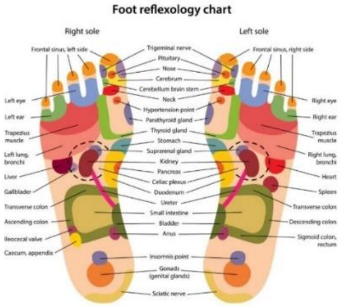

Foot reflexology chart
Right sole
Left sole
Frontal sinus, left side
Trigeminal nerve
Frontal sinus, right side
Pituitary
Nose
Cerebrum
Cerebellum brain stem
Left eye
Neck
Right eye
Hypertension point
Left ear
Parathyroid gland
Right ear
Trapezius muscle
Thyroid gland
Trapezius muscle
Stomach
Suprarenal gland
Left lung, bronchi
Right lung, bronchi
Kidney
Liver
Pancreas
Heart
Celiac plexus
Gallblader
Duodenum
Spleen
Transverse colon
Ureter
Transverse colon
Small intestine
Ascending colon
Bladder
Descending colon
Ileocecal valve
Anus
Sigmoid colon, rectum
Caecum, appendix
Insomnia point
Gonads (genital glands)
Sciatic nerve

Fig. 1. Foot reflexology chart [2]

To overcome the above problems, the author used EBNN to detect foot acupuncture points. The key steps consist of three parts. First, a camera is used to capture the sole of the users' soles. An image preprocessing procedure is performed to segment ROI of the sole. To get the position of the reference massage, the foot acupuncture points are mapped onto the foot image. Second, to complete the segmentation of the foot image, the YCbCr color space is used in skin detection to separate the brightness. Furthermore, the shoe sole image set is trained using EBNN to improve the success rate of image segmentation. Finally, to realize all acupuncture points massage function, an FPGA/DSP embedded system will be used for image processing and multi-axis PWM control. The results verify the superiority and practicality of the proposed image detection method for foot massage machines.

2. Hardware architecture design

2.1. *Mechanism design*

To improve the rate of image recognition and user convenience, a foot massage machine is re-designed, shown as Fig. 2. In our design, the foot massage machine is portable. Fig. 2(a)-(b) shows the drawbars pull up and drop down, respectively.

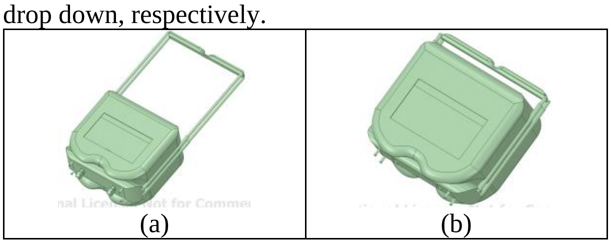

(a) (b)

Fig. 2. Foot massage machine body 3D design

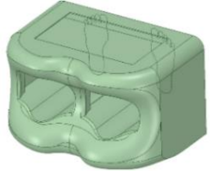


Fig. 3. Holder of ankles and feet

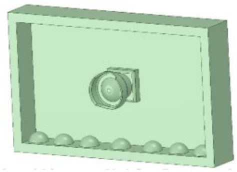


Fig. 4. Image capture system

The overall design includes four parts: a holder of ankles and feet (shown as Fig. 3), an image capture system (shown as Fig. 4), a massage institution (shown as Fig. 5), and embedded electronic control devices.

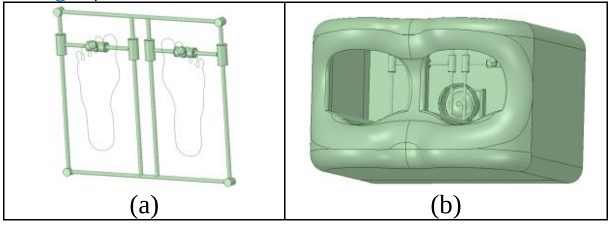

(a) (b)

Fig. 5. Massage institution

The holder is designed to meet the general adult feet. Fig. 5(a) shows the massage buttons and x-y axis connecting bars. The x-y axis connecting bars drive by servo motors. The massage buttons can reach all foot acupuncture points by x-y axis connecting bars. The foot massage machine assembly is shown as Fig. 5(b). Moreover, to facilitate the implementation of image recognition system, some special airbags are used to fix the different size feet. After above procedures, we use a DSP-FPGA hybrid embedded system to get the coordinates of the plantar acupuncture points. At the same time, the embedded system drives a proposed mechanism in order to make precise compressions on these acupuncture points. In addition, an LCD module (shown as Fig. 6) will show the location of acupuncture points, treatment time, and the related physical information.

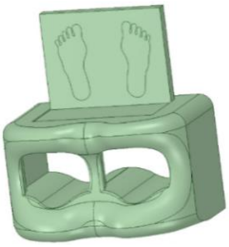


Fig. 6. LCD module

2.2. *DSP-FPGA hybrid embedded system*

In this paper, the DSP-FPGA hybrid embedded system is proposed. FPGA is used for video encoding and image preprocessing calculations. DSP is used for acupuncture points mapping and massage mechanism controlling. The architecture is shown in Fig. 7.

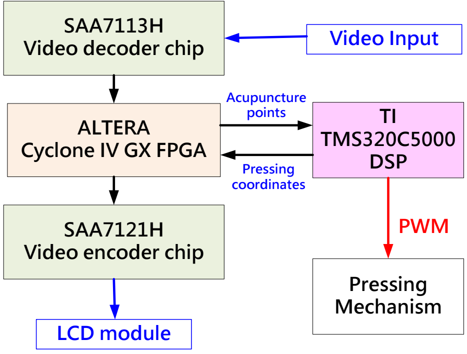


Fig 7. DSP-FPGA Hybrid Embedded System

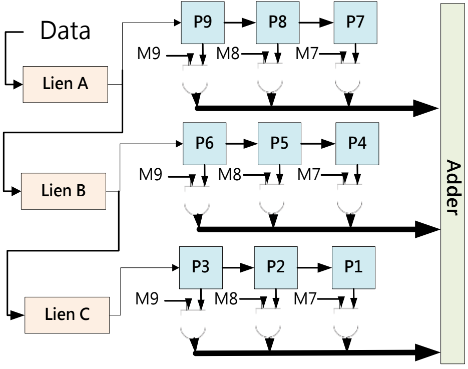


Fig. 8. 3×3 Mask Convolution Operation

A 3 by 3 masked convolution operation is used for edge detection. Data extraction [3] is shown in Fig. 8. In this paper, line buffer is used to fetch row data to reduce the register requirement. Furthermore, the length of line buffer can be regulated for different image size.

3. Acupuncture points detection via computer vision

3.1. *Foot reflexology image segmentation*

Color model and brightness compensation are the two main processes of the foot reflexology image segmentation. The YC_bC_r color model is used to achieve plantar reflexology image segmentation [4]. (Y is the brightness, C_b and Cr are blue and red intensities, respectively.) Besides, given that light compensation is still not done for most images. For this reason, Zhang [5] proposed an adaptive brightness compensation method. The formula is as follows.

$$L_c = \frac{H_{std}}{\mathrm{D}_{avg}} \tag{1}$$

where

$$H_{std} = \frac{\sum_{1}^{i}\left[\max(V_R,V_G,V_B)+\min(V_R,V_g,V_b)\right]}{2\times j} \tag{2}$$

$$j = i - \sum_{1}^{i}(V_R = V_G = V_B = 0)$$

H_{std} can be changed with different images. i is the amount of pixels in an image. j is the number of non-black pixels in the image. V_R, V_G and V_G are the red, green, and blue components of the pixel, respectively. D_{avg} is the non-black pixels of V_R, V_G and V_G. In addition, to obtain training data, EBNN [6][7][8][9] is used. The EBNN segmentation process is shown in Fig. 9.

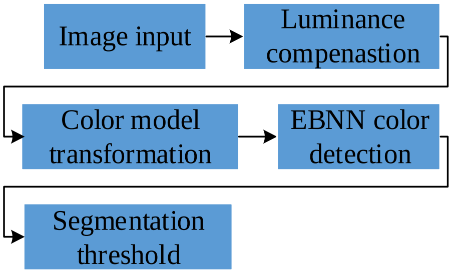


Fig. 9. EBNN Segmentation Process.

The algorithm of EBNN is described in the following steps.

Step 1. Set all the weights and threshold levels of the neural network so that the random numbers can be evenly distributed over a small range.

【Feedforward】

Step 2. Deliver the training data (YCbCr) to the input node of the training cycle.

Step 3. Calculate the value of hidden nodes.

Step 4. Calculate the value of output nodes.

【Backpropagation of the network error】

Step 5. Compares the actual value of each output node with the desired output value and calculates the squared error for each output node. (For each output node, send the squared error back to the hidden node.)

Step 6. Adjusts the weights linked to each output node based on the calculated squared error for each output node.

Step 7. The weights on the links to each hidden node are adjusted according to the squared error from the output node.

Step 8. When "stop condition" is false, return to step 2.

3.2. *Foot image obtaining and edge detection*

To reduce computation time, the ROI of the plantar image is segmented. Futhermore, to find out the mapping range of acupuncture points, the edges of the foot image need to be calculated. Because the foot images are all curves, the detection results will produce many discrete segments. Therefore, LoG (Laplacian of Gaussian) operator is used for curves edge detection in this paper.

Second-order Laplace function is

$$\nabla^2 L = \frac{\partial^2 L}{\partial x^2} + \frac{\partial^2 L}{\partial y^2} \tag{3}$$

and Gaussian function is

$$G(r) = -e^{\frac{d^2}{2\sigma^2}} \tag{4}$$

where $d^2=x^2+y^2$, σ is the standard deviation. Laplacian of Gaussian operator is as follows.

$$\nabla^2 LoG(r) = -\left[\frac{d^2-\sigma^2}{\sigma^4}\right]e^{-\frac{d^2}{2\sigma^2}} \tag{5}$$

3.3. *Acupuncture points detection*

The projective transformation of a two-dimensional plane [10] can be represented by a homogeneous vector as shown below.

$$\begin{bmatrix} a'_1 \\ a'_2 \\ a'_3 \end{bmatrix} = \begin{bmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{bmatrix}\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \tag{6}$$

Assume that the corresponding non-homogeneous coordinates of two points on the plane are (a, b) and (a ', b'). Then (a', b') can be calculated by Eq. (6), as shown in Eq. (7).

$$a' = \frac{a'_1}{a'_3} = \frac{v_{11}a+v_{12}b+v_{13}}{v_{31}a+v_{32}b+v_{33}}$$

$$b' = \frac{a'_2}{a'_3} = \frac{v_{21}a+v_{22}b+v_{23}}{v_{31}a+v_{32}b+v_{33}} \tag{7}$$

Solving equation (7) requires at least six corresponding points. Thus, standard foot acupuncture points can be mapped to the user's foot.

4. Experimental results

In the paper, the foot image is cut into three mapping ranges. Range 1: Range 2: Range 3=3.5:3:3.5 (Fig. 10). Range 1, Range and Range 3 cover 12, 7 and 4 main acupuncture points respectively.

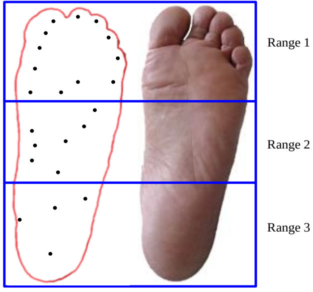
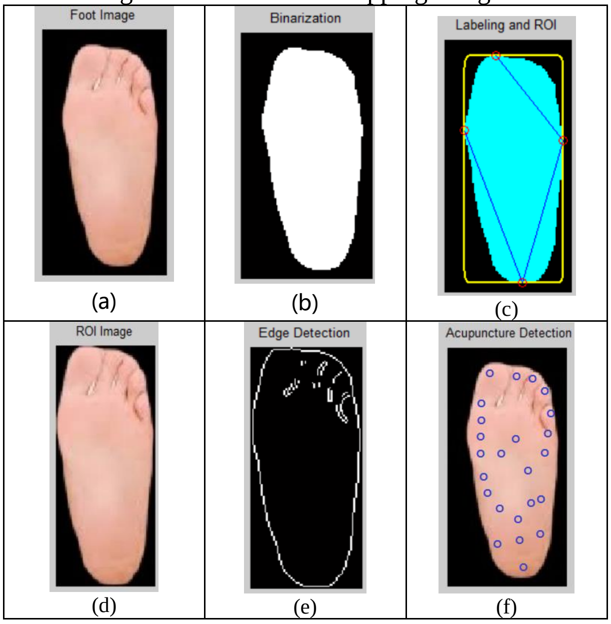


Fig. 10. Definition of Mapping Ranges



(a) (b) (c)

(d) (e) (f)

Fig. 11. Results of Image Processing and Acupuncture Points Mapping

Fig. 11 shows the results of image processing and acupuncture points mapping. Fig. 11(b) shows the skin color segmentation results of EBNN. The results verify that the foot image segmentation is clear.

5. Conclusion

In this paper, a novel computer vision technique is proposed to identify foot reflexology points via EBNN. A camera is used to capture the sole of the users' foot. Moreover, an image preprocessing procedure is performed to segment ROI of the soles. Foot reflexology points are mapped to foot images to obtain reference massage positions. In addition, the YCbCr color space is used to separate the brightness to complete the segmentation of the foot image in the skin detection. Moreover, EBNN is used to train the user's shoe sole image set to improve the success rate of image segmentation. Not only that, but a foot massage machine has been redesigned to improve the image recognition rate and user convenience. The experimental results verify the superiority and practicability of the proposed image recognition method for foot massage machines.

References

[1] Kaye A., Mathan DC (1980), Reflexology for Good Health, Wilshire Book Co.

[2] https://cloudmassage.com/blogs/news/the-main-pressure-points-on-our-feet-and-what-they-mean-a-basic-guide

[3] P.Y. Hsiao, C.H. Chen, H. Wen and S.J. Chen (2006), Real-time realisation of noise-immune gradient-based edge detector, IEE Proceedings on Computers and Digital Techniques, 153(4): 261-269.

[4] Zhang S, Jing X, Zhang B, Sun S (2010), An Adaptive Fingerprint Image Segmentation Algorithm Based on Multiple Features, Advanced Computer Control , pp.191-194.

[5] Zhang D, Zhoua ZH, Chen S (2005), Diagonal principal component analysis for face recognition, Pattern Recognition Society, 39(1):140-142.

[6] Rojas R. (1996), Neural Networks - A Systematic Introduction, Springer.

[7] Anthony M., Bartlett P. L. (2009), Neural Network Learning: Theoretical Foundations, Cambridge University Press

[8] Andrés B.C., Pedro J. G.L., José-Luis S.G. (2013), Neural architecture design based on extreme learning machine, Neural Networks, 48:19-24.

[9] L.G. Wright, T. Onodera, M.M. Stein, T. Wang, D.T. Schachter, Z. Hu, P.L. McMahon (2022), Deep physical neural networks trained with backpropagation, Nature, Vol. 601, 549-555.

[10] Hartley R, Zisserman A(2004), Multiple View Geometry in Computer Vision, 2nd Edition, Cambridge University Press.

Authors Introduction

Prof. Chun-Chieh Wang

He is a professor of National Yunlin University of Science and Technology. His areas of research interest include robotics, image detection, electromechanical integration, innovative inventions, long-term care aids, and application of control theory. He is now a permanent member of Chinese Automatic Control Society (CACS) and Taiwan Society of Robotics (TSR). He is also a member of Robot Artificial Life Society.

Cross-domain Sharing of Robots in the Community Caring and Practice of University Social Responsibility

Jia-Ming Hsiao
Bachelor Program in Intelligent Robotics, National Yunlin University of Science and Technology,
123 University Road, Section 3, Douliou, Yunlin 64002, Taiwan, R.O.C.
E-mail: hsiaojm@yuntech.edu.tw
https://irobot.yuntech.edu.tw/

Abstract

This paper focuses on community caring and social practice that assisting robot education development in rural schools to echo the two goals of the United Nations' SDGs (Sustainable Development Goals)–"Goal 4: Quality education" and "Goal 10: Reduce inequality". At the same time, in response to the scientific and technological literacy-oriented needs of the Curriculum Guidelines of 12-Year Basic Education in 2019, it is shown how to assist rural schools in Tainan City of Taiwan to develop computational thinking and cross-domain applications of robots under the COVID-19. Four robot education suites including LEGO Mindstorms EV3, Makeblock mBot, mBot2, and VEX IQ are applied in the rural schools. The dilemma and solutions of robot education in rural schools are also discussed in this article. Universities, non-profit organizations and robot equipment vendors work together to practice social responsibility.

Keywords: Rural Education, Robot education, University Social Responsibility

1. Introduction

The University Social Responsibility (USR) program was promoted by Ministry of Education (MOE) since 2017 to encourage colleges and universities to reflect on the content and goals of higher education in Taiwan[1]. The USR program had entered its second phase during 2020 and 2022. This project–"Rural Information Application Cultivation and Robot Cross-domain Sharing Project in Tainan City" was supported by 2020-2022 USR program to focus on community caring and social practice that assisting robot education development in rural schools to echo the two goals of the United Nations' SDGs (Sustainable Development Goals)[2]–"Goal 4: Quality education" and "Goal 10: Reduce inequality". On the other hand, in response to the scientific and technological literacy-oriented needs of the Curriculum Guidelines of 12-Year Basic Education in 2019, the lacking dilemma of teachers, amount of robot suite, teaching materials and technical supports is arising as well as the demands on developments of computational thinking and cross-domain applications of robots. It is more difficult during COVID-19 epidemic to teach and learn robot and programming in classroom. On-line learning is applied more and more to avoid the epidemic, however, without the robot suite, the learning effectiveness of robot operation and programming decreases.

In this paper, we describe the three lacking dilemmas of robot education in rural schools in Section 2. In Section 3, the solutions and corporation of universities, non-profit organizations and robot equipment vendors are discussed. Four robot education suites including LEGO Mindstorms EV3, Makeblock mBot, mBot2, and VEX IQ are applied in the rural schools. Two canonical examples are shown in Section 4. Conclusions are shown in Section 5.

2. The dilemma of robot education

There are three deficiencies in robot education in rural schools. They are lack of equipment/resources, lack of technology and lack of teachers.

2.1. *Lack of equipment/resources*

Due to the students of rural schools are less than the urban schools, the funding of robot equipment is not sufficient to meet the requirement of teaching and learning activities. The quantity of robot suite cannot meet class size. It is also difficult to choose which robot is suitable with the consideration of lesson preparation, applications and subsequent maintenance. After the lifetime cycle which may be three to five years, where is the source and how to arrange the funding of updating next generation robot are the other two important issues.

2.2. *Lack of technology*

In general, the rural school teachers lack for experiences of robot hardware design, sensors and motors application, control program developing, and coaching a team to joint robot competitions. The continuity problem of staying in the same rural schools enhances this lacking issue in Taiwan. On the other hand, the study or training courses are deficient for rural school teachers.

2.3. *Lack of teachers*

In spite of the continuity problem of staying and deficient study or training courses, insufficient time and enthusiasm from teachers is another factor for the lack of suitable teacher of teaching and applying robot in the class. The principal’s and colleagues’ support can influence the teacher’s willing to learn to teach and apply robot in and/or after class.

3. Solutions

3.1. *Robot education sources*

There are seven robot education sources including official funding, school experience exchange, robot vendors, robot classroom in the after-class education market, network communities, parents and family, and corporate/alumni sponsor.

3.1.1. *Official funding*

There are many official programs or projects every year for schools to apply. The funding of official programs or projects can support the purchase of robot suite and parts, and handling of study courses.

3.1.2. *School experience exchange*

The experience exchanges of different schools are very important. Different levels of school can share their robot teaching and applying experiences from each other. It is useful to share resources such as equipment, teaching material, technology, and project applying. The coaching experience to joint robot competitions is another interested issue.

3.1.3. *Robot vendors*

Robot vendors are the most important equipment support sources. They can not only make advice for the purchase of robots and lessons, but also can provide technical support. Some vendors may make discount or donate to rural schools for their Corporate Social Responsibility (CSR).

3.1.4. *Robot classroom in the after-class education market*

There are abundant teaching and robot competition experiences existing in the robot classroom of after-class education market. They can provide extensive materials for the teaching activities and robot competition.

3.1.5. *Network communities*

Network communities can provide channels for the introductions of robot equipment, technical discussion without the limitation of when and where.

3.1.6. *Parents and family*

With the parents’ support, teachers can develop in-class and after-class learning activities to explore the applications of robot. Some robot applications arise from living fields or needs, parents may provide the off-campus visit for the class. Parents and family need to accompany students to study after class to enhance the learning effect. During competition seasons, coach or teacher needs the parents’ help for the pick-up of students, even the technical supports.

3.1.7. *Corporate/alumni sponsor*

Another unofficial robot education source is the corporate or alumni sponsorship for funding or scholarship. The sponsor may provide funding for robot teaching activities and competition, or provide technical or labor support. With scholarship, the students could be encouraged to study robot applications, and to join further volunteer service.

3.2. *Strategy*

We employ the following four strategies to overcome the three lacks in the rural school robot education. They are learning bases, integration of resources, strive for sponsor of parents and alumni, and professional service learning volunteering.

3.2.1. *Learning bases*

As central learning bases, Far East University and National Yunlin University of Science and Technology provide the robots such as LEGO EV3, Makeblock mBot and mBot2, VEX IQ, teaching materials, and student volunteers to support the project. Some rural schools are chosen to be satellite learning bases to provide the local robot education opportunities. The learning bases already have robot suites with enthusiastic teachers. They can lead the rural schools around them to step into and try to develop robot education. Our USR project supply funding for the maintenance of robot and study courses. The learning bases also provide spaces for the robot study and competition preparation. Two facebook communities (Fig. 1, Fig. 2) are established to link rural and urban schools in Tainan and Yunlin in Taiwan.

台南校園機器人教育社群
私密社團 · 90 位成員
+ 邀請
瀏覽
管理
社群首頁
總覽
管理員工具
社群聊天室
+ 建立專區
編輯
mBot
台南校園機器人教育社群

Fig. 1: Tainan Robot Education Community

雲林縣偏鄉機器人教育推廣計畫-大眾科學教育
公開社團 · 26 位成員
+ 邀請
瀏覽
管理
社群首頁
總覽
管理員工具
社群聊天室
+ 建立專區
編輯
雲林縣偏鄉機器人教育推廣計畫-大眾科學教育

Fig. 2: Yunlin Robot Education Community

During COVID-19 epidemic, on-line learning is applied for classes. For robot learning, it is difficult to test the robot sensors, motors and mechanism without robot. Our USR project borrows robots to students for the robot learning activities.

3.2.2. *Integration of resources*

We invited robot vendors and non-profit organizations to join this USR project. The learning lessons, robot suite, and competition information are integrated to learning bases. We introduce robot vendors to the interested rural schools to plan for the possibility of implementing robot learning and applications. The learning bases and classroom in the after-class education market can work together to preparing robot competitions. Universities and the classrooms in the after-class education market contribute their experience and technic to lead the rural school teachers for coaching teams to join competitions.

3.2.3. *Strive for sponsor of parents and alumni*

The sponsors from parents and alumni like the two rocket boosters of space shuttle. They boost the development of robot education in rural schools. We support lessons and volunteers for satellite learning bases to create highlights of robot education. Then satellite learning bases can strive for official and unofficial project funding with these highlights. Also the highlights can be used to convince parents to encourage the students for joining the after-class learning club.

Universities and rural schools can enhance their relationship between them and local companies. Universities also can strive for the student scholarship and internship to achieve the object of cultivating talents.

3.2.4. *Professional service learning volunteering*

In order to achieve the object of cultivate talents, the student practice internship is valued gradually. However, the internship opportunity may not satisfy students. The professional service volunteering is integrated into some professional classes to encourage students to provide professional service for rural area or local communities. Via the professional service volunteering, students can provide professional assistance for the rural school robot education.

4. Canonical examples

4.1. *Tainan Municipal Houbi District Houbi Elementary School*

Houbi District is the north countryside of Tainan City, and is famous for rice and rural cultures. Houbi Elementary School is one of satellite learning bases and our most important partner for practicing USR project. It provides Saturday Robot Club activities with average frequency of one course per two weeks for the Xinying District and Houbi District. It is the central robot education of north Tainan City, Taiwan. There are 20 suites of LEGO EV3, 20 suites of mBot, and 20 suites of mBot2 to support club class and competitions. Students of Houbi Elementary School, as well as the outside students from other schools, can borrow robot for home exercise or summer robot competition season. The funding and robot are sponsored by alumni. The youth university Student volunteers form the special achievement class in robotics of National Yunlin University of Science and Technology provide their professional service for the robot lessons. We also provide basic maintenance and repair service for it so that saving the waiting time and cost.

4.2. *Tainan Municipal Cigu District Hougang Elementary School*

Fig. 3: FLL Junior challenge Activity in Hougang Elementary School

Cigu District is an area near the sea in Tainan City, and was famous for salt pans. Hougang Elementary School is also one of satellite learning bases and is the more active school for developing special classes in local caring with robot application. The teacher Kuo guide and coach students to joint 2021-2022, 2022-2023 FLL Junior challenge (Fig. 3) with the Cigu District local materials, and win the prizes. The teacher Kuo develops many special teaching materials and shares without any cost.

5. Conclusions

In this paper, we have discussed the dilemma and solutions of robot education in rural schools. The issue about rural school robot education is worthy of continued care and assistance. The university teachers and students should pay more effort on local care to realize the university social responsibility and to achieve the object of cultivating talents.

Acknowledgment

This work was supported by "The 2020-2022 USR program" hold by the Ministry of Education, Taiwan, Republic of China.

References

1. " 2021 USR ONLINE EXPO: University Social Responsibility", https://2021usrexpo.org/
2. "THE 17 GOALS - Sustainable Development Goals", https://sdgs.un.org/goals

Authors Introduction

Prof. Jia-Ming Hsiao

He received the B.S. and Ph.D. degree in Electrical Engineering from National Sun Yat-Sen University, Kaohsiung, Taiwan in 1994 and 2002, respectively. He is an Assistant Professor in the Bachelor Program in Intelligent Robotics at National Yunlin University of Science and Technology, Yunlin, Taiwan. He is also the 2022-2023 youth group vice presiding judge of mobile robotics at national skills competition in Taiwan. His current research interests include system and control theory, internet of things, mobile robotics and artificial intelligence.

Research on Design of Implementation Mechanism for Similar Production Line

Chia-Nan Ko
Department of Automation Engineering, Nan Kai University of Technology, Taiwan

Yi-Yu Li
Department of Automation Engineering, Nan Kai University of Technology, Taiwan

Ting-Ru Ko
Department of Automation Engineering, Nan Kai University of Technology, Taiwan

Ting-Yi Chen
Department of Automation Engineering, Nan Kai University of Technology, Taiwan
E-mail: t105@nkut.edu.tw

Abstract

Recently, enterprises are actively investing in the research and development of practical technologies for similar production line organizations. Faced with various technological innovations in intelligent automation. In this paper, the object is to research and integrate intelligence through practical functional solutions to design and establish a production line implementation mechanism. We use a myRIO controller as the main core to plan the factory automation production line structure through the practical function research for building similar production line implementation organizations. In this paper, combining theoretical methodology and practical experiments to construct institutional design, electromechanical system and intelligent automation procedure based on LabVIEW2018. Evaluate the efficiency of the design automated production lines to improve the capabilities of similar production lines.

Index Terms: Similar Production Line, Intelligent Automation, myRIO, Image Recognition

I. INTRODUCTION

In recent years, the application of intelligent automation production has become more and more extensive, and intelligent automation education has gradually been valued and promoted in the field of factory automation. The intelligent automation industry is growing year by year in Taiwan, and it is also testing the new education methods of the industry [1]. Intelligent automation education not only inspires student's interest and creativity in science and technology, but also cultivates cross-field practical talents in intelligent automated production. With the promotion of Productivity 4.0, the industry has transformed to seek intelligent automated production mode, so that factory mass production and warehouse management can replace a large number of human resources through artificial intelligence, improve production efficiency, reduce operating costs, and improve industrial competitiveness [2]. The intelligent automation education is the Ministry of Education continues to promote science and technology education syllabus, from primary school to cultivate basic education of information technology. Then cultivate robot practical technology and application talents, in the face of the future high-tech society gradually popularized various scientific and technological thinking [3].

The NI myRIO platform has been listed as the main core controller of the robot platform in various world competitions in recent years. Intelligent automation is a highly cross-field discipline, but also the integration of precision machinery, electrical electronics, information communication, automation and other academic theories and application technologies, its industrial correlation is high (such as IC industry, service industry, manufacturing), with locomotive industry characteristics, a wide range of applications, for intelligent life value, improve the quality of life, and automation academic and technology application development is one of the most important indicators of national competitiveness [4].

II. Methods

A. *myRIO Controller*

The controller uses the main core of the myRIO controller, with Tetrix_Robotics and Metrix Base Set metal kit, suitable for intelligent automation product line. The system is designed with NI LabVIEW2018 program as the design platform, allowing users to more easily integrate peripheral motors, sensor control and metal kit creative assembly, and then match NI LabVIEW software writing program can free users from the complexity of traditional general programming writing, programming writing is similar to C language, the only difference is that LabVIEW will programming, algorithms, visual development, measurement tools, communication mode, image processing in one, the function is relatively powerful, LabVIEW is different from other programming languages writing, easier for users to do development integration.

This article develops intelligent automation production line using LabVIEW2018 (Laboratory Virtual Instrumentation Engineering Workbench) program platform, LabVIEW is a graphical program compilation platform developed by National Instrument. LabVIEW2018 consists of three parts: Block Diagram, Front Panel, and Icon/Connector. LabVIEW has programming, programming and interface design functions, which can design the functions of the robot as a whole, and the diagram-controlled drag-and-drop design is easier for beginners without programming. When the software enters the LabVIEW window as shown in Fig. 1.

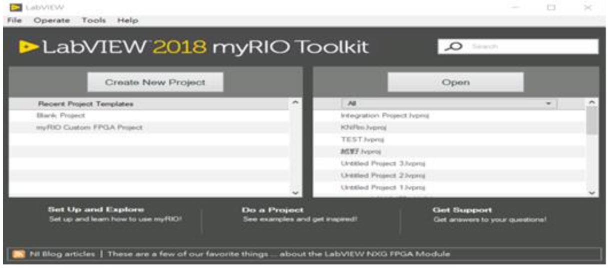

LabVIEW
File
Operate
Tools
Help
LabVIEW 2018 myRIO Toolkit
Search
Create New Project
Open
Recent Project Templates
Blank Project
myRIO Custom FPGA Project
All
Integration Project.lvproj
KNRm.lvproj
TEST.lvproj
Untitled Project 3.lvproj
Untitled Project 2.lvproj
Untitled Project 1.lvproj
Set Up and Explore
Set up and learn how to use myRIO!
Do a Project
See examples and get inspired!
Get Support
Get answers to your questions!
NI Blog articles | These are a few of our favorite things ... about the LabVIEW NXG FPGA Module

Fig. 1 VIEW 2018 window

The myRIO adapter board design allows the robot architecture to be easily combined with various controllers, which can provide various interfaces such as DC motor, RC motor, sensor and power supply, and the myRIO controller can quickly connect to various controllers.

B. *Image Recognition*

Using the Logitech C310 network camera and Microsoft LifeCam Cinema camera, the recognition of object graphics allows the robot to distinguish between similar graphics and colors, and develops a human-machine interface to immediately present the recognition results.

C. *Motor*

RC servo motor in the control of the robot, usually used to do robot gripper, but RC motor and DC motor is different, is that RC motor is usually used to do high torque low speed use, DC motor is used to do low torque high speed use, RC servo motor is divided into angle type and continuous type.

The angle servo motor is an angle control of 180°, and the circuit configured by the motor is simpler than the DC servo motor with three lines that can be controlled, GND-black, 5V-red, SIG-white (yellow), and the function of its foot position, which is to change the PWM pulse width to do angle control.

The difference between continuous RC servo motor and angle RC servo motor is that although they are both control PWM signal adjustment control, one is to use PWM pulse width to adjust the angle, is to control the direction of rotation and speed.

D. *sensors*

Infrared distance sensor is a distance measurement sensor assembly, composed of a PSD (position sensitive sensor), IRED (infrared light-emitting diode) and signal processing circuit, using triangulation method, the detection distance is not easily affected by the reflectivity of various objects, the temperature of the environment and the continuous operation time, the output voltage corresponds to the detection distance, so this sensor can also be used as a proximity sensor [5].

The ultrasonic wave is connected to the circuit board, confirm whether the pin position is correct, if there is no problem, the computer connects with myRIO and uses the project to find myRIO hardware, connect and turn on new VI, write the ultrasonic program in the program interface, after the writing is completed, you can test the ultrasound, and you can view the value on the human-machine interface to observe whether there is an error with the actual distance [6].

III. Experiments

Use Vision Assistant (Image Assistant) as shown in Fig. 2 to edit the desired items. RGB uses the three primary colors to find the respective ranges of the corresponding colors. Use Vision Assistant to edit the desired items. Select Color Threshold, RGB uses the three primary colors to find the respective ranges of the corresponding colors. Press Select Control again to display the input and output adjustment page, select the desired input and output, and press Finish, which is complete. Open the new VI and write the shape recognition program in the terminal (as shown in Fig. 3. At first, the target is taken by Vision Acquisition, then edited by Vision Assistant, the shape name is taken out by Index Array, and finally the corresponding shape photo is found on the computer side by using the name.

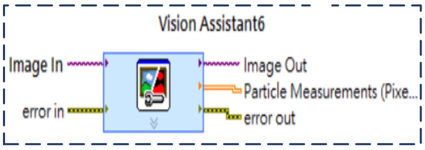

Vision Assistant6
Image In
error in
Image Out
Particle Measurements (Pixe...
error out

Fig. 2 Vision Assistant

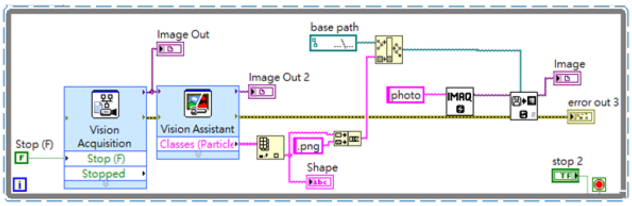


Image Out
base path
Image Out 2
photo
IMAQ
Image
error out 3
Stop (F)
Vision Acquisition
Stop (F)
Stopped
Vision Assistant
Classes (Particle
.png
Shape
stop 2

Fig. 3 Pattern Recognition Program

Similar Production Line Steps

The experimental structure of the production line is shown in Fig. 4.

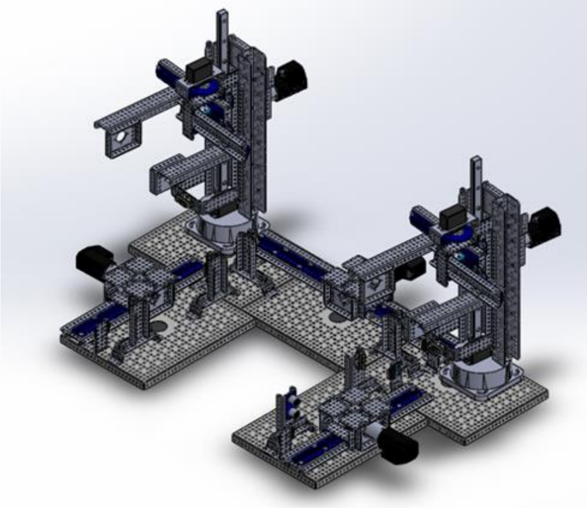
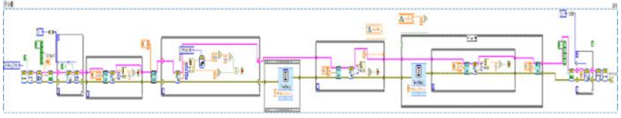


Fig. 4 Experimental structure of the production line

Similar Production Line Steps are shown as follows:

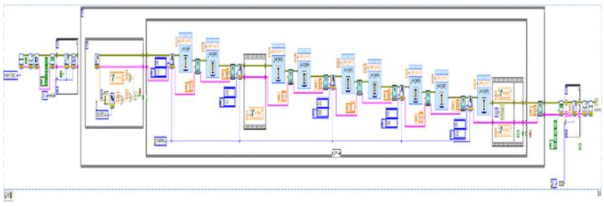
Step1

1.1 The object is put into the load platform
1.2 Infrared sensing object
1.3 Slide rail movement
1.4 Micro (limit) switch
1.5 Slide rail stop



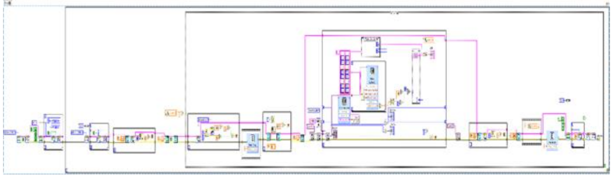
Step 2

2.1 Infrared sensing object
2.2 Arm descent
2.3 Grip object
2.4 Arm rise
2.5 Slide rail movement (return)
2.6 Turntable rotation
2.7 Arm descent
2.8 Object placement load platform
2.9 Arm return



Step 3

3.1 Object on the carrier platform
3.2 Infrared sensing object
3.3 Slide rail movement
3.4 Move under the lens (how to judge the stop?)
3.5 Lens recognition barcode and color
3.6 Micro (limit) switch
3.7 Slide stop



Step 4

4.1 Infrared sensing object
4.2 Arm descent
4.3 Gripping object
4.4 Arm rise
4.5 Turntable rotation (F left turn, T right turn)
4.6 Arm drop
4.7 Object placement load platform
4.8 Arm return

Step 5

5.1 Infrared sensing object
5.2 Slide rail movement
5.3 Micro (limit) switch

5.4 Slide rail stop
5.5 Ultrasonic sensing object (no object)
5.6 Slide rail movement return

IV. Results and Discussion

We use a myRIO controller as the main core to plan the factory automation production line structure through the practical function research for building similar production line implementation organizations. In the experiments, combining theoretical methodology and practical experiments to construct a structure design, electromechanical system and intelligent automation procedure based on LabVIEW2018. In the experiments, all 15 experiments can complete shape and color recognitions.

The efficiency of the designed automated production lines has verified the availability of designed structure. Then, the designed automated production lines can improve the capability of similar production lines.

V. Conclusion

This article mainly uses the myRIO controller developed the use of trapezoidal acceleration and deceleration control. We illustrated the results and used NI LabVIEW Vision Assistant to analyze image recognition. This program completes 15 units of teaching materials, which can provide basic teaching materials for intelligent automation applications and cultivate technical talents in cross-field intelligent automated production. The completion of similar production line materials in this program can enable promote the intelligent automation education. The experimental teaching materials continue to be developed and promoted, and it is hoped that in the future, the application function of image recognition can be extended to automated production technology to cultivate cross-field technical talents.

References

[1] Duncan McFarlane, Ajith Parlikad, Andy Neely, and Alan Thorne, "A Framework for Distributed Intelligent Automation Systems Developments," Proceedings of the 14th IFAC Symposium on Information Control Problems in Manufacturing, vol. 45, no. 6, pp. 758-763, 2012.

[2] Diana Kwon, "Self-Taught Robots," Science American, vol. 318, no. 3, pp. 26-31, March 2018.

[3] Janna Anderson, Lee Rainie, and Alex Luchsinger, "Artificial Intelligence and the Future of Humans," Artificial Intelligence, December 2018.

[4] Ren C. Luo, Yu-Ting Hsu, Yu-Cheng Wen, and Huan-Jun Ye, "Visual Image Caption Generation for Service Robotics and Industrial Applications," 2019 IEEE International Conference on Industrial Cyber Physical Systems (ICPS), 2019.

[5] En-Chen Chen, Cheng-Yang Shih, Ming-Zhi Dai, Han-Cheng Yeh, Yu-Chiang Chao, Hsin-Fei Meng, Wei-Ren Liu, Yi-Chen Chiu, Yao-Tsung Yeh, Chien-Jen Sun, Sheng-Fu Horng, and Sheng-Fu Horng, "Polymer Infrared Proximity Sensor Array," IEEE Transactions on Electron Device, vol. 58, no. 4, pp. 1215-1220, 2011.

[6] Fabio Ardiani, Mourad Benoussaad, and Alexandre Janot, "Improving Recursive Dynamic Parameter Estimation of Manipulators by knowing Robot's Model integrated in the Controller," IFAC Papers On Line, vol. 55, no.20. pp. 223–228, 2022.

Authors Introduction

Prof. Chia-Nan Ko

He is currently a Professor in Department of Automation Engineering, Nan Kai University of Technology, Tsaotun, Nantou, Taiwan. His research interests include neural networks, PID control, fuzzy logic control, support vector regression, and wavelet theory.

Yi-Yu Li

He is currently a third-year college student in Department of Automation Engineering, Nan Kai University of Technology, Tsaotun, Nantou, Taiwan. His research interests include automation engineering, design of automation, and mechanical drafting.

Ting-Ru Ko

He is currently a second-year college student in Department of Automation Engineering, Nan Kai University of Technology, Tsaotun, Nantou, Taiwan. His research interests include automation engineering and mechanical drafting.

ing-Yi Chen

He is currently a fifth-year grade student in Department of Automation Engineering, Nan Kai University of Technology, Tsaotun, Nantou, Taiwan. His research interests include automation engineering and mechanical drafting.

Using Multithreaded Load Balancer to Improve Connection Performance in Container Environment

Pang-Wei Tsai
Department of Information Management, National Central University, No. 300, Zhongda Rd., Zhongli Dist.
Taoyuan City 320, Taiwan.

Hong-Yu Wei
Department of Information Management, National Central University, No. 300, Zhongda Rd., Zhongli Dist.
Taoyuan City 320, Taiwan.

Yu-Chi Hsu
Department of Information Management, National Central University, No. 300, Zhongda Rd., Zhongli Dist.
Taoyuan City 320, Taiwan.

E-mail: pwtsai@ncu.edu.tw, 109423069@cc.ncu.edu.tw, yuchi.hsu@g.ncu.edu.tw
www.ncu.edu.tw

Abstract

The virtualization technology has been widely used in cloud service for a long period of time, it provides resource allocation, flexibility and efficiency by using virtual machine. Nevertheless, for lightweight workload, using container gains more advantages on transferability, rapid deployment, and robust ecosystem supports. Because of such characteristics, container is commonly used in edge computing and micro-service to fulfill essential resource requirements. However, when service deployed on the container receives too much requests without limitation, it may meet degrading effect which lowers down its performance and makes user experience to be poor. Hence, this research aims to investigate container request-response issue and implement multithreaded load balancer to improve HTTP connection performance equally for web-based application service.

Keywords: Virtualization, Container, Multithreaded, Load Balancer.

1. Introduction

With the growth of internet, requirements of networking service and application such as on-demand, quality assurance, and user experience are getting more and more important [1]. To provide stable, smooth, and persist data exchanges, how to design and develop a scalable and efficient system framework is a challenge [2]. For infrastructure provider, finding suitable solutions to support upper-layer service and application gains more benefits in production operation. Furthermore, the budget is an important thing as well. Not only hardware but software take cost in both deployment and maintenance. Therefore, it is difficult to find a fixed strategy to fulfill requirements of every applications and services for approaching cost-efficiency and high-performance.

Generally, user may access the application service from time to time. Depending on user amount, the resource consumption during rush and off-peak hours could be very different, especially for web-based services [3]. Hence, web service providers might be in doubt about paying more money to prepare extra resources to support uncertain flooding connections in a short time. Fortunately, the virtualization technology provides more flexibility and adaptation in infrastructure deployment. Considering the used computing component for running service software could be a physical server or virtual machine, enabling a control mechanism to allocate adequate resource to fulfill the requirement is necessary. Such resource on-demand concepts [4] has been widely used in cloud-based operation models. By doing this, service providers are able to proceed their preferred business models with performance-priority, essential low-cost, resilience or other strategies.

For most cloud resources, bare-metal, virtual dedicated, and virtual private servers are common units in selection. To allocate computing and networking resources with cost-efficiently, many researches [5][6] focus on auto-scaling and migration mechanisms to gracefully manage both spending and performance in balance. However, for lightweight micro-services, the overcapacity resource of virtual machine is too much. Hence, container solutions like Docker [7] has become a solution in supporting such operation scenes during deployment.

To improve the performance of lightweight web-based services running on container environment, this paper presents an investigation and uses multithreaded design to dispatch incoming connection to the pod in Kubernetes [8] environment. It aims to make adequate allocation for satisfying user HTTP requests to allocated pods.

2. Backgrounds and Related Work

2.1. *Container*

Even though the virtual machine solves the problem of isolation and security in sharing host resource [9], to prepare a production environment, operator still requires to treat virtual machine as a physical one (like installing OS). The hardware-level of virtual machine may gain more benefits from virtualization, while the OS-level is not exactly [10]. The resource of virtual machine might be too much for running some micro-services and their corresponding applications. Due to above reasons, the container starts to substitute traditional virtual machine for handling such lightweight tasks in recent years [11]. While there are still several issues about using container, such as software dependency and limited capability.

2.2. *Dockerized Environment*

For enabling container environment, using Docker framework [7] is a common choice. There is no longer need for cloud service provider to setup host environment for running application services separately. The container image could be prepared initially to support scale-up and flexible deployment in production use. With Kubernetes, the control mechanism is able to supervise created containers transparently [8]. Since created containers with the same purpose can be grouped as a pod, operators are able to monitor workload and network for making adaptation adjustment to satisfy user requirements well.

2.3. *Discussion*

According to the study made by S. K. Mohanty et al. [12], in different open-source container system architecture, the response latency of during the bursty has meet the bottleneck due to the in-efficient dispatching. Meanwhile, J. Li et al. [13] also compares several (e.g., Knative, Kubeless, OpenFaaS) frameworks and traces their API communication interactions. Both two researches indicate how to dispatch ingress flows to available pod is the key-point to make load balance in connection route. Hence, this research aims to investigate above issues, trying finding an equal strategy to gain more efficiency for user HTTP request-response.

3. System Design

To be capable of dispatching incoming request to assigned pod in Kubernetes, authors of this paper have modified the packet handling process of Kubernetes environment. The designed system architecture is shown as Fig. 1, and the added components are listed below:

- **Flow Controller:** It is the supervisor used to manage incoming request and determine forwarding action. The rule will be updated to packet filter and router in Kubernetes. It also keeps monitoring loading stats of each pod to approach balancing control.
- **Ingress Setup Module:** This module aims to receive incoming request and clarify the target path.
- **Pod Allocation Module:** This module maintains the available pod list for making forwarding decision.

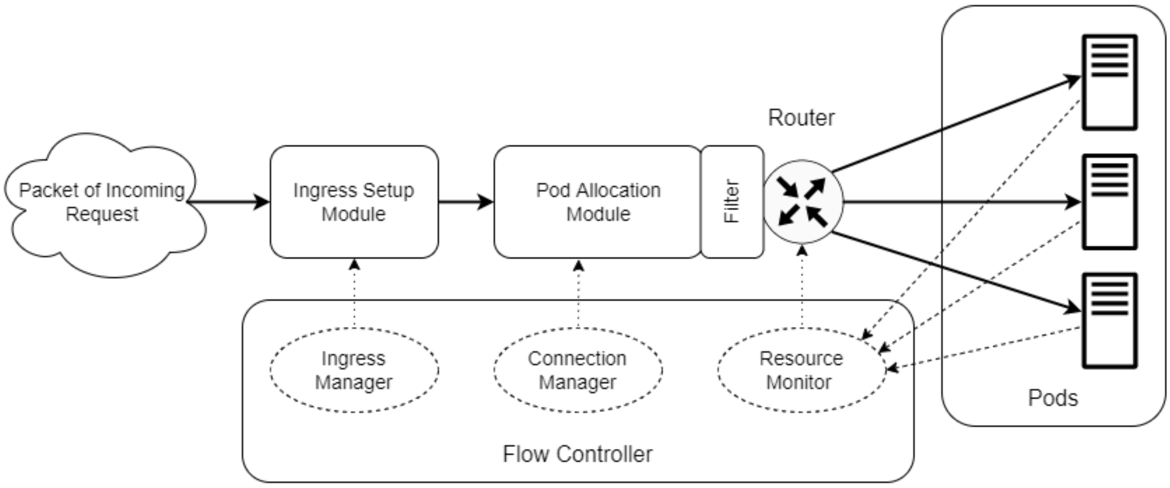


Fig. 1. The modified connection route architecture in Kubernetes.

- **Pod(s):** The component responds user requests. Each pod can be consisted of one or more containers.

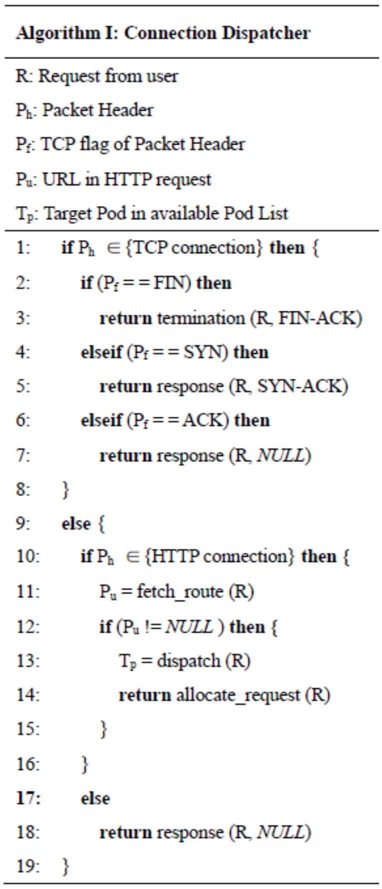
For every incoming request, the packet header will be used to make determination. The new request is going to be investigated by flow controller. After completing ingress setup and finding available pod, the request will be guided to the assigned pod through the established socket. All actions are supervised by the flow controller module, and its resource monitor keeps querying pod stats to understand current loading.

3.1. *Multithreaded Balancer*

During the operation, the balancing thread is expected to handover the request to flow control process, and the request will be guided by new forked load balancing thread. By doing this, it saves the waiting time for completing overall connection. Once the resource monitor finds any pod is getting overload or becoming malfunctioned, the flow controller should alter the routing strategy or ask physical host to create more pods for sharing the request.

3.2. *Request Allocation Algorithm*

The developed algorithm for allocation is shown below:



Algorithm I: Connection Dispatcher

```
R: Request from user
P_h: Packet Header
P_f: TCP flag of Packet Header
P_u: URL in HTTP request
T_p: Target Pod in available Pod List

if $P_h \in$ {TCP connection} then {
    if ($P_f$ = = FIN) then
        return termination (R, FIN-ACK)
    elseif ($P_f$ = = SYN) then
        return response (R, SYN-ACK)
    elseif ($P_f$ = = ACK) then
        return response (R, NULL)
}
else {
    if $P_h \in$ {HTTP connection} then {
        $P_u$ = fetch_route (R)
        if ($P_u$ != NULL ) then {
            $T_p$ = dispatch (R)
            return allocate_request (R)
        }
    }
    else
        return response (R, NULL)
}
```

3.3. *Balancing Strategy*

For making simple proof-of-concept, the applied strategy in development is using equal-cost policy to determine which pod should serve the new incoming request. According to monitoring data, flow controller will take consideration with ingress path, IP address, request URL, available pods (including workload and memory utilization) to decide which pod should be selected.

Hence, any pod with lowest utilization has priority chance to be allocated. In addition, although the default policy is trying to approach equality, it is still possible to change to be the weighted one for fulfilling purposes in different operation scenarios.

4. Experiment and Evaluation

To make evaluation, the experiment environment was conducted by one client and one hosting PCs (with Intel I7-10700 CPU, 64GB DDR4 RAM, WD NVME SSD, and Intel 82576 Network Adapter). The developed modules aim to enable load balancing mechanisms to dispatch the connection sent from client PC (emulated by JMeter [14] - a stress test toolkit) to the assigned pod.

4.1. *Experiment Design*

The used Docker template in experiment involves Ubuntu 18.04 and pre-configured Apache2. There are 5 customized HTTP documents placed in folders. The designed load balancing mechanism will guide each request to a small scale resource pool with 25 pods equally. To determine whether the request-response action made by JMeter is success, based on a technical report announced by Google [15], authors decided to set 3 seconds as the maximum waiting period of threshold.

4.2. *Evaluation Results*

In experiment, the number of JMeter request was started at 100 (and added 100 in batch) randomly. According to the experiment result, the system is capable of satisfying about 600-700 requests in 1 minute with 25 pods. Over this value, the waiting time will be increased rapidly and determined as failed request-response interaction (see Fig. 2 and Fig. 3). To understand what will be the changes when increasing more pods, authors also created additional pod during 700 requests. While it seems like the bottleneck is not load balancing module because of the number of failed interaction was lowering down when changing from 25 to 50 pods in experiment.

5. Conclusions

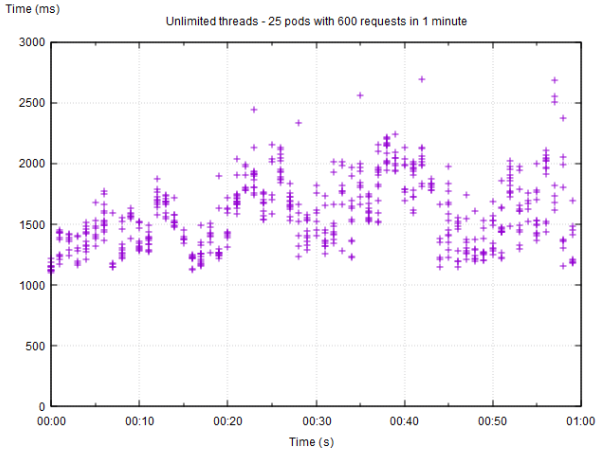


Fig. 2. The JMeter stress test results of random 600 requests in 1 minutes with 25 pods.

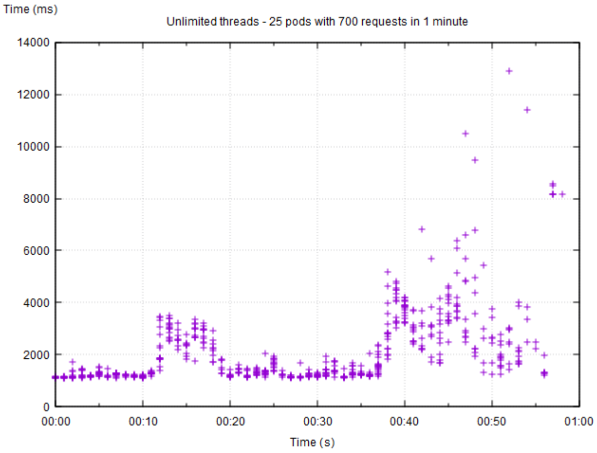


Fig. 3. The JMeter stress test results of random 700 requests in 1 minutes with 25 pods.

This research proposed a multithreaded load balancing development in container environment. By using JMeter to evaluate the success rate of request-response interaction, the experiment results show that the multithreaded balancer is able to dispatch hundreds of connections equally with pre-defined mechanism.

Acknowledgments

This research was supported in part by NSC of Taiwan (109-2222-E-008-005-MY3 and 111-2218-E-006-010-MBK). Authors are also grateful to TWAREN SDN research teams for their great help.

References

1. T. Zhao et al., “QoE in video transmission: A user experience-driven strategy,” IEEE Communications Surveys & Tutorials, vol. 19, no. 1, pp. 285-302, 2016.

2. S. Lehrig et al., “Scalability, elasticity, and efficiency in cloud computing: A systematic literature review of definitions and metrics,” in proceedings of the international ACM SIGSOFT conference on quality of software architectures, pp. 83-92, 2015.
3. A. L. Lemos et al., “Web service composition: a survey of techniques and tools,” ACM Computing Surveys vol.48, no.3, pp. 1-41, 2015.
4. Y. Tokusashi et al., “The case for in-network computing on demand,” in proceedings of the EuroSys conference, pp. 1-16, 2019.
5. E. Roloff et al., “High performance computing in the cloud: Deployment, performance and cost efficiency,” in proceedings of the IEEE International Conference on Cloud Computing Technology and Science Proceedings, pp. 371-378, 2012.
6. K.R.R. Babu and P. Samuel, “Interference aware prediction mechanism for auto scaling in cloud,” Computers & Electrical Engineering, vol. 69, pp. 351-363, 2018.
7. D. Bernstein, “Containers and cloud: From lxc to docker to kubernetes,” IEEE Cloud Computing, vol. 1, no. 3, pp. 81-84, 2014.
8. B. Burns et al., “Kubernetes: up and running,” O'Reilly Media, Inc., 2022.
9. A. V. Cleeff et al., “Security implications of virtualization: A literature study,” in proceedings of the international conference on computational science and engineering, pp. 353-358, 2009.
10. J. Sahoo et al., “Virtualization: A survey on concepts, taxonomy and associated security issues,” in proceedings of the international conference on computer and network technology. pp. 222-226, 2010.
11. I. Baldini et al., “Serverless computing: Current trends and open problems,” Research Advances in Cloud Computing, Springer, pp. 1-20, 2017.
12. S. K. Mohanty et al., “An evaluation of open source serverless computing frameworks,” in proceedings of the IEEE International Conference on Cloud Computing Technology and Science, pp.115-120, 2018.
13. J. Li et al., “Analyzing open-source serverless platforms: characteristics and performance,” arXiv:2106.03601, pp. 1-7, 2021.
14. B. Erinle., “Performance testing with JMeter,” Packt Publishing Ltd, 2015.
15. D. An, “Find out how you stack up to new industry benchmarks for mobile page speed,” [Online]. Available: https://www.thinkwithgoogle.com/marketing-strategies/app-and-mobile/mobile-page-speed-new-industry-benchmarks/. [Dec. 10 2022].

Authors Introduction

Dr. Pang-Wei Tsai

He received the B.S. degree in Electrical Engineering and the M.S. / Ph.D. degrees in Computer and Communication Engineering from National Cheng Kung University. His research interests include SDN, cloud computing, information security, and network testbed.

Mr. Hong-Yu Wei

He received the M.S. degree in Department of Information Management from National Central University, 2022. His research interests include Internet technology, cloud application, and Kubernetes.

Mr. Yu-Chi Hsu

He received the B.S. degree in Information Management in 2019 form the Fu Jen Catholic University. He is currently studying in National Central University for M.S. degree. His research interests are Internet technology, network management, and information security.

Electronic Biometric Detector and body composition index in predicting disease risk

Wen-Fu Yang
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

Chung-Te Ting
Department of Tourism, Food and Beverage Management,
Chang Jung Christian University,
No. 1, Changda Rd., Gueiren District, Tainan City, 711301, Taiwan
E-mail: leonyang899@gmail.com, ctting@mail.cjcu.edu.tw

Abstract

According to the World Health Organization, COVID-19 has killed 14.9 million people worldwide by 2021. According to statistics from the World Obesity Alliance, if the domestic obese population exceeds 50% (three high diseases), the country's COVID-19 mortality rate will be 10 times higher. There seems to be a close relationship between obesity and the risk of hospitalization and treatment. In this study, the electronic biological information detector was used to collect the response values of the body cells of adults in Taiwan. Through the comparison and analysis of big data, the functional status of each organ system in the human body was calculated and compared with the body composition index. It was found that there is a high correlation between the two. The results of this study may provide feasibility of different health risk assessments.

Keywords: COVID-19, electronic biological information detector, big data, obesity, body composition index

1. Introduction

The aging of the global population has made people's lifestyles and living environment more complex and difficult. Sub-health issues derived from dietary content and habits are becoming more and more serious, leading to an increasing number of people with chronic Triple H (hypertension, hyperlipidemia, and hyperglycemia), rising major diseases and mortality rates. According to the 2017-2020 Nutrition and Health Survey in Taiwan, NAHSIT [1] conducted by the Health Promotion Administration, MOHW, the prevalence of obesity among adults over 19 years of age in Taiwan averaged 80.65% for men and 91.77% for women (Table 1). Researchers at the Chicago Medicine found that patients hospitalized for COVID-19 were more likely to die if they also had obesity, even when age, gender, and underlying disease were taken into account [2].

There appears to be a strong relationship between obesity and the risk of hospitalization and treatment. According to the information on the national health financial situation of the National Health Insurance Administration, MOHW, the medical expenses between 2011 and 2021 enhanced from 463.3 billion to 671.1 billion, with a significant increase of 45% in 10 years [3]. Therefore, it is a matter of great urgency to make it easier for people to understand their health risks and for health care professionals to help them make adjustments to stay away from the risk of obesity and reduce the country's financial burden on health care.

Table 1. Prevalence of obesity in adults (standardized by percentage of body fat)

Gender	Age (years old)	Number of samples	Excess fat percentage (%)
Male	19-30	250	65.06
	31-44	277	86.32
	45-64	621	82.95
	65 -	1008	85.99
	≥19	2,156	80.65
Female	19-30	256	89.94
	31-44	294	86.39
	45-64	678	94.25
	65 –	920	96.07

≧19	2,148	91.77

Source: 2017-2020 Nutrition and Health Survey in Taiwan, NAHSIT.

2. Theoretical Background

As the population continues to age, lifestyles and environments become more complex, leading to an increasing number of people suffering from chronic diseases and rising cancer incidence and mortality rates [4]. Much of the research in modern medical technology is focused on disease assessment and patient care. However, even if some of the risk factors are within normal limits, individuals are not necessarily completely healthy [5]. In the modern medical model, complex tests are used to obtain detailed data for diagnosis, which can be associated with high costs and wastage of resources. Diagnosing cancer early doesn't really improve the chances of survival, and it just lets patients know early when they get cancer [6]. Therefore, if health policies and resources are used for the purpose of treating the disease, it does not contribute significantly to improving people's quality of life and reducing the burden of national health care costs. COVID-19, which has emerged in recent years, is a highly contagious disease with viruses. There are currently 48,539,872 infections in 215 countries and 1,232,791 deaths [7]. National health and education policies around the world are under severe scrutiny and clinicians are facing major challenges, leading everyone to rethink the importance of preventive medicine and health promotion. Historically, healthcare has been biased towards curative medicine at the expense of preventive care, resulting in high levels of preventable disease and mortality [8].

With the development of information technology, collecting physiological data and health behavior data for analysis, and then providing health adjustment suggestions and services, has become a fairly complete service model. It is a trend in the development of big health industry [9] to build and cross-reference biological information data by using big data, artificial intelligence health management and cloud monitoring so as to observe changes in test values for disease risk prediction.

3. Materials and Methods

In this study, we applied the field of empirical research, both qualitative and quantitative, and combined theoretical and practical applications. The research content is that after using bioenergy test technology to detect the response signals of the subjects' body organs, through database comparison and analysis, the risk of hypertension and hyperlipidemia in chronic Triple H diseases (hypertension, hyperlipidemia, and hyperglycemia) is analyzed, and the data were then cross-checked with the body composition indices of the subjects to confirm the correlation between the two data. In this study, we attempted to construct an innovative service model for disease risk prevention, and the following is a detailed description of the research methodology.

3.1 Subjects

A total of 10 academic and project management workers, ranging in age from 26 to 70 years old, six females and four males, were enrolled in this study. Before participating in the study, the participants were informed of the study content and the precautions to be taken and signed a consent form for the use of their personal data. The results of the study will be used for academic research purposes only.

3.2 Research tools

A. Bio-Energy Testing Analyzer

The bio-energy testing analyzer technology originated from the health care sector of the former Soviet space agency program, and its research purpose was to monitor the body condition of astronauts in weightlessness. Over 35 years, tens of billions of dollars of national research funds have been spent to collect clinical data on tens of millions of test cases of different genders, ages and ethnicities. It is a scientific instrument that uses AI test and big data statistical analysis and comparison.

The principle of the device is to use the bio-cell resonance method to collect data for diagnosis and adjustment. The instrument uses a signal transmitter to send a signal to the brain. The corresponding organ part will generate the corresponding signal and return to the transmitter. The signal feedback time varies in speed and strength. The computer collects the relevant data, calculates and compares the similarity (data and line pattern) with the database to find out the risk of related diseases.

The instrument can test the 12 major systems of the body without the need to fast in advance, dressing or undressing, with a short testing time. If necessary, it can be used in conjunction with medical equipment for review and verification for the best treatment plan of the patient to achieve maximum therapeutic effect and minimum side effects and to reduce unnecessary waste of resources.

B. Body Composition Index (BCI)

The body composition index analyzer is the most commonly used device for weight management by ordinary people. The principle of obtaining the body composition index is the application of bioelectrical impedance analysis (BIA). Impedance refers to the resistance to the flow of electricity (electrical resistance). Compared to muscle and blood, body fat and skin are less electrically conductive and have a higher impedance. Therefore, if the body weight is the same, the higher the body fat, the higher the impedance data will be. In this study, four items, namely body weight, visceral fat, body fat and basal metabolic rate, which are highly correlated with the risk of hypertension and hyperlipidemia, were selected for comparison.

3.3 *Research process*

Ten subjects were recruited and tested by the bio-energy testing analyzer and body composition index, and the obtained data were then cross-analyzed.

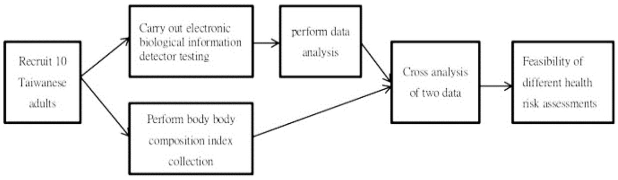

Recruit 10 Taiwanese adults
Carry out electronic biological information detector testing
perform data analysis
Perform body body composition index collection
Cross analysis of two data
Feasibility of different health risk assessments

Fig. 1. Research process

3.4 *Description of data interpretation criteria*

There are two types of test data in this study, each of which has criteria for determining the risk of hypertension and hyperlipidemia or for warning signs of impending risk, as described below.

A. Bio-energy testing analyzer

In the three sub-categories of "Disease Classification," "Growth Factors" and "Pathological Pattern," there are items related to hypertension and hyperlipidemia, such as "atherosclerosis" and "vasculitis" representing possible triggers of hypertension, and "gallbladder dysfunction" representing possible triggers of hyperlipidemia. Please refer to Table 2 for the determination criteria.

Table 2. Bio-energy testing determining criteria

Number of testing	< 0.425	0.425 ~ 1	>1
Representative meaning	High risk	Medium risk	Low risk

B. Body Composition Index

Body Composition Index (BMI) is one of the most common references in modern medicine for determining body condition. Overweight, high body fat, high visceral fat and reduced basal metabolic rate are all clinically associated with the induction of hypertension and hyperlipidemia. When the value changes gradually increase or decrease, it is usually a reference for the time before or after the disease onset. Therefore, if the test data is not within the standard value range, an early warning will be given to the case. Please refer to Table 3 for the determination criteria.

Table 3. Body Composition Index (BCI) value testing criteria

Item	Name	Description
1	Weight	Normal body weight: The test value is within plus or minus 10% of the standard value. Obesity or underweight: The test value is more than plus or minus 20%.
2	Visceral fat	Normal visceral fat values for males are 4 to 6 and for females 2 to 4. Values between 10 and 14 are considered obese; values above 15 are considered dangerous.
3	Body fat	The normal body fat rate for men is between 15% and 25%, and the normal body fat rate for women is between 20% and 30%; adult male body fat rate more than 25%, adult female body fat rate more than 30% is obese.
4	Basal metabolic rate	The standard value is (13.7×weight (kg)+5×height (cm)-6.8×age) for male and (9.6×weight (kg)+1.8×height (cm)-4.7×age)+655 for female. If the basal metabolic rate is lower than the standard, it means weight gain occurs.

4. Research Results

The above two test methods, based on their academic theories or design principles, warnings to the risk of patients suffering from chronic diseases of hypertension and hyperlipidemia are provided. The results of the bio-energy testing analyzer test and the body composition index test for risk warning are shown in Table 4.

Table 4. Comprehensive evaluation form

Item	Bio-energy testing analyzer warning		Body Composition Index warning	
Number	Hypertension-related	Hyperlipidemia related	Hypertension-related	Hyperlipidemia related
A	■	▲	■	▲
B	■	▲	■	▲
C	■			
D	■		■	▲
E	■	▲	■	▲
F	■		■	▲
G	■	▲	■	▲
H	■	▲	■	▲
I	■	▲	■	▲

J	🟥	▲	🟥	▲

Note: "🟥, ▲" appearing in the same item in different testing methods represent the same recommendation for warning.

Cross-comparison results: For the seven cases, including A, B, E, G, H, I and J, the results of the bio-energy testing analyzer and the Body Composition Index (BCI) test for hypertension and hyperlipidemia were in 100% agreement.

Analysis of Case C: The bio-energy testing analyzer issued an alert for hypertension, mainly detecting the signals of "vegetative nervous system (VNS) vascular" and "vascular dysfunction." This symptom was related to autonomic nerves and should be related to project management work stress after consultation. However, the body composition indices were within the standard values and this result should be discussed.

Analysis of Case D: The body composition index was 4.3 kg, 3% and 2% above the standard values for body weight, body fat and visceral fat respectively, representing a slight risk. However, the basal metabolic rate was higher than the standard value of 42, which means that the body had good metabolic capacity. Therefore, the risk should be a short-term state.

Analysis of Case F: Body Composition Index (BMI) values for body weight, body fat, visceral fat and basal metabolic rate were 17.1 kg, 14.9% and 5% above the standard values respectively, representing a high risk of disease. Afterwards, the case was asked to provide a blood report, and the values of triglyceride (TG), cholesterol (TC) and low-density lipoprotein cholesterol (LDL-C) are all within the standard values (Table 5). This result should be discussed as there was no risk of hyperlipidemia in the case.

Table 5. Blood test items

Item	TG	TC	HDL-C	LDL-C
Test value	118	171	44.7	116.7
Test standard	30-150	130-200	35-85	0-130

5. Conclusions and Suggestions

The bio-energy testing analyzer measured the warning results for hypertension and hyperlipidemia, and the body composition index. The cross-comparison results of hypertension and hyperlipidemia were 90% and 80%, respectively.

The purpose of this study is to investigate the comparative analysis of electronic biomarkers and body composition index in predicting disease risk, and to verify the consistency through the data of the test results. It is expected to be applied to preventive medicine health management. The bio-energy testing analyzer uses test technology and big data to collect and compare data to provide detailed functional information of various systems of the body. If the bio-energy testing analyzer results can be used as the entry point for the first stage of a health management plan, and the warning information can be used to provide recommendations for follow-up care, it is the most complete service framework for preventive medicine.

The research scope of this project is only set to verify the hypertension and hyperlipidemia items of the chronic Triple H diseases, as the bio-energy testing analyzer provides very diverse and detailed information. It is recommended to increase the number of participating cases, test and analyze different occupational categories and age groups in the future research.

Conflicts of Interest

The author declares no conflicts of interest.

References

1. https://www.hpa.gov.tw/Pages/Detail.aspx?nodeid=3999&pid=11145 (In Chinese)
2. https://www.worldobesity.org/resources/policy-dossiers/obesity-covid-19/data-and-case-studies
3. https://www.nhi.gov.tw/Content_List.aspx?n=D529CAC4D8F8E77B&topn=23C660CAACAA159D (In Chinese)
4. Li Ai, Wireless Communications and Mobile Computing, 4727340, 2021.
5. Wang, C.-C.; Chang, C.-D.; Jiang, B.C. Developing a Health Risk Evaluation Method for Triple H. Int. J. Environ. Res. Public Health, Vol. 16, No. 7, 1168, 2019.
6. Tseng, P. H., Lin, H. J., Chiu, H. M., Lee, B. C., Wu, M. S., & Chen, M. F. Self-paid health examination: An evidence-based perspective. J. Intern. Med. Taiwan, Vol. 20, pp. 532-543, 2009.
7. Khan M, Adil SF, Alkhathlan HZ, Tahir MN, Saif S, Khan M, Khan ST. COVID-19: A Global Challenge with Old History, Epidemiology and Progress So Far. Molecules, Vol. 26, No. 1, 39, 2021.
8. Abdulhammed Opeyemi Babatunde, Habib Ayomide Shobanke, Abdulsabur Aderemi Akinade, Anayo James Michael, Modupe Osadare, Oluwatomisin Khadijat Akanbi, Obiorah Chioma Favour, Yusuff Adebayo Adebisi, Enhancing preventive medicine over curative medicine: Role of telemedicine, Public Health in Practice, Vol. 2, 100130, 2021.
9. Hu, X., Abdulghani, A. M., Imran, M., & Abbasi, Q. H. Internet of Things (IoT) for Healthcare Application: Wearable Sleep Body Position Monitoring System Using IoT Platform. In Proceedings of the International

Conference on Computing, Networks and Internet of Things, CNIOT 2020, No. 6, pp. 76-81,2020.

Authors Introduction

Dr. Wen-Fu Yang

He is a senior professional health risk analyst and psychological consultant in the preventive medicine industry, with about 20 years of industry experience. The main research directions are preventive medicine management, big data analysis and psychological counseling.

Dr. Chung-Te Ting

He is a doctor of Applied Economics Department of National Chung Hsing University. He is currently an associate professor at the Department of Tourism and Catering Services, Chang Jung University, Taiwan. His research interests focus on tourism industry performance measurement and tourism consumption behavior analysis.

Exploring Consumers' Intention to Use Mobile Payment APPs Based on Technology Acceptance Models - Taking Line Pay as an Example

Chun-Nan 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
E-mail: blakebob66@mail.cjcu.edu.tw

Yu-Sheng Huang
Department of Tourism, Food and Beverage Management,
Chang Jung Christian University,
No. 1, Changda Rd., Gueiren District, Tainan City, 711301, Taiwan

Abstract

According to the statistics of the Financial Supervisory Commission, as of the end of March 2021, the cumulative transaction amount of the latest five mobile payments has reached 503.2 billion， LINE, a communication software that is inseparable from our lives, has launched a service - "LINE Pay", which has become the most commonly used mobile payment by Taiwanese consumers. This research mainly takes people who have installed LINE PAY as the research object, and uses three factors that may affect consumers' consumption patterns, such as " Brand Association ", " Consciousness Risk " and "social impact".) to explore consumers’ usage intentions for mobile payment apps, hoping to gain a deeper understanding of their usage intentions.

Keywords: action payments, technology acceptance models, brand association, consciousness risk, social influence

1. Introduction

Since the launch of the three major international payment platforms in 2017, the advance in mobile technology and the development of the Internet have led to considerable innovation in consumer services. The impact of mobile payments on consumers is significant, in terms of not only increasing the productivity of the financial system or related industries but also enhancing convenience for both merchants and consumers.

It is found in the consumer survey concerning mobile payment conducted by the Institute for Information Industry (III) that consumer preference for mobile payments has increased significantly, from 37% to 50% in 2020 [1]. In the last two years, the impact of COVID-19 has changed consumers' payment habits and spending behavior, with people beginning to accept contactless payment. According to the Financial Supervisory Commission (FSC) [2], as of the end of March 2021, the cumulative value of the five largest mobile payment transactions has reached NT$503.2 billion. The survey conducted by 1111 Job Bank also shows that as many as 90% of the respondents have used mobile and electronic payment tools [3]. LINE Pay (65.5%) is the most frequently used, followed by EasyCard (54.0%), Apple Pay (25.6%), Jiekou Pay (22.9%) and Taiwan Pay (18.0%). LINE, a communication software inseparable from our lives, has launched its service LINE Pay, which is the most popular mobile payment service used by consumers in Taiwan.

In addition, among the various theories of innovative technology acceptance models, the most widely used model framework is the Technology Acceptance Model (TAM) proposed by Davis in 1989. In this study, we focus on people who have installed LINE PAY, and use the Technology Acceptance Model (TAM) to explore consumers' usage intent mobile payment apps through three factors that may influence consumers' consumption patterns, namely ‘brand association,’ ‘consciousness risk’ and ‘social impact.’ We hope to gain a better understanding of behavioral intention to use.

2. Literature Review

2.1 Technology Acceptance Model (TAM)

The Technology Acceptance Model (TAM) was first proposed by Davis in 1986. Based on Fishbein and Ajzen's Theory of Reasoned Action (TRA), he developed a model to explain and predict how users would behave in an information acceptance system. Davis (1989) [4] pointed out in his research that among the variables affecting system use, the two most important determinants were perceived usefulness and perceived ease of use. Therefore, the model focused on the perceived usefulness and perceived ease of use of IT to analyze users' behavioral intention to use IT. And in his research, it was also pointed out that perceived ease of use would affect users' perceived usefulness of new technologies, and the two were in a positive relationship. In other words, when users subjectively perceive that the use of a new technology will improve their work performance or results, then the operation or ease of use of the technology will affect the user's attitude and intention to use the technology. The technology acceptance model proposed by Davis is shown in Figure 1.

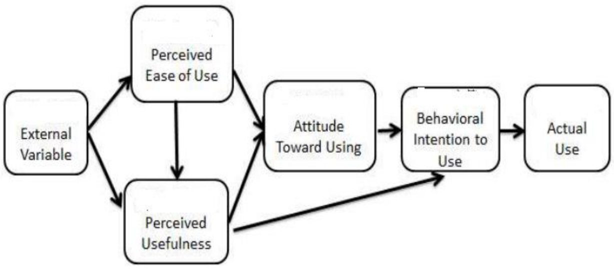


Figure 1.Technology Acceptance Model, TAM

2.2 Definition of Mobile Payments

In *Innovations in Retail Payments* published by the Bank for International Settlements in 2012 [5], ‘mobile payment’ is defined that ‘as long as the payment is initiated over a mobile network, whether by voice, SMS or near-field communication, it can be called a mobile payment.’

2.3 Brand Association

"Brand Association" means that when consumers know a brand of a party, they can trigger relevant evaluations and ideas about the brand from all his memory impressions. Aaker (1996) advocated that brand equity can be viewed from two aspects of product and market, and proposed eleven indicators to measure brand equity [6].

2.4 Consciousness Risk

‘Consciousness risk’ is the risk of unsatisfactory consumption when consumers recognize the unpredictable sense of uncertainty. Therefore, the phenomenon of consumer behavior can be explored through the concept of consciousness risk. According to the literature review on consciousness risk, it is found that consumers' consciousness risk is a multi-faceted structure (Jacoby, J., & Kaplan, 1972) [7].

2.5 Social Impact

Social impact is the change in an individual's opinion or behavior as a result of the opinion or behavior of others or groups. Deutsh & Gerard (1955) [8] defined social impact as an individual's acceptance of information provided by others as a factual reference. Rogers (1995) believed that whether an individual adopts a new technology, in addition to personal decision-making patterns and technological characteristics, social impact plays an important role [9].

3. Research Methods

3.1 Modified Delphi method

The basic framework of this study is the technology acceptance model proposed by Davis in 1989, and the external factors that may influence consumers' intention to use mobile payments were compiled by referring to the relevant literature on mobile payment usage intent. The three external factors are referred to ‘brand association’ in Aaker's (1996) measure of brand equity. The ‘social impact’ proposed by Deutsh & Gerard in 1995 and the ‘consciousness risk’ proposed by Stone & Gronhaug in 1993 are used to analyze the current domestic mobile payment usage status and future development trends, so as to understand the factors that hinder consumers from adopting mobile payments. Consumers' intention to use mobile payment will be affected by not only their attitude towards using but also consciousness risk and social impact, as shown in Figure 2.

3.2 Operational Definitions and Measurement Questions of the Study Variables

3.2.1 Questionnaire design

This study focuses on the factors that influence the use of mobile payment by Taiwanese consumers. The questionnaire questions were adapted from the scales proposed by Ram (1987) [10], Thong, J. Y., & Xu, X (2012) [11] and others.

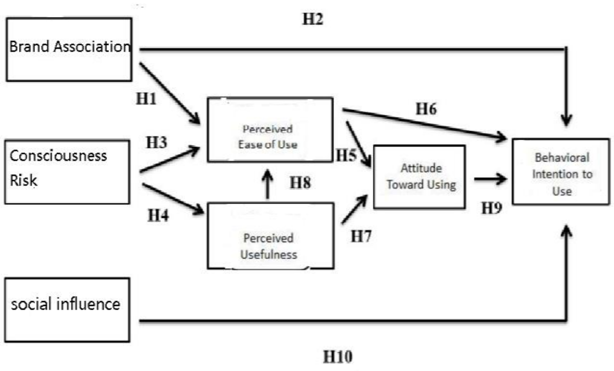


Figure 2 Research Architectur

3.2.2 *Questionnaire content*

This questionnaire is divided into two parts. In the first part, the questionnaires were developed in accordance with the research framework. The second part is the demographic variables. Each question was quantified using the five-point Likert scale, which was divided into ‘strongly agree,’ ‘agree,’ ‘undecided,’ ‘disagree,’ and ‘strongly disagree,’ and was assigned a score of 5,4,3,2,1 in order to facilitate subsequent statistical analysis.

4. Research Results

4.1 Descriptive Statistical Analysis

Regarding the socioeconomic background of the 323 respondents, the statistical results are shown in the table. Among the respondents, the majority of them were female, with 35.9% being male and 64.1% being female. In terms of age, the majority of respondents were 31-40 years old, accounting for 25% of the respondents, followed by 41-50 years old (23%) and 21-30 years old (21%). In terms of education level, the college and university level have the highest proportion (44%), followed by the (vocational) high school level (21.0%) and the junior college level (18%). The most common payment system used by respondents is "Line Pay,” with 96 respondents, accounting for 30% of the total valid sample, followed by "Apple Pay.” Respondents' experience in using mobile payment is mostly ‘1-2 years’ and ‘2-3 years,’ accounting for 36% and 26% of the overall valid samples respectively. as shown in table 1.

Table 1. Demographics of the respondents

Overall socioeconomic background information	**Category**	**Number of samples**	**Percentage**
Gender	Male	116	35.9
	Female	207	64.1
Age	Under 20	13	4
	21-30	69	21
	31-40	81	25
	41-50	76	23
	51-60	64	20
	61+	20	6
Education level	Below elementary school	6	1
	Junior high school	8	2
	Senior high school and vocational high school	68	21
	Junior college	59	18
	University	142	44
	Graduate school and above	40	12
Personal annual income	Below 200,000 (inclusive)	18	5
	200,000-400,000 (inclusive)	72	22
	400,000-600,000 (inclusive)	103	32
	600,000-800,000 (inclusive)	64	20
	800,000-1,000,000 (inclusive)	38	12
	More than 1 million	28	9
The most-often-used payment system	Apple pay	82	25
	Google pay	64	20
	Samsung pay	53	16
	Line pay	96	30
	Taiwan Pay	28	9
Mobile payment experience	Less than one year	44	14
	1-2 years	85	26
	2-3 years	115	36
	3-4 years	52	16
	4-5 years	27	8

4.2 Regression Analysis

Regression analysis was conducted to understand the relationship between the study variables and to test the hypotheses proposed in this study, as shown in table 2.

5. Conclusion

This research framework is based on technology acceptance patterns; three variables are used to examine the intent of consumers to use mobile payments through the narrative statistical analysis and regression analysis of dimensions. The recommendations are as follows.

(1) There is a significant positive correlation among 'perceived ease of use,' 'perceived usefulness' and 'attitude towards using.' When consumers feel that mobile payment is easy to use and can quickly solve their payment needs, their attitudes towards using mobile payment will also be affected.

(2) On the consciousness risk dimension, consumers may be exposed to potential risks when using mobile payment transactions, as examined by regression analysis. The hypothesis that consciousness risk has no effect on perceived usefulness and usage intent is not valid when examined by regression analysis. This phenomenon may be due to the fact that consumers are less sensitive to risk perception, and when the loss occurs, the existence of the risk will be realized. Consciousness risk for usage intention may be that the reward generated by mobile payment far exceeds the risk that one can bear.

(3) In terms of social impact, if the user is affected by important events or celebrity endorsements, it will affect their intention to use mobile payment. In a regression analysis, there is also a significant positive impact, representing important festivals or the influence of others will increase the intention to use mobile payment.

Hypothesis	Regression Analysis	R^2	F Value	T Value	Accepted or rejected
H1	Brand Association → Perceived Usefulness	.145	.146	.381	Rejected
H3	Consciousness Risk → Perceived Usefulness	.051	12.388**	3.521**	Accepted
H8	Perceived Ease of Use → Perceived Usefulness	.198	55.939**	7.478**	Accepted
H4	Consciousness Risk → Perceived Ease Of Use	.078	19.467**	4.413**	Accepted
H6	Perceived Usefulness → Usage Intent	.126	32.620**	5.712**	Accepted
H9	Attitude towards Using → Usage Intent	.225	65.652**	8.102**	Accepted
H2	Brand association → Usage Intent	.058	6.993**	-0.171	Rejected
H10	Social Impact → Usage Intent			3.744**	Accepted
H5	Perceived Usefulness → Attitude towards Using	.318	52.972	7.823**	Accepted
H7	Perceived Ease of Use → Attitude towards Using			2.523*	Accepted

(4) To increase consumers' behavioral intention to use mobile payments, how to introduce more practical points (or cash) rewards with low risk and how to allow consumers to use mobile payments with peace of mind and get rewards are the issues that the practitioners must consider.

Table 2. Regression Analysis

Note: * denotes P value < 0.05; ** denotes P value < 0.01 at significant level.

References

1. https://mic.iii.org.tw/aisp/news-content?sno=617 (In Chinese)
2. https://www.fsc.gov.tw/ch/home.jsp?id=487&parentpath=0,7,478 (In Chinese)
3. https://www.1111.com.tw/news/jobns/142660 (In Chinese)
4. Davis, F. D. (1989), Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, pp. 319-340.
5. https://www.bis.org/cpmi/publ/d102.htm
6. Aaker, D. A. (1996), Measuring Brand Equity Across Product and Markets, California Management Review, pp. 102-120.
7. Jacoby, J., & Kaplan, L. (1972). The components of perceived risk. In M. Venkatesan (Ed.) Annual conference of the association for consumer research: pp. 382-393. Chicago, IL: Association for Consumer Research.
8. Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. The Journal of Abnormal and Social Psychology, 51(3), pp. 629-636.
9. Rogers, E. M. (1995), Diffusion of Innovation-Fourth Edition. The Free Press. New York.
10. Ram, S and Sheth, J.N. (1989), Consumer Resistance to Innovations: The Marketing Problem and its Solutions, Journal of Consumer Marketing, 6(2), pp. 5-14.
11. Venkatesh, V., Thong, J. Y., & Xu, X. (2012), Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. MIS Quarterly, 36(1), pp. 157-178.

Authors Introduction

Mr. Chun-Nan Chen

He is a senior professional health risk analyst and psychological consultant in the preventive medicine industry, with about 20 years of industry experience. The main research directions are preventive medicine management, big data analysis and psychological counseling.

Dr. Yu-Sheng Huang

He is a doctor of Applied Economics Department of National Chung Hsing University. He is urrently an associate professor at the Department of Tourism and Catering Services, Chang Jung University, Taiwan. His main research topics in recent years mainly focus on: natural environmental resources, and local connections to society topics.

Optimization of robot path and IoT communication path based on artificial intelligence.

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

E-mail: jrhung@yuntech.edu.tw, khhsia@yuntech.edu.tw

Abstract

The Internet of Things (IoT) and robotics are very popular research topics, and they have begun to enter people's daily life, and both robots and the IoT have the problem of path optimization. For robots, Although the map can be established in advance and the robot can avoid obstacles, the robot's travel map is likely to change at any time, and a new path needs to be generated at this time. The IoT will provide users or monitoring system information with the best transmission speed path. The traditional optimal path is usually the shortest distance or shortest time. But because it ignores some variables, such as the probability of accidents. Or with the shortest time or distance, the results are planned in this way, but the best plan becomes the worst. Based on the above requirements, this paper uses artificial intelligence to optimize the path between robots and the IoT. question. The method expected in this paper is to parameterize the length of each path, the number of obstacles, whether there will be collisions, etc., and then use the artificial intelligence algorithm of multi-tree and Long short-term memory (LSTM) to find the best path.

Keywords: Path Optimization, Internet of Things (IOT), Robot, Recurrent Neural Network (RNN), Long short-term memory (LSTM).

1. Introduction

Path planning is one of the important topics in robotics and Internet of Things (IOT) applications. In the case of a robot, in order to allow the robot to complete the task quickly, it is hoped that the robot will complete the task with the shortest path. In the application of the Internet of Things, it is also hoped to provide services with the best transmission quality and the fastest transmission speed. But is the so-called shortest path really the best path? For example, on the robot path, there is a shortest path, so all robot path planning uses this path. Communication in the Internet of Things is also similar, all communications are concentrated in one communication node. The result of this is the "congestion" of robot or Internet of Things communication. At this time, the so-called shortest and best path is likely to become the "worst" path. Therefore, how to avoid such a result and make path planning achieve the best result is what this paper tries to discuss and solve. For example, Inkyung Sung [1] et al. proposed a neural network-like algorithm to use offline path data for real-time path planning. Its main purpose is to avoid collisions with obstacles in the environment when the robot is operating. Fatin Hassan Ajeil [2] and others proposed (Aging-Based Ant Colony Optimization, ABACO) algorithm for grid-based mobile robot path planning. In addition, in the research of Radmanesh [3] et al., it is proposed that although the A* path algorithm is very simple, if it is a very complex map, or if there are many variables in the entire map, it will consume a lot of computing resources. However, in the above research, after the path planning, the condition of the actual robot during operation has not been taken into consideration. For example, the originally planned path is no longer available, or all path planning calculates the same result, which will make the original optimal path planning result worse. In addition, in the robot part, although the current SLAM [4] (Simultaneous Localization and Mapping, SLAM) has been quite mature, but there are still several problems. For example, sensors such as lidar or image, if they are not detected, or because of sensor errors, the environment is constructed incorrectly, these will cause problems in the follow-up robot path planning, and there are similar problems in wireless communication. For example, it is necessary to optimize the communication quality or speed through other communication nodes, but it is impossible to know the status of the bandwidth and number of connections of the surrounding communication nodes. In this way, if only the signal strength is used to process the communication data, there may also be a problem that the best efficiency cannot be obtained. Therefore, this thesis tries to use artificial intelligence. On the issue of path planning, in addition to finding the results of the shortest path or the shortest time, it will also include possible states, such as whether all planning results are The same, or the probability of possible path conditions, as well as issues such as path quality and update rate are taken into account. Through this method, the level of the entire path planning can be made more complete, and when new variables appear, they can also be added to the planning conditions in real time. Through this method, it is hoped that the path planning can optimize the result and be able to Taking into account many possible situations and probabilities, instead of just pursuing the shortest path to obtain the shortest time, so that the entire path planning can actually be closer to the real situation in the communication between the robot and the Internet of Things, and truly achieve the goal of optimization.

2. System Architecture

In the traditional path planning method, the length of the calculated path and the length of time are usually used to plan, but there are usually several problems in this way, (1) there will be a problem of path concentration, that is, all the shortest paths are used in planning Or plan in the shortest time, but this will cause all planning results to be concentrated on one path, which will cause congestion. (2) It is impossible to consider the temporary state, such as a car accident or road damage, which will also cause a detour even though the optimal path is planned. (3) The path data has not been updated for a long time, so the original shortest path is no longer available. (4) Although it is the shortest path or the path with the shortest time, but because the road condition is unstable or the quality of the road condition is not good, it cannot be driven at the fastest speed, which will also deteriorate the result of the shortest path. Based on the above problems, we are thinking about using artificial intelligence to deal with these problems. On the basis of path planning, we still use A*[5] as the basis of path planning, because this algorithm is simple to calculate, although it can only be used to find the shortest path, but we use this basis, coupled with artificial intelligence processing, try to find a path that considers various variables, and then let the user or use specific options to find the most suitable path. In the part of artificial intelligence algorithm, we choose LSTM [7] based on recurrent neural network RNN [6]. The conditions and the establishment of the model can achieve different results such as one-to-one, one-to-many, many-to-one, and many-to-many, and such characteristics are just suitable for the application of path planning, that is, according to different needs and settings, find out a variety of options, and then find out the best feasible solution according to the actual situation and user needs, instead of only a single shortest path or shortest time. Figure 1 is the basic structure diagram of RNN, x represents the input data, h is the output, the result of yesterday's prediction can be used as the data of today's prediction, forming a long chain. But in practice, RNNs do not perform as expected in long-term memory. For example, in the known RNN model, if there are situations that appear in a long period of time, these long-term situations will be "forgotten" during the long-term learning process, that is, they will not appear in the end. Now in the final result, these "forgotten" factors will appear.

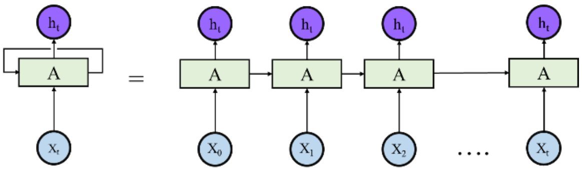


Fig. 1. The RNN basic structure diagram.

Therefore, LSTMs are designed to improve the short-term memory of RNN. Although the RNN constitutes a huge neural network, if we enlarge the internal units of the standard RNN, it is a rather simple architecture. Its architecture is shown in Figure 2, usually only one layer, including an equation called the activation function, there are several choices for this equation, which is represented here by the hyperbolic function tanh.

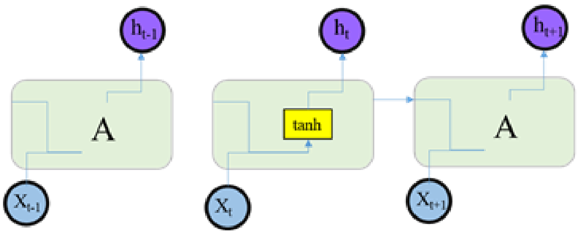


Fig. 2. The LSTMs basic structure diagram.

In LSTMs, it is more complicated and is mainly divided into several parts. Its architecture is shown in Figure 3. In this Figure 3, the first area is the core concept of LSTMs: adjustment through gates. The horizontal line from left to right in area 1 is used to represent the cell state. It can be imagined as a road that runs through all cells, that is, to bring information from one cell to the next. For example, yesterday used A condition, the "memory" of planning the path using condition B the day before yesterday.

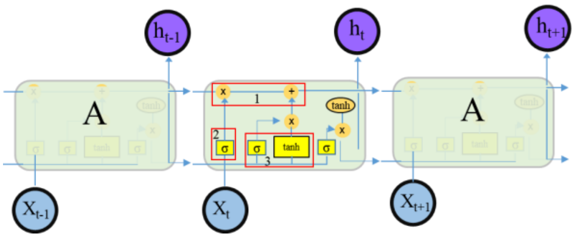


Fig. 3. The LSTMs structure diagram.

But this memory can be adjusted through the orange gate. In this paper, the control of this gate will be handled by using conditions such as accuracy rate and demand degree. The range is 1%~99%. It is mainly used as the ratio of different states or requirements, or the length of the path, The choice of the length of time, etc., and since we will not turn all the ratios into 0, the conditions used in the past will be retained, and can still be used as the result of planning at an appropriate time. The part in area 2 is used to determine how much the judgment results of the previous layer should be retained. Similarly, we also use the range of 1% to 99%. The area 3 is used to determine how much information is used for judgment. Since LSTMs are based on the RNN algorithm, the current-end neurons have accumulated a lot of calculated data, which will cause a lot of calculations for operations and decision-making. Time, so there must be a trade-off, and this area is for this purpose. In this article, we use the conditional structure of the multi-element tree. The biggest difference between this and the traditional binary tree is that the binary tree only has "0" and "1" ", "true" and "wrong" judgments, while the multiple tree will provide more options, and these options are also derived from the user or the past planning results, here we are for each condition The setting range of is 0%~100%.

For example, if the user only needs the shortest path and the shortest time, and other conditions are not considered, then the relevant conditions can be set to a maximum value or ignored. Through this mechanism, the entire path planning can be made more flexible. It can appropriately reduce the calculation time required due to too much data. Through such a mechanism, LSTMs will eliminate too long data for RNN, and can not provide more conditional screening mechanisms and other shortcomings. This is very helpful for us to use LSTMs algorithm for robot path planning and node path planning for Internet of Things communication. Next, we will explain how to use LSTMs to plan the communication path between robots and the Internet of Things.

2.1. *Experimental methods and results*

In the path planning part, this article uses the Titu database of Yunlin and Chiayi, Taiwan. The reason for choosing this database is that many places in this area are mountainous, so road conditions often have to be closed due to weather or engineering Or slow down, and we use this database to randomly select data in small areas, and then use LSTMs to plan and test the entire route. In the path planning part, this paper not only uses the basic data such as each path in the database and the time required for the driving path as planning parameters, but also adds the following parameters for path planning:

1. Data update rate: This parameter refers to how often the map database is updated. This is one of the key points that affect the planning results. The higher the update rate, the better the accuracy of the data, and these data can be used through actual paths. And update.
2. Accident occurrence rate: This refers to the possibility of accidents such as closure and blockage of the path, usually due to weather or unwarranted conditions, and this parameter is also adjusted through the accumulation and update of actual data.
3. Allocation rate: mainly to avoid that all planning results are the same, which will cause path blockage and reduce the efficiency of planning results.
4. Expectation indicators: mainly users can choose the shortest path, the shortest time, or choose to take a longer detour and spend a little more time, but it can avoid possible congestion caused by heavy traffic during commuting hours or when everyone uses the shortest path and other issues.

In addition, after the path planning in advance, the test status will be sent back to the database in real time during the actual test. If there is congestion, slow down, or other conditions, the relevant parameters will be updated in real time. Through such a mechanism, the status of the database can be kept up to date at any time. In the part of the experimental platform, it is divided into robots and IoT modules. In the part of the IoT module, we use the test platform designed by the STM32 microprocessor. This platform supports wireless communication functions such as WI-FI, Bluetooth, and LoRa, so it can be used Different wireless communication makes data transmission faster and more reliable. This module is shown in Figure 4. In the mobile robot part, we use the robot shown in Figure 5. This is a robot that can perform tasks outdoors. It also has the functions of laser positioning and 3D laser environment construction. It also uses STM32 microprocessors. The communication module of the machine, so the execution of the entire path planning, as well as the subsequent update of the map database, etc., can be updated accurately and in real time.

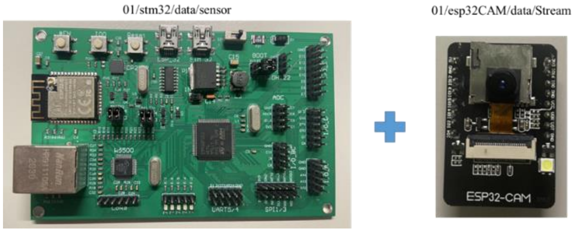


Fig. 4. Base on STM32 chip IoT Module Block Diagram.

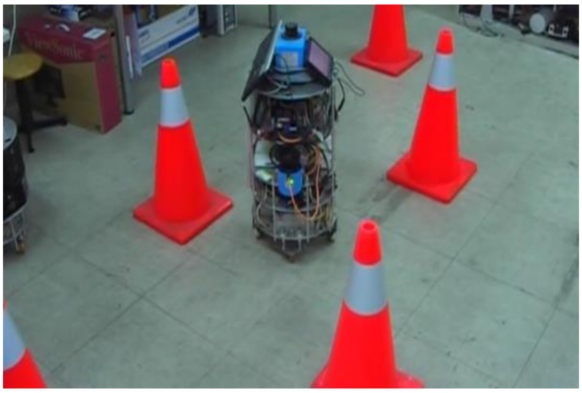


Fig. 5. Mobile robot.

This article uses the above-mentioned IoT modules and mobile robots to test the results of path planning using LSTMs. The overall operation is in good condition. During the test process, we also encountered path blocking, unexpected path interruption, and others mentioned at the beginning of this article. Different situations can be resolved and re-planned normally in this system.

3. Conclusions

This paper proposes an artificial intelligence path planning method using LSTMs. This method is based on the RNN algorithm, but through the adjustment of relevant data and parameters, the less frequently occurring data will not be excluded as when using RNN. Moreover, the use of data and parameters can make the overall planning result better through learning and parameter control. Through this method, we actually tested the IoT module of the STM32 microprocessor and the mobile robot, and the preliminary results were normal. However, due to the relationship between time and related equipment, it has not been tested with various platforms and map databases. We will continue to test in the future and hope to be able to develop it in the end. A path planning system that considers the real environment and can be used in practice.

References

1. SUNG, Inkyung; CHOI, Bongjun; NIELSEN, Peter. On the training of a neural network for online path planning with offline path planning algorithms. International Journal of Information Management, 2021, 57: 102142.
2. AJEIL, Fatin Hassan, et al. Grid-based mobile robot path planning using aging-based ant colony optimization algorithm in static and dynamic environments. Sensors, 2020, 20.7: 1880.
3. RADMANESH, Mohammadreza, et al. Overview of path-planning and obstacle avoidance algorithms for UAVs: A comparative study. Unmanned systems, 2018, 6.02: 95-118.
4. DURRANT-WHYTE, Hugh; BAILEY, Tim. Simultaneous localization and mapping: part I. IEEE robotics & automation magazine, 2006, 13.2: 99-110.
5. https://en.wikipedia.org/wiki/A*_search_algorithm#cite_note-:2-1,2023/01/10
6. MEDSKER, Larry R.; JAIN, L. C. Recurrent neural networks. Design and Applications, 2001, 5: 64-67.
7. HOCHREITER, Sepp; SCHMIDHUBER, Jürgen. Long short-term memory. Neural computation, 1997, 9.8: 1735-1780.

Authors Introduction

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.

Key Success Factors Affecting Family Members’ Intention to Withdraw from Life-sustaining Treatment for Long-term Ventilator-dependent patients: Nursing Professionals’ Perspective

Hsiao-Fang 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
Jin-Yuan Chern
Dept. of Health Care Administration, Chang Jung Christian University
No. 1, Changda Rd., Gueiren District, Tainan City, 711301, Taiwan
E-mail: chern@mail.cjcu.edu.tw

Abstract

This study aimed to explore, from the perspective of nursing professionals, the key success factors which would significantly affect VDPs family members’ decision about withdrawal from LST. A structured questionnaire composed of three constructs was developed and the analytic hierarchy process was adopted to collect and analyze the responses. Overall, “subjective norms” considered relatively important; family consensus and quality assurance were the two key factors when family members confronted the decision of withdrawing the LST for their beloved person.

Keywords: ventilator-dependent patient (VDP), analytic hierarchy process (AHP), life-sustaining treatment (LST), shared decision making (SDM), nursing professionals

1. Introduction

Mechanical ventilation after tracheal intubation is an important medical treatment and life-sustaining treatment (LST) for critical patients. While medical treatment may help delay death, for many vegetative people who have been dependent on respirators for a long time and are unconscious, it would cause long-term torture for patients and their families as well. Studies have found that about 60% of ventilator-dependent patients have reduced cognition, ability to understand information, and reduced communication skills. In addition, patients' life activities are completely dependent on the assistance of others, whose pain and stress are no less than patients. The use of ventilators may not improve the end-of-life (EoL) quality of terminally ill patients, usually only prolonging low-quality life, and may cause survival or psychological distress to the patient's family [1]. Therefore, after the third revision of the "Regulations on Tranquility and Alleviation" was promulgated and implemented on January 9, 2013, for patients who remain in a coma after intubation and cannot breathe on their own, if they are confirmed as terminally ill by two specialist physicians, they can be treated with withdrawal of LST with consent by a family member. However, in clinical practice, first-line medical personnel, especially nursing professionals, often found that there are divergent opinions among family members, and it is difficult to come up with an effective final decision. In fact, the evaluation and selection of life-sustaining medical treatment for patients with long-term ventilator dependence often involves multi-faceted considerations, which are multi-criteria decision-making problems involving multiple levels, and there are dependencies and complex relationships among the elements at each level. Therefore, to reach a best decision, how to objectively and effectively assign appropriate weights to each decision-making criterion, and find out the possible relative relationship between each criterion so as to assist decision-makers (medical staff and family members) in the evaluation and withdrawal of life-supporting medical methods is an important issue. The purpose of this study is to use the well-known analytic approach AHP to construct a set of evaluation and decision-making models, from the perspective of nursing professionals, for family members of patients with long-term ventilator dependence on withdrawal of life-support medical decisions.

2. Literature Review

When a patient falls into a coma and is unconscious, he/she is incapacitated and unable to communicate with others. Usually the family members must make decisions on behalf of the patient. Under such a circumstance, no matter whether the patient's willingness has been delivered, the family members will fall into a decision-making dilemma. This could be due to the fact that the patient and his/her family had not discussed the important decision before reaching the end of life. Therefore, they had no choice but to rely on life sustaining treatment (LST). Nevertheless, while the LST may prolong the patient's life, the quality of life suffers seriously. Meanwhile, the family members might suffer as well. All these highlight the importance of advance medical decision-making [2].

2.1. *Ventilation-dependent*

Ventilator-dependent patients refer to those who have used the respirator continuously for more than 21 days (inclusive), and the interruption time of the use of the respirator does not meet the definition of successful weaning off the respirator [3]. A study by Steinz et al. (2022) found that 42.2% of prolonged mechanical ventilation (PMV) patients admitted to the chronic respiratory care unit were in vegetative state/ minimally conscious state (VS/MCS); 32.5% had severe cognitive impairment, and 11.0% had mild to moderate cognitive impairment. Mortality was associated with poorer awareness and cognitive status. These findings underscore the importance of discussing end-of-life decisions regarding resuscitation and/or intubation in these chronically intubated patients [4].

The American Thoracic Society (ATS) issued a statement on the withdrawal of life-sustaining medicine: Respecting the patient's autonomy is the ethical basis for withdrawing and not granting life-sustaining treatment. When a patient lacks decision-making capability, the designated healthcare agent can make the decision on behalf of the patient to withdraw life-sustaining treatment when it is ineffective.

2.2. *Withdrawal from life-sustaining treatment*

"Withdrawal of life-sustaining treatment" literally means passive inaction. When a terminally ill patient cannot avoid death even with the use of life-sustaining treatment, withdrawal of LST might indicate a way of respect of life and sustaining of dignity for the patient [5]. Solomon et al. (2005) found that medical staff's acceptance of withdrawal of life-sustaining medicine was lower than that of non-administration of life-sustaining treatment. This could be attributed to their subjective perception that withdrawal of life-sustaining medicine implies a negative action and allowing patients to die naturally is otherwise [6].

2.3. *Key factors of withdrawal from life-sustaining treatment*

There are no clear standards for when life-sustaining treatment should be terminated or withdrawn for terminally ill patients. Many factors play a role individually and collectively, such as ethics, socioeconomic status, culture, legal regulation, and family concerns [7], which could be classified into three categories as follows.

2.3.1. Behavior and attitude

The key factors include: (1) families' perception towards withdrawal of life-sustaining medicine, especially "regulation knowledge" and "attitude regarding promotion of palliative care" [8, 11]; (2) expected outcome of disease, i.e. the patient's prognosis[4], with "poor prognosis" boosting the willingness to withdrawal [9]; (3) personal and social values, especially guilt related to filial piety, decision-making and responsibility[10]; (4) willingness to care, referring to family cohesion, which serves as a predictor of decision making; (5) preference for disease treatment (without other chronic comorbidities) [8, 11]; and (6) culture and religious belief by taking into account the patient's cultural background [11, 12].

2.3.2. Subjective norms

There are six key factors, including: (1) family consensus [10]; (2) medical personnel attitudes (values) towards the withdrawal of life-sustaining medicine [13]; (3) medical personnel's experience with palliative medical [10]; (4) family dynamics, referring to different opinions among family members and decision-making conflicts; (5) assurance of quality of life, with focus on relieving patients' pain, respecting the patient's autonomy, avoiding excessive ineffective or inhumane medical treatment, allowing patients to pursue the rights and interests of hospice [13]; and (6) holding cross-team family meetings and ethical consultation to come up with shared decision-making [14].

2.3.3. Perceived behavioral control

Included are six key factors: (1) economic considerations associated with socioeconomic status, cultural belief, legal regulation and family concern [7]; (2) ethical and legal issues involving a complex balance between medical, legal, and ethical considerations; (3) knowledge related to end-of-life care to deal with symptom control after withdrawal/termination of life-support treatment [2]; (4) timing of communication [10]; (5) family members support with emotional comfort[2] and (6) grief counseling for dying family members.

3. Materials and Methods

This is a quantitative survey research study. Nursing professionals from a government-affiliated region-level teaching hospital in southern Taiwan were potential participants. First, based on an action-research design, a structured questionnaire composed of three constructs (18 sub-constructs) were developed through a literature review. Second, senior nursing professionals with abundant experience in caring VDPs were recruited for data collection. Third, the analytic hierarchy process (AHP) was adopted to collect and analyze the participants' responses.

3.1. *Research tools*

This study aimed to explore the key success factors of withdrawal of life-sustaining medical treatment for long-term ventilator-dependent patients, and provide a reference for clinical medical assistance to family members in medical sharing decision-making. At the first stage, a thorough review of literature was conducted and three main facets and 18 secondary facets were collected. At the second stage, a questionnaire was developed and distributed to the targeted nursing professionals. At the third stage, the Analytic Hierarchy Process (AHP) was adopted for data analysis.

AHP is a multi-criteria decision making (MCDM) approach by using both empirical data as well as subjective judgements of the decision-maker [15]. It decomposes a complex problem into several different meaningful constructs (criteria) and further into measurable sub-constructs (indicators), if necessary.

The computational procedures are described as follows.

Step 1. Pairwise compare the constructs and sub-constructs for their importance in the decision;

Step 2. Compute eigenvalues and eigenvectors;

Step 3. Compute inconsistency ratio.

Through the process, an overall priority value (weight) will be derived for each sub-constructs (alternatives). It is suggested that inconsistency ratios should be less than 0.1 or so to be considered reasonably consistent in completing serial pairwise comparisons.

In this study, the package software Expert Choice 11 (Expert Choice, Inc.) was adopted for further data analysis.

3.2. *Research process*

Figure 1 illustrates the research process for this study. After data were collected and built into the analytic model, inconsistency ratios were checked first to ensure the acceptable consistency. In this study, all ratios were less than 0.1, indicating reasonable pairwise judgement of priority between each pair of indicators for the solicited 10 questionnaires.

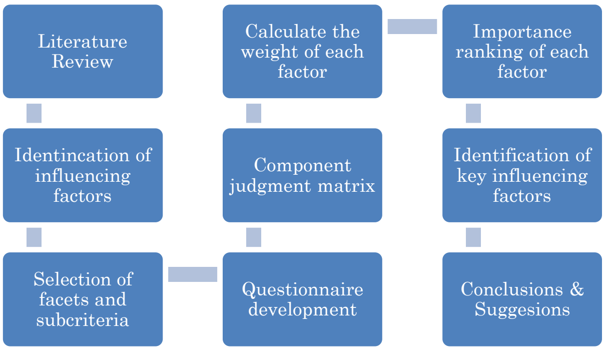

Literature Review
Calculate the weight of each factor
Importance ranking of each factor
Identincation of influencing factors
Component judgment matrix
Identification of key influencing factors
Selection of facets and subcriteria
Questionnaire development
Conclusions & Suggesions

Figure 1. Illustration of research process.

The process of this study can be detailed as follows.

Phase1. Through a literature review, 18 sub-criteria applicable to the withdrawal of life-sustaining treatment for long-term ventilator-dependent patients were collected, which can be grouped into three main constructs (facets).

Phase 2. Design the questionnaire based on the facets and sub-criteria (factors) and distributed to the targeted nursing professionals.

Phase 3. Use the Expert Choice 11 to calculate the relative weight. Based on the results of relative weight, explain the meaning of each weight index.

Phase 4. Analyze relevant weights, and further discuss the application of research results in practice.

4. Research Results

Targeted on nursing professionals with abundant experience in caring VDPs, 10 copies of questionnaires composed of three constructs (18 sub-constructs) were collected and retained for this survey research study. The analytic results of relevant weight (priorities) for each sub-construct are demonstrated in Table 1.

Table 1. Relevant weights for each construct and sub-construct.

Constructs	Sub-constructs (factors)	Wc	Wi	Wci
Behavior & attitude		0.323		
	family perception		0.097	0.046
	exp. disease outcomes		0.264	0.068
	values		0.438	0.158
	willingness to care		0.081	0.015
	preference for treatment		0.076	0.021
	culture and region		0.045	0.015
Subjective norms		0.577		
	family consensus		0.373	0.199
	staff values		0.109	0.086
	staff caring experience		0.036	0.028
	family dynamics		0.177	0.139
	quality assurance		0.269	0.212
	family meeting		0.038	0.030
Perceived behavioral control		0.101		
	economic considerations		0.237	0.024
	ethical and legal issues		0.135	0.014
	EoL caring knowledge		0.218	0.022
	communication timing		0.082	0.008
	family emotion		0.295	0.030
	grief counseling		0.035	0.004

Wc: construct weight; Wi: sub-construct weight; Wci: overall weight (= Wc x Wi).

Overall, “subjective norms” (Wc=.577) was considered relatively important than “behavior/attitude” (Wc=.323) and “behavior control” (Wc=.101). Further, “family consensus” (Wi=.199) was considered as the first priority, followed by “quality assurance” (Wi=.196) and “individual value” (Wi=.159). On the contrary, “grief counseling” (Wi=.004), “communication timing” (Wi=.008) and “ethics and legal” (Wi=.014) were with the least priorities. Family consensus and quality assurance are the two key factors when family members confront the decision whether to withdraw the LST for their beloved person.

5. Conclusions

Families often face difficult decisions when caring for terminally ill beloved ones. Better communication between health care providers and family members, especially primary caregivers, to reduce negative perceptions towards palliative care can help reduce psychological confusion and distress. Through education, training and role-playing, medical personnel may realize the importance of the role in end-of-life care, and have less moral remorse when implementing it. Establishing a standardized medical order for withdrawing life-sustaining treatment will also help medical personnel play a better role in shared decision making scenarios.

As is said, "No matter how advanced medicine is, it will never be able to cure all diseases." The traditional first aid process for terminally ill patients not only fails to achieve the purpose of prolonging life, but instead aggravates the pain of terminally ill patients and prolongs the process of dying. For those who have to face complete disability and must continue to rely on life-support medical care for the rest of their lives, quality of life sometimes becomes an even more urgent issue.

In summary, “subjective norm” is perceived as a relatively important criterion in making decisions regarding withdrawal of life-sustaining treatment. Among the sub-constructs, “family consensus” and “quality assurance” outweigh the other factors. It is thus suggested that encouraging family members to reach consensus for their beloved ones’ end-of-life decisions becomes a top priority.

References

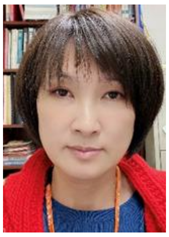
1. Meeker, M. A., & Jezewski, M. A. (2009). Metasynthesis: withdrawing life-sustaining treatments: the experience of family decision-makers. Journal of Clinical Nursing, 18(2), 163-173. 10.1111/j.1365-2702.2008.02465.x
2. Pei-Ni Chuang., Jaw-Shiun Tsa., Hsien-Liang Huang. (2018). Ethical Dilemmas in Withdrawing Life Sustaining Treatments in Geriatric Patients. Geriatrics gerontology, 13(1), 28-44.
3. Health insurance consultation (2022). https://www.nhi.gov.tw/Content_List.aspx?n=A27FC07E1D6ACC13&topn=5FE8C9FEAE863B46
4. Stein, D., Sviri, S., Beil, M., Stav, I., & Marcus, E. (2022). Prognosis of Chronically Ventilated Patients in a Long-Term Ventilation Facility: Association with Age, Consciousness and Cognitive State. Journal of Intensive Care Medicine, , 08850666221088800.
5. Savelkoul, C., de Graeff, N., Kompanje, E. J., & Tjan, D. H. (2016). Treatment in the Intensive Care Unit: continue o withdraw? Nederlands Tijdschrift Voor Geneeskunde, 160, A9694.
6. Solomon, M. Z., Sellers, D. E., Heller, K. S., Dokken, D. L., Levetown, M., Rushton, C., Truog, R. D., & Fleischman, A. R. (2005). New and lingering controversies in pediatric end-of-life care. Pediatrics, 116(4), 872-883.
7. Morgan, J. (2015). How do you decide when to withdraw life support? The Lancet Respiratory

Medicine, 3(6), 430-431.

8. Lobo, S. M., De Simoni, F. H., Jakob, S. M., Estella, A., Vadi, S., Bluethgen, A., Martin-Loeches, I., Sakr, Y., & Vincent, J. (2017). Decision-making on withholding or withdrawing life support in the ICU: a worldwide perspective. Chest, 152(2), 321-329.
9. Hoel, H., Skjaker, S. A., Haagensen, R., & Stavem, K. (2014). Decisions to withhold or withdraw life-sustaining treatment in a N orwegian intensive care unit. Acta Anaesthesiologica Scandinavica, 58(3), 329-336.
10. Chen, Y., Loh, E., & Huang, T. (2020). Humanity behind the intention of primary caregiver to choose withdrawing life-sustaining treatment for terminating patients. Patient Education and Counseling, 103(12), 2477-2482. 10.1016/j.pec.2020.06.011
11. Huynh, T. N., Walling, A. M., Le, T. X., Kleerup, E. C., Liu, H., & Wenger, N. S. (2013). Factors associated with palliative withdrawal of mechanical ventilation and time to death after withdrawal. Journal of Palliative Medicine, 16(11), 1368-1374.
12. Chai, H. Z., Krishna, L. K. R., & Wong, V. H. M. (2014). Feeding: what it means to patients and caregivers and how these views influence Singaporean Chinese caregivers' decisions to continue feeding at the end of life. American Journal of Hospice and Palliative Medicine®, 31(2), 166-171.
13. Brown-Saltzman, K., Upadhya, D., Larner, L., & Wenger, N. S. (2010). An intervention to improve respiratory therapists' comfort with end-of-life care. Respiratory Care, 55(7), 858-865.
14. Kon, A. A., Davidson, J. E., Morrison, W., Danis, M., & White, D. B. (2016). Shared decision making in intensive care units: an American College of Critical Care Medicine and American Thoracic Society policy statement. Critical Care Medicine, 44(1), 188.
15. Saaty, T. L. (1980). The Analytic Hierarchy Process McGraw Hill, New York.

Authors Introduction

Hsiao-Fang Chen, MHA, RN



She received her MHA from the Department of Health Care Administration, Chang Jung Christian University (CJCU), Tainan, Taiwan. Currently, she works full time as a Nursing Supervisor with specialization in respiratory therapy, medical administration, and health consultation and management at a government-affiliated hospital in southern Taiwan. She is also a doctoral program student in the College of Management, CJCU.

Jin-Yuan Chern, Ph.D.

He received his Ph.D. degree from the Medical College of Virginia/Virginia Commonwealth University, Virginia, US. Currently, he is an assistant professor at the Department of Health Care Administration, Chang Jung Christian University. His research interests include organization theory and management, and data applications and decision making.

Android Based Educational Mobile Robot Design and Pilot Evaluations

Elvira Chebotareva, Maksim Mustafin
Intelligent Robotics Department, Kazan Federal University, 420008, Kazan, Russian Federation
E-mail: elvira.chebotareva@kpfu.ru
kpfu.ru/robolab.html

Abstract

Educational robotics is a powerful tool for STEM and STEAM education. However, in practice, while trying to introduce educational robotics into a well-established educational process, a number of significant problems might arise. These include a high cost of an equipment and a lack of necessary competencies in the field of educational robotics among teachers. In this paper, we describe a concept of an inexpensive mobile robot equipped with a mobile device running Android operating system. We present sample projects that demonstrate how to extend capabilities of the educational mobile robot through the use of the mobile device. The proposed approach was preliminary evaluated in “Educational robotics” Bachelor level course and “Introduction to robotics” Master level course at Kazan Federal University.

Keywords: Educational robotics, Educational robot, Android based robot, Arduino based robot

1. Introduction

Currently, educational robotics is successfully used as an effective tool for STEM education at various educational levels[1][2][3]. Educational robotics as an interdisciplinary area within STEM and STEAM education[4] forms students' comprehensive knowledge, skills and abilities based on the integration of science, technology, engineering, art and mathematics.

Some of the technologies used in robotics and related industries are available for study as part of STEAM education[5][6][7]. Of particular interest is the implementation of elements of machine vision and artificial intelligence in robotic projects of students[8].

At the same time, in practice, when trying to introduce educational robotics into the educational process, many educational organizations face a number of problems that do not allow them to achieve the positive effects expected from the introduction. Among these problems, there is an acute shortage of specialists with the necessary knowledge in the field of educational robotics and a shortage of the necessary educational and methodological materials for the training of such specialists. In addition, in the difficult economic situation observed around the world, often educational institutions are faced with the high cost of logistics for educational robotics.

The purpose of this study is to find and test a low-cost technical solution suitable for studying the basics of educational robotics in higher educational institutions.

2. Project description

In order to level out the difficulties arising from the high cost of components, we use a following set of inexpensive components in the preparation of students studying educational robotics: Arduino Uno based controller or similar, DC motors, L298N motor driver module or similar, ultrasonic rangefinder HC-SR04 or similar, Bluetooth module HC-05 or similar, MPU-6050 gyroscope and accelerometer module, SG90 servo drive or similar, LEDs, RGB LED, piezo buzzer, resistors and photoresistors, optical encoders, connecting wires accumulator batteries 18650 with the holder and power button, rubber tire wheels.

Using this kit students can build a mobile robot for robotics classes. These components are easy to find on sale and they are easily replaceable using analogues. The cost of such a set, in market prices for November 2022, is about $50.

To assemble a mobile robot from the components listed above, students will also need robot's body parts. Currently, various options for inexpensive structural elements for assembling a robot chassis are offered for sale, however, in practice, we have encountered the fact that assembling a robot from these elements takes students a lot of time. Therefore, we have developed our own model of a mobile robot, the body elements of which are easy to manufacture using a 3D printer (Fig. 1). This model is adapted to the set of parts listed above, but can be easily modified by students to fit other similar components.

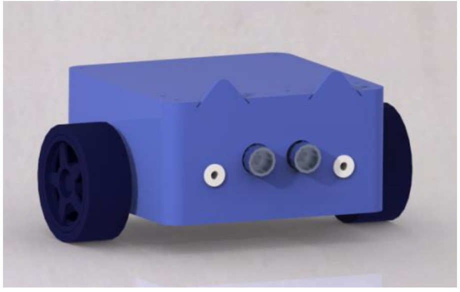


Fig. 1 The 3D-model of the robot

It should be noted that in the course of our work, we noticed that using the Arduino platform to build a robot encourages students to be creative and develop their own design of a robot body from improvised materials.

In order to expand the capabilities of a mobile robot assembled from the above components, we propose to additionally equip it with a mobile device running the Android operating system.

To solve our tasks, any device based on the Android 5 operating system and higher is suitable. We assume that these devices are accessible to many students through widespread adoption.

Thanks to mobile device an inexpensive mobile robot gets additional sensors, including cameras and microphone, additional processing power, as well as an improved human-machine interface.

Note also that additional equipment of robots with mobile devices can be used not only in educational projects, but also in scientific research[9][10].

As part of our work, we use the capabilities of smartphones or tablets based on the Android operating system to equip an educational robot with computer vision and human-machine interface elements. To exchange data between a mobile device and a mobile robot, it is convenient to use a wireless connection via Bluetooth or Wi-Fi.

Professional development tools for the Android platform, such as Android Studio, are not always suitable for use in the educational process. In practice, such tools often have a too high entry threshold for students, and their development requires a lot of time and special skills.

Therefore, to program the mobile robot we propose to use software adapted to the educational process.

For students who do not specialize in Android development, we suggest using the App Invnetor cloud platform[11]. MIT App Inventor is a handy tool for integrating a smartphone into robotics and IoT projects.

Another tool that can be used in the educational process for programming Android devices is the Processing programming language and development environment. The Android implementation of Processing[12] allows students to quickly create GUI applications.

We used the above described concept of a robot equipped with an Android mobile device to develop an "Educational Robotics" course for university students.

The proposed approach was preliminary evaluated in "Educational robotics" Bachelor level course and "Introduction to robotics" Master level course at Kazan Federal University.

Also, fragments of this course were presented as part of the advanced training program for teachers of additional general educational programs of a technical orientation in the IT-Cube digital education centers, conducted by the Academy of the Ministry of Education of the Russian Federation.

In the next section, we provide examples of educational projects implemented using the robot described above. These projects are part of the educational robotics course developed by us for students of higher educational institutions and are included in this course in the form of training assignments.

3. Examples of expanding the capabilities of the robot using a mobile device

3.1. *Mobile robot control*

This project allows to implement the control of a mobile robot in the teleoperation control mode. Students are encouraged to develop software that allows them to control the actions of a mobile robot equipped with a servo grip. Fig. 2 shows an example of an assembly scheme for such a robot from the components presented above. Fig. 3 shows an example of an application interface that allows to control the robot via a Bluetooth connection.

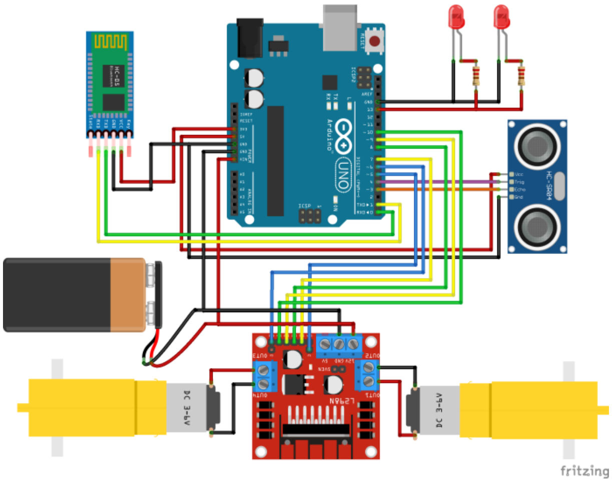


Fig. 2 Robot assembly diagram

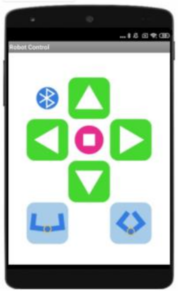


Fig. 3 Robot control application interface

The application interface (Fig. 3) contains 8 buttons: a button to connect to the robot via Bluetooth, 4 buttons to control the movement of the robot, a button to stop the robot and 2 buttons to control the grip of the robot.

3.2. *Search for a specified object*

This example demonstrates how the implementation of computer vision, artificial intelligence and human-machine interface elements can be combined in a training project.

In the conditions of the task, various objects are in front of the robot. The user says a voice command indicating the object to be found (for example, "Where is the cat?"). The robot must search for this object and drive up to it, and then inform the user that the object has been found. If the object could not be found, the robot also informs the person about this in the form of a voice message.

In Fig. 4 an example of an application interface for a mobile device installed on a robot is presented. The application interface contains a button to connect to the robot chassis via Bluetooth, a voice control activation button and a camera turn button. The application screen also displays the image received from the device's camera.

Visual search for objects is based on the application of the MobileNet neural network[13], previously trained on the ImageNet dataset. This network is undemanding to resources and is suitable for devices with low performance. In Fig. 5 a fragment of an experiment with searching for an image of a cat is presented.

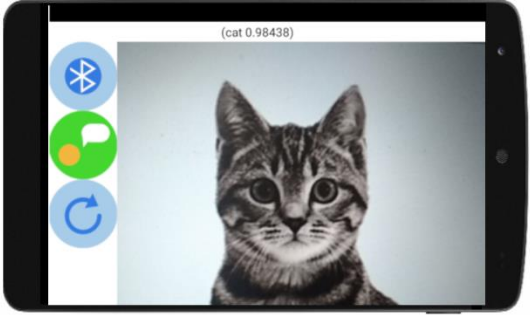


Fig. 4 Target object search application interface

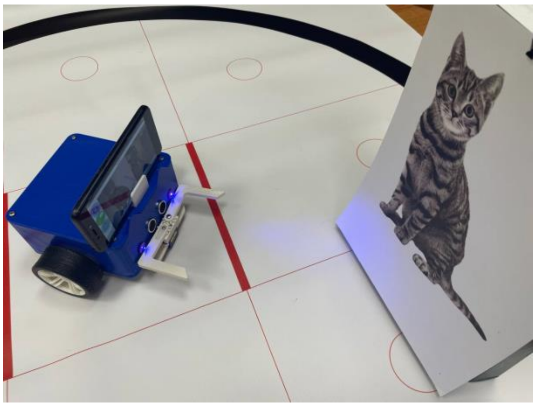


Fig. 5 Experiments with finding the cat

4. Conclusion

In this paper we proposed a conceptual scheme of an inexpensive mobile robot equipped with a mobile device based on the Android operating system. A feature of the proposed solution is a low cost and availability of components. Equipping the robot with a mobile device allows using additional sensors and computing resources in educational projects and expanding possibilities for a human-machine interface implementing. This concept was implemented on the model of a mobile robot developed by us and designed to study the basics of educational robotics in higher educational institutions involved in the training of teachers in the IT profile and related areas.

Acknowledgements

This paper has been supported by the Kazan Federal University Strategic Academic Leadership Program ("PRIORITY-2030").

References

1. Benit F B V, Exploring the educational potential of robotics in schools: A systematic review[J]. Computers & Education, 2012, Vol. 58(4): 978-988.
2. Zha S, Jin Y, Wheeler R, Bosarge E. A mixed-method cluster analysis of physical computing and robotics integration in middle-grade math lesson plans[J]. Computers & Education, 2022, Vol. 190: 104623.
3. Tsoy T, Sabirova L, Magid E. Towards effective interactive teaching and learning strategies in robotics education[C]//10th International Conference on Developments in eSystems Engineering (DeSE). IEEE, 2017: 267-272.
4. Aguilera Morales D, Ortiz-Revilla J. STEM vs. STEAM education and student creativity: a systematic literature review[J]. Education Sciences, 2021, Vol. 11: 331.
5. Tsoy T, Sabirova L, Lavrenov R, Magid E. Master program students experiences in robot operating system course[C]//11th International Conference on Developments in eSystems Engineering (DeSE).IEEE: 186-191.
6. Tsoy T, Sabirova L, Abramsky M, Magid E. Establishing effective teaching for robotics: a comparison study of Bachelor students participated in Introduction to Robotics course[C]//International Conference on Artificial Life and Robotics, 2018: 212-215.
7. Tsoy T, Sabirova L, Magid E. Effective robotics education: surveying experiences of master program students in introduction to robotics course[C]//MATEC Web of Conferences, Vol. 220(06005), 2018.
8. Tsoy T, Sagitov A, Magid E. Experiences of robotics students in machine vision course being taught in a foreign language: comprehension, self-efficiency, and active learning strategies improvement[C]//International Conference on Artificial Life and Robotics, 2019: 128-131.
9. Botzheim J, Kubota N. Spiking neural network based emotional model for robot partner[C]//IEEE Symposium on Robotic Intelligence in Informationally Structured Space, 2014: 1-6.
10. Woo J, Botzheim J, Kubota N. A modular cognitive model of socially embedded robot partners for information support[J]. ROBOMECH Journal, 2017, Vol. 4(10).
11. Patton E W, Tissenbaum M, Harunani F. MIT AppInventor: Objectives, Design, and Development. Computational Thinking Education[B], 2019: 31-49.
12. Colubri A. Processing for Android: Create Mobile, Sensor-Aware, and VR Applications Using Processing [B], 2017.
13. Howard A, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, Adam H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. ArXiV [OL], 2017.

Authors Introduction

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.

Mr. Maksim Mustafin

Maksim Mustafin is a bachelor student at Institute of Information Technology and Intelligent Systems, Kazan Federal University.

Virtual Collaborative Cells Modeling for UR3 and UR5 Robots in Gazebo Simulator

Ramir Sultanov, Shifa Sulaiman, Tatyana Tsoy, Elvira Chebotareva
Intelligent Robotics Department, Kazan Federal University, 420008, Kazan, Russian Federation
E-mail: sultan.ramir@it.kpfu.ru
kpfu.ru/robolab.html

Abstract

This paper presents virtual models of collaborative cells for two industrial collaborative robots UR3 and UR5 in the Gazebo simulator. Typically, the UR3 and UR5 robots are used by enterprises for packaging, assembly and sorting. Modeling and virtual experiments are an important stage in production processes planning, which involves joint human-robot work. Such models allow to plan safe human-robot interactions within a joint workspace and, if required, to rearrange the workspace. Our models of collaborative cells were adapted to several typical cases of joint human-robot operation scenarios and could be used in engineering design and testing for human-robot interaction in the field of production processes.

Keywords: Gazebo simulator, ROS, Collaborative industrial robot, Collaborative robot cell

1 Introduction

Simulation studies are often used for modeling the human-robot interaction in different processes[1][2][3], including the manufacturing processes involving collaborative robots[4][5][6]. Simulation models help to secure the safety of the human-robot interactions[7]. In particular, simulations are used for avoiding collisions in human-robot interactions[8][9]. Virtual models are also required for efficient and secure algorithm testing of collaborative robots before incorporating the robots in a real-world environment[10][11].

Nowadays, there are numerous simulation software available for simulating human robot interactions such as[12][13][14]. The selection of a software for evaluating human operator interactions with robots depends not only on the effectiveness of the modeling, but also on other elements like the software's accessibility and usability.

In this work, virtual collaborative cells for two industrial collaborative robots UR3 and UR5[15] using the open-source environment Gazebo[16] are developed for simulating industrial tasks. Gazebo is used to develop a variety of robotic systems since it can be integrated with the Robotic Operating System (ROS)[17][18].

UR3 and UR5 developed by the Universal Robots are used by large, medium-sized and small-scale industries. Development and testing of collaborative interaction of UR3 and UR5 robots with human operators are made easier using simulation studies before the implementation of these robots in industrial environments.

2 Typical use cases for collaborative industrial robots UR3 and UR5

During the last 10 years, a lot of research has been reported on the studies and implementation of collaborative robots in various industrial sectors[19][20]. As part of our research, we analyzed more than 100 cases of collaborative interaction of UR family robots with humans as part of the manufacturing process[15]. As a result, various applications of UR3 and UR5 robots and workspace prerequisites for incorporating the robots are identified.

We have identified the following common types of cases in which UR3 and UR5 robots are involved: collaborative assembly or processing, packaging and sorting. The most common cases of human-robot interaction work include transfer of objects from robot to person or vice versa and working on a common process simultaneously. The transfer of the part can be carried out either by shifting the part to a certain location or using a conveyor.

We have identified two options for arranging the collaborative workspaces. Workspaces of human operator and robot are located in adjacent zones in the first option. In the second option, the human operator and a robot work in a common area. In the second case, the workspace is difficult to distinguish.

Three cases of human robot workspace simulation models are developed in this work. A person sitting and standing in one location are considered as the first and second cases respectively. The third case involves the movement of a person to different locations inside the workspace of the robot.

A collaborative work area in which a robot and a human do a common job can also be integrated into a common space in various ways. A small area of a shared workspace with other people, in a shared office or office with conditional boundaries Part of a production workshop with other people and robots or a separate room can also be considered a collaborative environment. It should be noted that UR3 and UR5 robots are successfully used in small-sized rooms.

3 Modeling the virtual collaborative cell in Gazebo simulator

The following common traits of a virtual collaborative cell are identified based on the investigation of various applications of UR3 and UR5 robots. Cell sizes, type of the robot, workspace sizes, location of the working area relative to the cell, location of the robot relative to the work area, the position and trajectory of a person relative to the work area are important for the development of a collaborative work. These characteristics are considered during the development of simulation models of work cells in the Gazebo environment.

Simulation models of virtual collaborative cells in the Gazebo environment, taking into consideration the main characteristics of the collaborative robotic cell are developed. By updating the above-mentioned characteristics, this model can be used for a wide variety of applications. Cells that satisfy the requirements of a certain assignment can be produced with the use of this model. This model can be adapted to various constraints changing the key characteristics of the workspace. With the help of this model, cells that meet the conditions of a particular task can be generated. In addition to the robot itself, an animated model of a person standing or sitting at the workplace is created inside the cell. The user can control the motion of the person inside the cell.

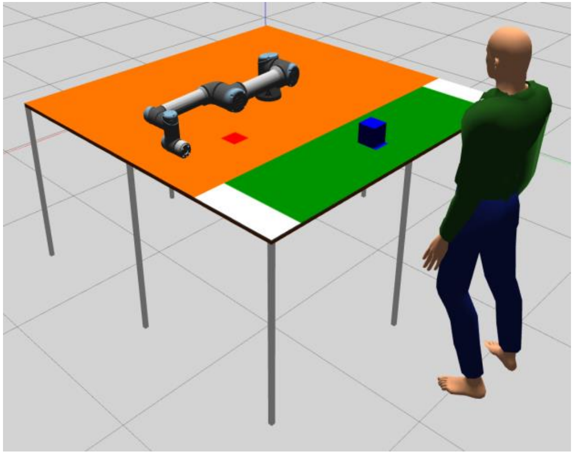


Fig. 1. The virtual model of collaborative workcell with the standing person and UR5 robot

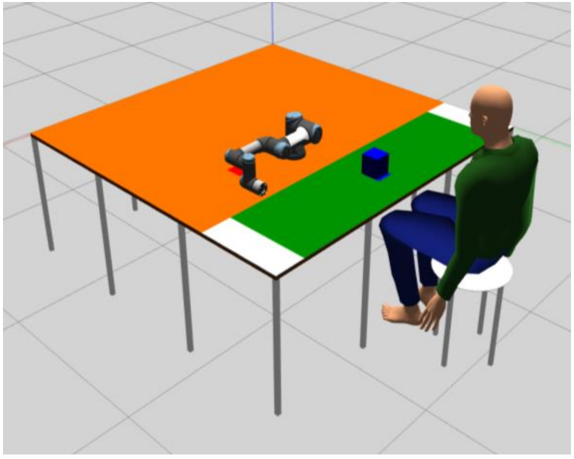


Fig. 2. The virtual model of collaborative workcell with the sitting person and UR3 robot

Fig. 1, Fig. 2, and Fig. 3 show various virtual collaborative cells constructed in Gazebo for simulating different workcell environments. Fig. 1 shows a person standing next to a table on which a UR5 robot is fixed. Fig. 2 shows a person sitting next to a table on which a UR3 robot is fixed. In Fig. 3 the integration of two collaborative workcells with UR3 and UR5 robots into the office space model is presented.

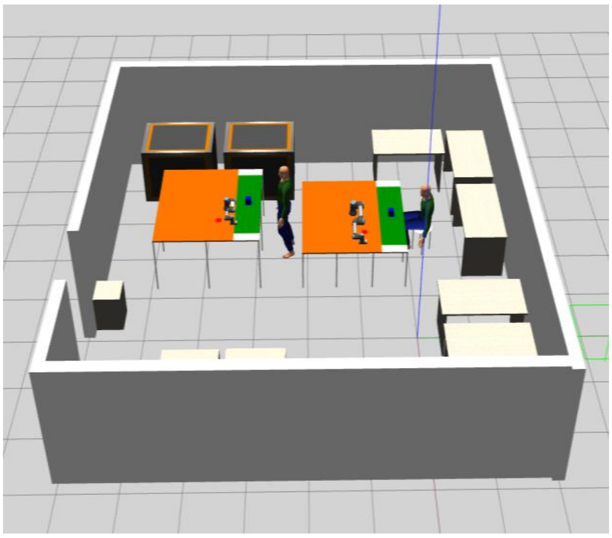


Fig. 3. The integration of two collaborative workcells with UR robots into the office space model

4 Conclusions and future work

Different applications of UR3 and UR5 robots are studied for determining the common characteristics of collaborative robots used in industrial applications along with human operators. UR3 and UR5 robots collaborate with human operators to complete tasks including assembly, packing, and sorting operations. In the Gazebo environment, a generic model is created which enables users to easily create virtual collaborative cells for UR3 and UR5 robots. These cells can be used to design and test different algorithms for the collaborative robots for robot-human interaction. The created virtual model can be expanded to different applications by including other types of manipulators and systems. The motion of human operators can also be planned for future works.

Acknowledgements

This paper has been supported by the Kazan Federal University Strategic Academic Leadership Program ("PRIORITY-2030").

References

1. Niehaus S, Ajoudani A, Bianchi M, Durandau G, Fritzsche L, Gaertner C, Mohamed I, Sartori M, Wang H, Wischniewski S. Human-centred design of robotic systems and exoskeletons using digital human models within the research project SOPHIA [J/OL]. Zeitschrift für Arbeitswissenschaft, 2022.
2. Kaonain T E, Rahman M A A, Ariff M H M, Yahya W J, Mondal K. Collaborative Robot Safety for Human-Robot Interaction in Domestic Simulated Environments[C]//IOP Conference Series: Materials Science and Engineering, 2021, 1096(012029).
3. Chebotareva E, Hsia K H, Yakovlev K, Magid, E. Laser rangefinder and monocular camera data fusion for human-following algorithm by PMB-2 mobile robot in simulated Gazebo environment[C]//Proceedings of 15th International Conference on Electromechanics and Robotics "Zavalishin's Readings", 2021: 357–369.
4. Benotsmane R, Kovács G, Dudás L. Economic, Social Impacts and Operation of Smart Factories in Industry 4.0 Focusing on Simulation and Artificial Intelligence of Collaborating Robots[J/OL]. Social Sciences, 2019, 8, 143.
5. Realyvásquez-Vargas A, Arredondo-Soto K C, García-Alcaraz J L, Márquez-Lobato B Y, Cruz-García J. Introduction and configuration of a collaborative robot in an assembly task as a means to decrease occupational risks and increase efficiency in a manufacturing company[J]. Robotics and Computer-Integrated Manufacturing, 2019, 57: 315-328.
6. Pieskä S, Kaarela J, Mäkelä J. Simulation and programming experiences of collaborative robots for small-scale manufacturing[C]//2nd International Symposium on Small-scale Intelligent Manufacturing Systems (SIMS), 2018: 1-4.
7. Spitzhirn M, Liedtke M, Grün G, Matheis C. Simulation of work environment factors for human-oriented and efficient workplaces[J/OL]. Zeitschrift für Arbeitswissenschaft, 2022.
8. Metzner M, Utsch D, Walter M, Hofstetter H, Ramer C, Blank A, Franke J. A system for human-in-the-loop simulation of industrial collaborative robot applications[C]//IEEE 16th International Conference on Automation Science and Engineering (CASE), 2020: 1520-1525.
9. Gläser D, Fritzsche, L, Bauer S, Leidholdt W. The quest to validate human motion for digital ergonomic sssessment - biomechanical studies to improve the human-like behavior of the human model "EMA"[C]// Applied Human Factors and Ergonomics International, 2020, 11.
10. Badia S B, Silva P A, Branco D, Pinto A, Carvalho C, Menezes P, Almeida J, Pilacinski A. Virtual Reality for Safe Testing and Development in Collaborative Robotics: Challenges and Perspectives[J]. Electronics. 2022; 11, 1726.
11. Erős E, Dahl M, Hanna A, Albo A, Falkman P, Bengtsson K. Integrated virtual commissioning of a ROS2-based collaborative and intelligent automation system[C]//24th IEEE International Conference on Emerging Technologies and Factory Automation, 2019: 407-413.

12. Fritzsche L, Ullmann S, Bauer S, Sylaja V J. Task-based digital human simulation with Editor for Manual work Activities - industrial applications in product design and production planning[B], 2019: 569-575.
13. Bobka P, Germann T, Heyn J K, Gerbers R, Dietrich F, Dröder K. Simulation Platform to Investigate Safe Operation of Human-Robot Collaboration Systems[C]// 32nd European Conference on Modelling and Simulation, 2016, Vol. 44: 187-192.
14. Coelho F, Relvas S, Barbosa-Póvoa A P. Simulation of an order picking system in a manufacturing supermarket using collaborative robots[C]//32nd European Conference on Modelling and Simulation, 2018.
15. Collaborative robotic automation. Cobots from Universal Robots[W], 2022, https://www.universal-robots.com/
16. Gazebo[W], 2022, https://gazebosim.org/home
17. Shimchik I, Sagitov A., Afanasyev I., Matsuno F, Magid E. Golf cart prototype development and navigation simulation using ROS and Gazebo[C]//MATEC Web of Conferences, 2016, Vol. 75, 09005.
18. Sultanov R, Sulaiman S, Li H, Meshcheryakov R, Magid E. A Review on Collaborative Robots in Industrial and Service Sectors[C]//Siberian Conference on Control and Communications, 2022
19. Dobrokvashina A, Sulaiman S, Zagirov A, Chebotareva E, Hsia K-H, Magid E. Human Robot Interaction in Collaborative Manufacturing Scenarios: Prospective Cases[C]//Siberian Conference on Control and Communications, 2022
20. Galin R R, Shiroky A A, Magid E, Mescheriakov R V, Mamchenko M V. Effective functioning of a mixed heterogeneous team in a collaborative robotic system[J]. Informatics and Automation, 2021, 20(6): 1224-1253.

Authors Introduction

Mr. Ramir Sultanov

He is 20 years old. Currently he is studying at the Institute of Information Technology and Intelligent Systems of the Kazan (Volga Region) Federal University as a fourth-year-student of the "Software Engineering" Bachelor degree program.

Ms. Shifa Sulaiman

In 2013, she received her Master's degree in Machine Design from Mahatma Gandhi (MG) University, India. In 2013-2017, Shifa worked as an Assistant Professor in various Indian Engineering Institutes. Since 2017, she has been doing PhD at National Institute of Technology, Calicut, India, specializing in Humanoid Robotics. She is currently working as a Research Associate at the Laboratory of Intelligent Robotic Systems (LIRS) at 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.

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.

New Features Implementation for Servosila Engineer Model in Gazebo Simulator for ROS Noetic

Alexandra Dobrokvashina, Shifa Sulaiman, Timur Gamberov
Intelligent Robotics Department, Kazan Federal University, 420008, Kazan, Russian Federation

Kuo-Hsien Hsia
Department of Electrical Engineering, National Yunlin University of Science & Technology, 64002, Douliou, Taiwan

Evgeni Magid
Intelligent Robotics Department, Kazan Federal University, 420008, Kazan, Russian Federation
Higher School of Economics University, Moscow, Russian Federation
E-mail: dobrokvashina@it.kfu.ru
kpfu.ru/robolab.html

Abstract

Virtual experiments play an important part in robotics allowing to reproduce complex environments, perform complicated and risky tasks. Yet, a virtual model is not always a one-time build action and it requires revisions in a timely manner as operating systems and dependent software evolves. This article presents a number of technical updates of the Servosila Engineer crawler type robot virtual model. The model evolution necessity was caused by a migration from an outdated robot operating system (ROS) of Melodic version to the modern ROS Noetic version. In addition to migration issues, for the robot virtual model a new onboard torch control unit and a robot head aligning unit were developed.

Keywords: ROS, Simulation, Servosila Engineer crawler type robot, Gazebo, Webots.

1. Introduction

A simulation is useful for evaluating new algorithms in complex virtual environments. It decreases chances of serious robot’s malfunctions during real time experiments due to conceptual or coding errors. Simulators are often employed in a wide variety of situations, including manufacturing[1][2], medical[3][4], and urban search and rescue operations[5]. Additionally, a simulation allows to easily replicate complex environments[6]. This article presents an updated virtual model of the crawler robot Servosila Engineer and demonstrates a methodology of enhancing the model in accordance with actual robot behavior. The virtual model is upgraded from ROS Melodic to ROS Noetic version.

2. Real robot and its virtual models

The Servosila Engineer robot (Fig.1) was designed by Russian company Servosila[7]. The robot is equipped with four cameras, a laser rangefinder, an inertial measurement unit sensor, and a light source that allows acquiring data about surroundings in absence of external light sources.

A robot virtual model for the Webots simulator[8] is one of our most recent research initiatives. Webots has an excellent physics-based realization, a high level of visual realism and already contains tracks for a crawler-type robot that are typically difficult to generate. The simulation model of the Servosila Engineer robot is shown in Fig.2.

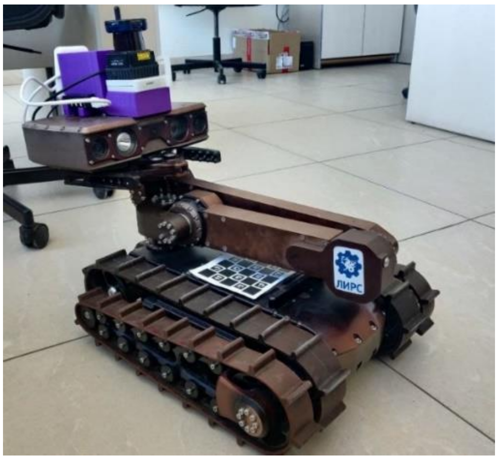


Fig.1. Servosila Engineer robot

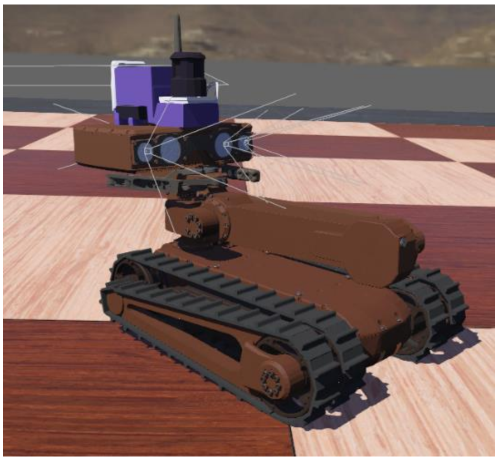
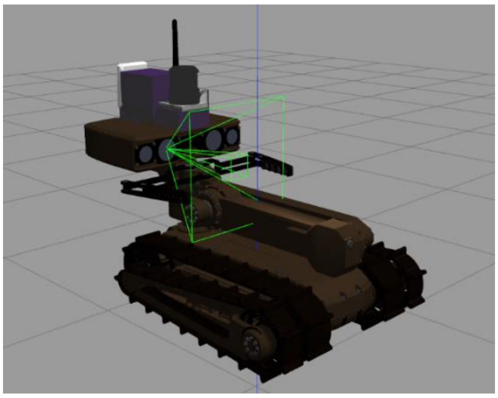


Fig.2. Webots simulation model

Even though various options are available for creation of different types of crawlers and ROS-based systems in Webots, it has a number of disadvantages, e.g., a lack of some types of plugins, including *mimic* joints that are widely presented in the Servosila Engineer model[9]. Separate scripts are needed to generate *mimic joints* to complete the robot model. The current version of the Servosila Engineer with all controllers and on-board sensors was presented in[10].

Several Servosila Engineer models are currently available in Gazebo simulator. Since there are no built-in virtual crawler models for a tracked base, solutions using various wheels to approximate traditional tracks were proposed[11]. A variety of plugins that enable creation of mimic joints and additional artificial light sources made it a valuable option. The Servosila Engineer model for Gazebo simulator (Fig.3) with controllers and a navigation algorithm[12] was demonstrated in[13].

Fig.3. Gazebo simulation model



3. New features of the Gazebo model

3.1. *ROS upgrade*

One of the goals was to migrate the Servosila Engineer model from ROS Melodic to Noetic version. If a ROS-project does not include specific packages, a migration typically easy and goes smoothly. Yet, one of common packages, *tf*, appeared to have issues while switching from Melodic to Noetic version. In all previous versions of ROS *tf* package allowed users to add namespaces for individual robots in order to launch several homogenous robots within a single simulation. However, the first version of ROS Noetic removed this function. To resolve this issue, we manually compiled an earlier version of *tf* package that is still functionally sound. Due to a standard type of ROS packages, a compilation process was successfully completed.

3.2. *Elbow-neck motion*

When the real robot moves, its head joint moves simultaneously with the elbow joint (up or down) to create the head motion and align the head with the horizon. While this feature had been missing in the previous Gazebo model of the robot, a new ROS-node was created to provide such motion. After calculations, the node resends to the head joint commands, which were originally delivered to the elbow joint. The control scheme for the elbow-neck joint motions is shown in Fig.4.

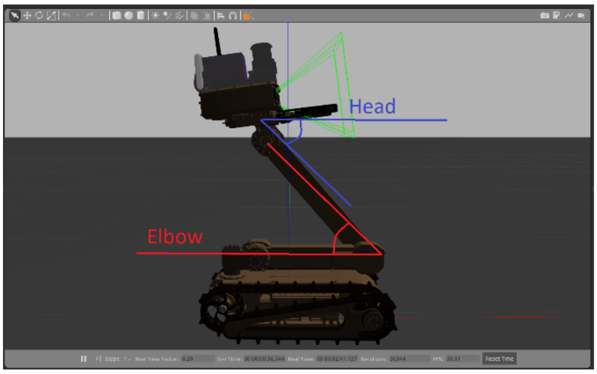


Fig.4. Elbow-neck joints movement scheme (red angle – elbow joint, blue angle – neck joint)

3.3. *Onboard torch*

Migration between ROS versions caused problems with the torch presented in Gazebo simulation model. In Melodic version to manipulate a torch status the following command is used:

```
rosservice call /gazebo/set_light_properties '{light_name: "spot", diffuse: {r: 0.0, g: 0.0, b: 0.0, a: 1.0}, attenuation_constant: 0.8, attenuation_linear: 0.01, attenuation_quadratic: 0.0}'
```

Table 1. Parameters for setting up the light in Gazebo in different versions of the ROS

Parameter name	*ROS Melodic*	*ROS Noetic*
light_name	yes	yes
diffuse	yes	yes
specular	**no**	**yes**
attenuation_constant	yes	yes
attenuation_linear	yes	yes
attenuation_quadratic	yes	yes
direction	**no**	**yes**
pose	**no**	**yes**

Using the above listed command, the torch calling service responsible for light sources in the scene could be turned off. The same service is available in Noetic version. However, the same command causes the Gazebo simulator crash. To fix this issue, the service structure was modified and the new light control command was formed as follows:

```
rosservice call /gazebo/set_light_properties '{light_name: "spot", diffuse: {r: 255.0, g: 255.0, b: 255.0, a: 255.0}, specular: {r: 230.0, g: 230.0, b: 230.0, a: 255.0}, attenuation_constant: 0.8, attenuation_linear: 0.01, attenuation_quadratic: 0.0, direction: {x: 0.0, y: 0.0, z: 0.0}, pose: { position: {x: 0.0, y: 0.0, z: 0.0}, orientation: {x: 0.0, y: 0.0, z: 0.0, w: 0.1}}}'
```

This command successfully turns the torch off. The light parameters used for the Servosila Engineer robot in ROS Melodic and Noetic are listed in Table 1.

4. Conclusion

This article presents a number of technical updates of the Servosila Engineer crawler type robot virtual model. The model evolution necessity was caused by a migration from an outdated ROS of Melodic version to the modern ROS Noetic version. In addition to migration issues, for the robot virtual model a new onboard torch control unit and a robot head aligning unit were developed and validated.

Acknowledgements

This paper has been supported by the Kazan Federal University Strategic Academic Leadership Program ("PRIORITY-2030").

References

1. Gunal M M. Simulation for industry 4.0. Past, Present, and Future[B]. 2019.
2. Kuts V, Otto T, Tähemaa T, Bondarenko Y. Digital twin based synchronised control and simulation of the industrial robotic cell using virtual reality[J]. Journal of Machine Engineering, 19(1), 2019: 128-145.
3. Shah D, Yang B, Kriegman S, Levin M, Bongard J, Kramer-Bottiglio R. Shape changing robots: bioinspiration, simulation, and physical realization[J]. Advanced Materials, 33(19): 2002882.
4. Cornejo J, Cornejo-Aguilar J A, Palomares R. Biomedik surgeon: surgical robotic system for training and simulation by medical students in Peru[C]//International Conference on Control of Dynamical and Aerospace Systems, 2019: 1-4.
5. Nirovi F, Zhang K, Kashino Z, Nejat G. Deep reinforcement learning robot for search and rescue applications: Exploration in unknown cluttered environments[J]. Robotics and Automation Letters, 4(2), 2019: 610-617.
6. Lewis M, Sycara K, Nourbakhsh I. Developing a testbed for studying human-robot interaction in urban search and rescue[C]//Human-Centered Computing, 2019: 270-274.
7. Michel O. Cyberbotics Ltd. Webots™: professional mobile robot simulation[J]. Journal of Advanced Robotic Systems, 1(1), 2004: 39-42.

8. Dobrokvashina A, Lavrenov R, Bai Y, Svinin M, Magid E. Sensors modelling for Servosila Engineer crawler robot in Webots simulator[C]//Moscow Workshop on Electronic and Networking Technologies, 2022: 1-5.
9. Dobrokvashina A, Lavrenov R, Magid E, Bai Y, Svinin M, Meshcheryakov R. Servosila Engineer Crawler Robot Modelling in Webots Simulator[C]//International Journal of Mechanical Engineering and Robotics Research, 11(6), 2021: 417-421.
10. Dobrokvashina A, Lavrenov R, Martinez-Garcia E A, Bai, Y. Improving model of crawler robot Servosila Engineer for simulation in ROS/Gazebo[C]//International Conference on Developments in eSystems Engineering (DeSE), 2020: 212-217.
11. Pecka M, Zimmermann K, Svoboda T. Fast simulation of vehicles with non-deformable tracks[C]//In 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2017:6414-6419.
12. Danilko A I, Stukach O V. Simulation of the Odometric Autonomous Navigation System of a Crawler Robot Optimal in Performance Criteria[J]. Automatics & Software Enginery (A&SE),1(39), 2022: 48-66.
13. Dobrokvashina A, Lavrenov R, Tsoy T, Martinez-Garcia E A, Bai Y. Navigation stack for the crawler robot Servosila Engineer[C]//Conference on Industrial Electronics and Applications, 2021: 1907-1912.

Authors Introduction

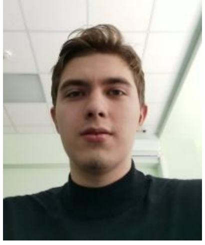
Ms. Alexandra Dobrokvashina

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.

Ms. Shifa Sulaiman

In 2013, she received her Master's degree in Machine Design from Mahatma Gandhi (MG) University, India. In 2013-2017, Shifa worked as an Assistant Professor in various Indian Engineering Institutes. Since 2017, she has been doing PhD at National Institute of Technology, Calicut, India, specializing in Humanoid Robotics. She is currently working as a Research Associate at the Laboratory of Intelligent Robotic Systems (LIRS) at Kazan Federal University, Russia.

Mr. Timur Gamberov



Timur Gamberov is a bachelor student at Institute of Information Technology and Intelligent Systems, Kazan Federal University.

Prof. 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. 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.

Prof. Kuo-Hsien Hsia

He is an Associate Professor at the Department of National Yunlin University of Science & Technology, Douliou, Taiwan. He received his PhD Degree from National Sun Yat-Sen University, Taiwan in 1994. His research interests include fuzzy systems, intelligent systems, and technical and vocational education.

Features of Interaction Between a Human and a Gestures-controlled Collaborative Robot in an Assembly Task: Pilot Experiments

Maksim Mustafin, Elvira Chebotareva
Intelligent Robotics Department, Kazan Federal University, 420008, Kazan, Russian Federation
E-mail: maksamustafin@kpfu.ru
kpfu.ru/robolab.html

Hongbing Li
Department of Instrument Science and Engineering, Shanghai Jiao Tong University, Shanghai, Minhang, 200240, China

Edgar A. Martínez-García
Department of Industrial Engineering and Manufacturing, Universidad Autónoma de Ciudad Juárez,
Juarez, 32315, Mexico

Evgeni Magid
Intelligent Robotics Department, Kazan Federal University, 420008, Kazan, Russian Federation
HSE University, Moscow, Russian Federation

Abstract

This paper presents results of pilot experiments that were run to study a human interaction with the UR5e collaborative 6-axis robot manipulator in a cooperative assembly task. The participants controlled the equipped with a screwdriver UR5e robot using computer vision and gestures. The purpose of the experiments was to identify the features of user interaction with the UR5e robot controlled with gestures in a task of a complex object assembly. Ten people took part in the experiments. The results of the experiments allowed to conclude on practical efficiency of cobots in joint assembly tasks. In addition, we identified preferable by the users location areas during the assembly task.

Keywords: Human–robot interaction, Human-robot collaboration, Collaborative robots, Collaborative Assembly

1. Introduction

Due to a significant progress of industrial automation, issues of human-robot interaction in a shared workspace[1] and efficient distribution of collaborative tasks between human and robotic agents[2] became critical in the past decades. In practice, a human-robot collaborative assembly task often arises in cases where a part of assembly process stages requires a human intervention[3]. An example is a situation when a part of assembly operations cannot be automated, or full automation of operations is impractical due to a high cost and complexity of a setup process. Of particular interest are cases in which human operations with assembly parts alternate with actions of a robot. In this case, the operator interacts with the robot through various communication technologies.

A positive user experience (UX) in human-robot interaction (HRI) is essential for an efficient organization of human-robot production processes. Despite a significant number of studies devoted to a methodology for assessing the UX in social robotics, a problem of choosing certain methods for assessing UX of a human-robot collaboration remains open[4].

Communication between humans and robots can take both verbal and non-verbal forms[5]. In practice, multi-modal interfaces can be used to communicate with industrial robots, including buttons and joysticks, a haptic control, speech recognition technologies, a gesture control, and a gaze recognition[6].

Specifics of a particular manufacturing process may impose some restrictions on the human-robot interaction. For example, tactile and button controls can distract an operator from an assembly process, and noise in a manufacturing area can prevent successful recognition of voice commands. In this case, an interaction with the robot through gestures seems to be the most convenient.

In this paper, we present results of pilot experiments that were run to study a human interaction with the UR5e collaborative 6-axis robot manipulator in a cooperative assembly task. Our study is aimed at identifying features of a user interaction with the gesture-controlled robot during the joint assembly task in order to further apply the results in practice when designing real manufacturing processes.

2. Related work

Collaborative robots are often considered safer for humans than industrial robots[7]. A number of review papers were devoted to a collaborative assembly implementation[8][9]. Papers[10][11][12] considered various aspects of the methodology for designing and implementing HRI in the context of assembly tasks.

Holm et al.[13] presented the results of an evaluation of a human-robot interaction at three industrial demonstrators: Mixed Packaging of Cheese, Aircraft Wing Rib Assembly and Automotive Engine Mass Balancing System Assembly. In [14] and [15], an influence of cognitive ergonomics on the interaction between a human and an industrial robot during a joint assembly was investigated. Neto et al. [16] proposed a gesture-based HRI structure in which a robot helps a person by passing tools and parts. Paper [17] presented the results of experiments with a collaborative robot controlled by gestures. For the gestures classification a taxonomy proposed in [18] was used.

Taking into account the literature review we focused on the following aspects in the course of preparing pilot experiments:

- Safety;
- Qualitative and quantitative evaluation of the human-robot interaction;
- Ergonomic requirements for the workspace.

3. Materials and methods

3.1. *Participants*

The experiments involved 10 people who are not professional operators of collaborative industrial robots. To guarantee an independent nature of the experiments the participants did not contact with each other during the experiment.

3.2. *Assembly task and workcell configuration*

During the experiment, participants were asked to assemble a chassis of a small mobile robot. The chassis consisted of several parts that should be fastened with screws. The experimental workcell (Fig. 1) included:

- UR5e robot equipped with a screwdriver;
- a work table;
- a camera for tracking the operator’s movement within a danger zone;
- a camera for a gesture recognition;
- a screen for displaying information for the operator;
- a chair for the operator.

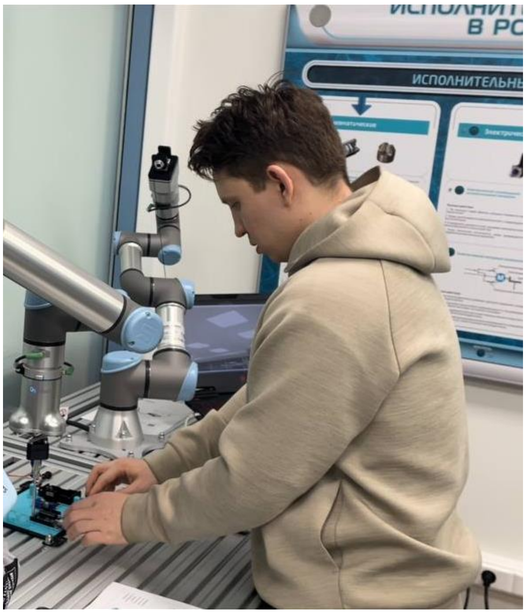


Fig. 1 The workcell during the experiment

3.3. *Experimental protocol*

The experiment consisted of three stages. At the first stage, the participants were trained to control the UR5e robot using predefined gestures. This stage was necessary in order to get acquainted with the system of robot control gestures. The set of gestures used consisted of symbols A, B, C of the American sign language ASL[19]: A was for power on of the UR5e robot, B was for playing or continuing the robot program, C was for pausing the program (Fig. 2). Participants also had an opportunity to test a safety system, which stopped the robot if the participant's hand entered a working area of the robot.

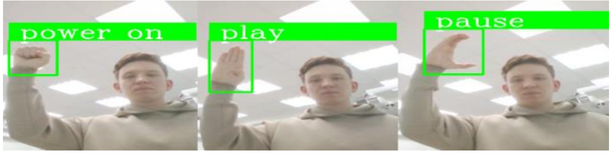


Fig. 2 The view from a camera for the gesture recognition. The participant shows gestures (from left to right): power on, play or continue the program, pause the program

At the second stage, the participant was asked to assemble the chassis manually (without the robot assistance) according to provided instructions.

At the third stage, the participants assembled the same chassis together with the robot. This stage was organized as follows. The participant placed several assembly parts on the chassis platform and passed them to the robot for tightening (Fig. 1) the screws. After the robot fixed the current parts on the platform, the participant placed new parts on the platform and again transferred the platform to the robot.

After the experiment, the participants were asked to take a survey that consisted of the following questions:

Q1: Was it easier for you to assembly the chassis with the UR5e robot assistance than without the robot?

Q2: How accurately did the UR5e robot execute commands based on your gestures?

Q3: How quickly did the UR5e robot respond to your commands?

Q4: How comfortable were you working with the UR5e robot?

Q5: How good did the UR5e robot perform its task?

Q6: What disadvantages in the robot operation could you note (if any)?

Q7: What gestures would you prefer to use?

We used a 5-point Likert-type scale [20] for questions Q1-Q5 and a free form for questions Q6 and Q7.

4. Results

Fig. 3 presents survey results (Q1-Q5). The x-axis (horizontal axis) indicates a question number. The y-axis (vertical axis) indicates a number of responses of participants for a particular answer to the questions (e.g., in the first question seven participants chose the forth answer option). A legend on the right side of the graph is indicates which color of the graph corresponds to a certain answer option in the questions. In general, the participants expressed satisfaction with both the work of the robot itself and the level of the comfort. However, two (of ten) participants noted that self-assembly of the chassis without a robot seemed faster to them.

In response to question Q6 two participants indicated the slow screw tightening by the robot, two participants noted the need to provide a more explicit feedback from the robot, e.g., with voice messages.

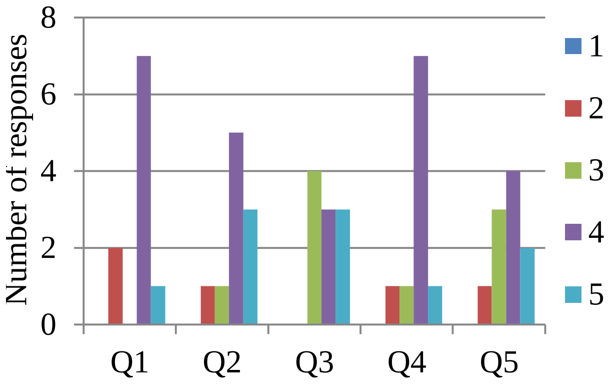


Fig. 3 The survey results. Avg: Q1 - 3.7, Q2 - 4.0, Q3 - 3.9, Q4 - 3.8, Q5 - 3.7

In response to question Q7, two participants suggested using gesture A to stop and gesture B to continue. Five participants indicated that gesture C was inconvenient to use. One participant suggested using gestures with one and two fingers instead of gestures A and B, respectively. The rest of the participants noted that all gestures were comfortable for them.

In the area of the workcell we placed a chair for the operator. Before the experiment, we informed each participant that he/she could sit or stand as desired. We observed that all participants periodically changed standing and sitting positions.

5. Discussion and conclusions

The conducted pilot experiments allowed us to formulate some hypotheses about the features of the user interaction with the gesture-controlled robot UR5e in the context of collaborative assembly.

The first hypothesis suggests to review the choice of the robot control gestures. We intentionally avoided using special gestures' sets for industrial robot control from a literature review (e.g., [21]) in order to test whether users could quickly adapt to gestures with an unusual semantic load. We noted that after the training phase, despite the fact that all users correctly used all gestures, some of them indicated their preferences. It can be assumed that in order to improve perception during the learning phase, along with the default gesture system, users should be prompted to select their own gesture system.

The second hypothesis relates to the speed of the robot's screw tightening actions. In our experiments, due to peculiarities of the end-effector, the robot performed the tightening rather slowly. This could negatively affect the perception of the robot efficiency by users. We assume that when using an automatic screwdriver end-effector and increasing the tightening rate, the negative estimates associated with the tightening rate would be leveled.

The third hypothesis concerns the operator position during the experiment in order to reduce a biomechanical load on the operator. In our experiments, the working surface on which the assembly parts were located allowed the operator to perform his/her work both sitting and standing. Most of the participants selected a sitting position during those stages of the experiment when the robot performed operations that did not require a human intervention and stood up when their intervention was required. Perhaps most users would prefer to do all work in a sitting position, if conditions of the work cell allow it. In addition, we concluded that the location of the camera that reads gestures should take into account the operator height in order to reduce the biomechanical load on his/her hands when showing the gestures.

The results of our experiments demonstrated that the use of the UR5e assisting robot for collaborative assembly tasks could be effective from a practical point of view. In our further studies, we plan to test the hypotheses we have put forward, as well as the degree of influence of the interaction features we have identified on the efficiency of the assembly process as a whole.

Acknowledgment

This paper has been supported by the Kazan Federal University Strategic Academic Leadership Program ("PRIORITY-2030").

References

1. Galin R, Meshcheryakov R. Review on human–robot interaction during collaboration in a shared workspace[C]//International Conference on Interactive Collaborative Robotics, 2019: 63-74.
2. Galin R, Meshcheryakov R, Kamesheva S. Distributing tasks in multi-agent robotic system for human-robot interaction applications[C]//Int. Conf. on Interactive Collaborative Robotics, 2020: 99-106.
3. Villani V et al. Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications[J]. Mechatronics. ScienceDirect, 2018, 55: 248-266.
4. Lindblom J, Alenljung B, Billing E. Evaluating the User Experience of Human–Robot Interaction[J]. Human-Robot Interaction. Springer, 2020, 12: 231-256.
5. Mavridis N. A review of verbal and non-verbal human–robot interactive communication[J]. Robotics and Autonomous Systems. ScienceDirect, 2015, 63(1): 22-35.
6. Gustavsson P, Holm M, Syberfeldt A, Wang L. Human-robot collaboration – towards new metrics for selection of communication technologies[J]. Procedia CIRP. ScienceDirect, 2018, 72: 123-128.
7. Sultanov R, Sulaiman S, Li H, Meshcheryakov R, E Magid. A Review on Collaborative Robots in Industrial and Service Sectors[C]//Siberian Conference on Control and Communications, 2022 (in press).
8. Rodríguez-Guerra D, Sorrosal G, Cabanes I, Calleja C. Human-Robot Interaction Review: Challenges and Solutions for Modern Industrial Environments[J]. IEEE Access. IEEE, 2021, 9: 108557-108578.
9. Dobrokvashina A, Sulaiman S, Zagirov A, Chebotareva E, Hsia K-H, Magid E. Human Robot Interaction in Collaborative Manufacturing Scenarios: Prospective Cases[C]//Siberian Conference on Control and Communications, 2022 (in press).
10. Mateus J C, Claeys D, Limère V, Cottyn J, Aghezzaf E-H. A structured methodology for the design of a human-robot collaborative assembly workplace[J]. The International Journal of Advanced Manufacturing Technology, 2019, 102(5): 2663–2681.
11. Gualtieri L et al. Safety, Ergonomics and Efficiency in Human-Robot Collaborative Assembly: Design Guidelines and Requirements[J]. Procedia CIRP. ScienceDirect, 2020, 91: 367-372.
12. Colim A et al. Human-Centered Approach for the Design of a Collaborative Robotics Workstation[B].

Occupational and Environmental Safety and Health II. Springer, 2020, 277: 379-387.

13. Holm M et. al. Real-World Industrial Demonstrators on Human–Robot Collaborative Assembly[B]. Advanced Human- Robot Collaboration in Manufacturing. Springer, 2021: 413-438.
14. Gualtieri L, Fraboni F, Marchi M D, Rauch E. Evaluation of Variables of Cognitive Ergonomics in Industrial Human-Robot Collaborative Assembly Systems[C]//Proceedings of the 21st Congress of the International Ergonomics Association (IEA 2021). Springer, 2021, 223: 266-273.
15. Fraboni F, Gualtieri L, Millo F, Marchi M, Pietrantoni L, Rauch E. Human-Robot Collaboration During Assembly Tasks: The Cognitive Effects of Collaborative Assembly Workstation Features[C]. Proceedings of the 21st Congress of the International Ergonomics Association (IEA 2021). Springer, 2021, 223: 242-249.
16. Neto P, Simão M, Mendes N, Safeea M. Gesture-based human-robot interaction for human assistance in manufacturing[J]. Int. J. Adv. Manuf. Technol. Springer, 2019, 101(1): 119–135.
17. Shukla D, Erkent Ö, Piater J. Learning Semantics of Gestural Instructions for Human-Robot Collaboration[J]. Frontiers in Neurorobotics. Frontiersin, 2018, 12: 7.
18. Quek F K H. Eyes in the interface[J]. Image and Vision Computing. ScienceDirect, 1995, 13: 511-525.
19. Wilbur R B, Nolkn S B. The Duration of Syllables in American Sign Language[J]. Lang speech. Sage journals, 1986, 29(3): 263-280.
20. Krägeloh C, Bharatharaj J, Kutty S K S, Nirmala P R, Huang L. Questionnaires to Measure Acceptability of Social Robots: A Critical Review[J]. Robotics. MDPI, 2019, 8(4): 88.
21. Shukla D, Erkent Ö, Piater J. A multi-view hand gesture RGB-D dataset for human-robot interaction scenarios[C]//25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN). IEEE, 2016: 1084-1091.

Authors Introduction

Mr. Maksim Mustafin

Maksim Mustafin is a bachelor student at Institute of Information Technology and Intelligent Systems, Kazan Federal University.

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.

Associate Professor Hongbing Li

He received his Ph.D. degree in Mechano-Micro Engineering at the Tokyo Institute of Technology, Japan. Currently he has been working at 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 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). His academic interests are mathematical modeling and dynamic control of robots.

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. 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.

Modeling of Human Actions in a Collaborative Robotic Space Using AR601M Humanoid Robot: Pilot Experiments in the Gazebo Simulator

Aidar Zagirov, Artem Apurin, Elvira Chebotareva
Institute of Information Technology and Intelligent Systems, Intelligent Robotics Department, Kazan Federal University, 420008, Kazan, Russian Federation
E-mail: ai.zag@it.kfu.ru
http://robot.kpfu.ru/eng

Abstract

To guarantee a safe human-robot collaboration, a collaborative system development requires a significant amount of real world experiments. Yet, it is critical to avoid injury risks for participants of such experiments. The risks could be reduced by introducing a virtual experiments' stage to detect mistakes in a robot behavior prior to the real world experiments. This paper presents a virtual model of a humanoid robot AR601M in the Gazebo simulator. Unlike the standard human models in the Gazebo, this model allows to simulate the gross and fine motor skills of a human and could be used when performing various human actions in collaborative robotic cells.

Keywords: Humanoid robot, Anthropomorphic robot, Gazebo simulator, Collaborative robotics

1. Introduction

Recently, human-robot interaction during collaboration in a shared workspace[1] and efficiency of distributing tasks between actors of different origin[2] attract great attention of researchers. Collaborative robots are often considered safer for humans than industrial robots[3]. However, unexpected failures, poorly thought-out security systems, and the human factor still maintain the risk of injury when a robot and a person work together. Expanding the possible cases of human interaction with a robot[4], building process maps and conducting experiments can increase the level of safety of the collaboration between a robot and a human.

Experiments with real people in the first stages of testing can be traumatic, and the use of anthropomorphic-type robots as human models can be unreasonably expensive. Therefore, the use of simulators and simulation models of a person for collaborative interaction modeling becomes an urgent task. One of the popular simulators used in robotics is Gazebo. However, when trying to apply ready-made animated human models available for the Gazebo[5][6] environment, we encountered several problems that do not allow them to be fully used in pilot experiments.

The first problem concerns the animation of the so-called actors in the Gazebo environment. Available by default Gazebo actor animations are limited to 9 types of actions. Creating custom actor animations requires to work with the COLLADA format files which is often inconvenient and time-consuming.

Another significant problem is that actors cannot interact with the environment. Because of this, it is impossible to simulate collisions of actors with physical objects. In addition, actors cannot be detected by most sensors.

Instead of Gazebo actors, we propose to use the AR601M anthropomorphic robot virtual model as human models. This robot can interact with the world. It has gross and fine motor skills, which is necessary for most pilot experiments.

The purpose of our research is to combine the related works into one model of AR601M robot, modeling human actions for this model, and perform pilot experiments.

2. Related Works

Some motor functions of the AR601M virtual robot model were partially implemented earlier. Khusainov et al. modeled AR601M locomotion with Walking Primitives approach in Simulink environment[7][8]. Walking primitives is a popular approach for biped robots that automatically generates joint trajectories[9]. The authors used simplified models of AR601M with 6 DoFs and 12 DoFs per legs. This approach showed good results in simulation, but it cannot be applied to the real AR601M robot because it doesn't support fast biped locomotion due to the leg's motor torque limitations.

Khusainov et al in the next work with AR601M robot model researched two walking stability algorithms in Simulink - virtual height inverted pendulum model (VHIPM) and preview control methods[10]. The preliminary VHIPM method has been modeled for AR601M robot and got onto his influences of robot height and step length[11]. As a result of the work, the preview control method turned out to be more reliable, than VHIPM method. The preview control method has feedback, it allows to avoid error accumulation for better robot locomotion. The authors decided to continue their work for both algorithms. They wanted to add calculations for peak walking speed simulations to VHIPM method and consider full body dynamics for high-speed walking to the preview control method.

Khusainov et al using knowledge of the above mentioned works addressed the issue of optimization gait pattern for a bipedal robot, which maximizes its locomotion speed under a joint angular velocity, actuator power limits and acceleration[12]. The authors compared kinematic[13] and dynamic approaches. They determined kinematic approach are applicable only if we use the upper body motion to balance the robot dynamics during walking. As opposed to kinematic approach, dynamic approach allows to use different acceleration and velocity limits depending on the payload applied to the actuated joint. This allowed them to reach physical limits for the swing leg.

Magid and Sagitov reviewed various approaches of fall detection and management procedures for different size humanoid robots and evaluated the possibility of implementations them for the AR601M robot[14].

Beside works with AR601M robot leg movements there is progress in hand control. Khusnutdinov et al. implemented household objects take and place tasks in Gazebo Simulator[15]. Authors configured mimic joints of right-hand fingers, focusing on the real AR601M robot possibility. In another work Khusnutdinov et al. presented grasping algorithm utilizes the simplicity of an antipodal grasp and satisfies force closure condition[16].

3. Robot AR601M and its Gazebo model

The AR601M robot shown in Fig. 1 is a bipedal walking robot of humanoid type. The robot is being developed by a Russian company Android Technics. Its height of 144 cm and weight of 65 kg. Robot has 41 active DoF. AR601M has the ability to grab objects and hold them. The load capacity of each grip reaches 1 kg. The robot is made with an anthropomorphic structure that provides kinematic characteristics that are approximate and inherent of a person.

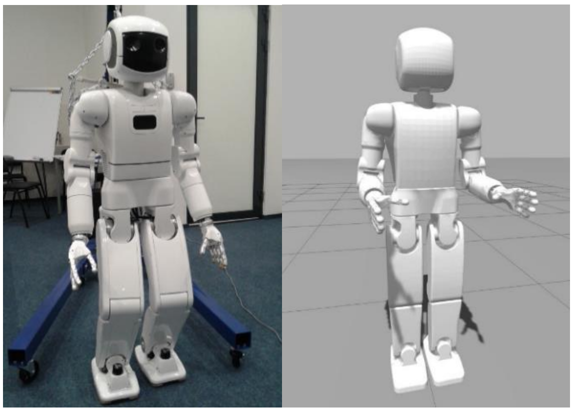


Fig.1 Real AR601M(left)[17] and its Gazebo model(right)

Model of AR601M in Gazebo should coincide with the real robot. We have two options for the robot model: model with Kinect is shown in Fig. 2 and without is shown in Fig. 1. Our model uses revolute joint controllers for the robot's body control and mimic joint controllers for the robot's fingers control. All controllers have the same limits, like on a real AR601M robot. With these controllers, the robot can walk, turn its head, move its arms, squeeze and unclench its fingers.

Additionally, for our pilot experiments, we modeled the following human actions:

1. Walking from a given point to a given point.
2. Movement of each hand along a given arbitrary trajectory (set of points).
3. Gripping a wrench.

4. Experiments

Using the results of the works presented in the literature review, we combined the fine and gross motor skills of the robot in one simulation model. In this section, we present the experiments with this model and the results of these experiments.

The first experiment shown in Fig. 2 simulates typical human actions in collaborative workspace such as walking, object take and place tasks. During the experiment AR601M robot needed to turn around, walk over to the table, pick up the wrench from a special holder, carry and place it to the conveyor used by the UR3 and UR5 robots.

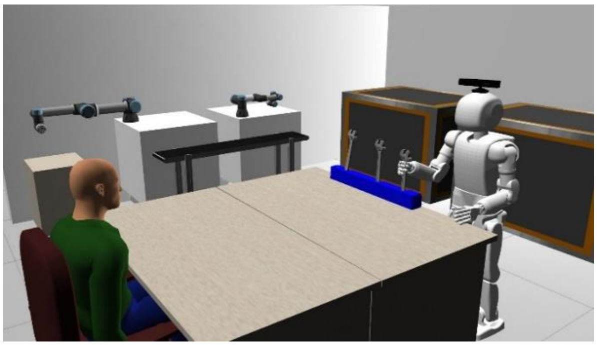


Fig.2 Frame of the first experiment, where the AR601M robot with Kinect grasped the wrench

The second experiment shown in Fig.3 presents a situation where the human and manipulator have co-working space - here is the table. Our AR601M robot stands at the table and simulates human actions close to the manipulator. Robot AR601M move apart small cube by hand, turns torso and moves its arms in front of UR5. The resulting model of AR601M in Gazebo coped with experiments. The robot performed human actions such as walking, turning to the side, hand movements and grasping objects in a collaborative workspace.

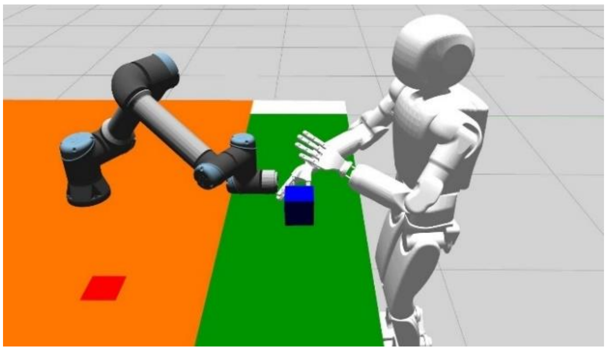


Fig.3 Frame of the second experiment, where the AR601M robot moves the work cube

5. Conclusion

Simulation modeling is an important stage in the design of the processes of collaborative interaction between a robot and a person. This research presented a model of AR601M robot that can be used in pilot experiments as a human in Gazebo Simulator. This model can interact with the environment unlike standard Gazebo “actors”, simulate the gross and fine motor skills of a human.

Acknowledgements

This paper has been supported by the Kazan Federal University Strategic Academic Leadership Program ("PRIORITY-2030").

References

1. Galin R, Meshcheryakov R. Review on human–robot interaction during collaboration in a shared workspace[C]//International Conference on Interactive Collaborative Robotics, 2019: 63-74.
2. Galin R, Meshcheryakov R, Kamesheva S. Distributing tasks in multi-agent robotic system for human-robot interaction applications[C]//Int. Conf. on Interactive Collaborative Robotics, 2020: 99-106.
3. Sultanov R, Sulaiman S, Li H, Meshcheryakov R, E Magid. A Review on Collaborative Robots in Industrial and Service Sectors[C]//Siberian Conference on Control and Communications, 2022 (in press).
4. Dobrokvashina A, Sulaiman S, Zagirov A, Chebotareva E, Hsia K-H, Magid E. Human Robot Interaction in Collaborative Manufacturing Scenarios: Prospective Cases[C]//Siberian Conference on Control and Communications, 2022 (in press).
5. He L, Glogowski P, Lemmerz K, Kuhlenkötter B, Zhang W. Method to integrate human simulation into gazebo for human-robot collaboration[C]//IOP

Conference Series: Materials Science and Engineering, 2020, 825(1): 012006.

6. Chebotareva E, Hsia K H, Yakovlev K, Magid E. Laser rangefinder and monocular camera data fusion for human-following algorithm by PMB-2 mobile robot in simulated Gazebo environment[C]//Proceedings of 15th International Conference on Electromechanics and Robotics "Zavalishin's Readings", 2021, 187: 357-369.
7. Kovalev A, Pavliuk N, Krestovnikov K, Saveliev A. Generation of walking patterns for biped robots based on dynamics of 3D linear inverted pendulum[C]//International Conference on Interactive Collaborative Robotics, 2019: 170-181.
8. Khusainov R, Shimchik I, Afanasyev I, Magid E. Toward a human-like locomotion: modelling dynamically stable locomotion of an anthropomorphic robot in simulink environment[C]//Proceedings of the 12th International Conference on Informatics in Control, Automation and Robotics (ICINCO), 2015, 2: 141-148.
9. Khusainov R, Shimchik I, Afanasyev I, Magid E. 3D modelling of biped robot locomotion with walking primitives approach in simulink environment[C]//Proceedings of the 12th International Conference on Informatics in Control, Automation and Robotics (ICINCO), 2016: 287-304.
10. Khusainov R, Afanasyev I, Sabirova L, Magid E. Bipedal robot locomotion modelling with virtual height inverted pendulum and preview control approaches in Simulink environment[J]. Journal of Robotics, Networking and Artificial Life, 2016, 3(3): 182-187.
11. Khusainov R, Afanasyev I, Magid E. Anthropomorphic robot modelling with virtual height inverted pendulum approach in Simulink: step length and robot height influence on walking stability[C]//Proceedings of International Conference on Artificial Life and Robotics, 2016.
12. Khusainov R, Klimchik A, Magid E. Kinematic and Dynamic Approaches in Gait Optimization for Humanoid Robot Locomotion[J]. Lecture Notes in Electrical Engineering, 2018, 430: 293-320.
13. Khusainov R, Klimchik A, Magid E. Swing leg trajectory optimization for a humanoid robot locomotion[C]//Proceedings of the 13th International Conference on Informatics in Control, Automation and Robotics (ICINCO), 2016, 2: 130-141.
14. Magid E, Sagitov A. Towards Robot Fall Detection and Management for Russian Humanoid AR-601[C]//KES International Symposium on Agent and Multi-Agent Systems: Technologies and Applications, 2017: 200-209.
15. Khusnutdinov K, Sagitov A, Yakupov A, Lavrenov R, Martinez-Garcia E A, Hsia K H, Magid E. Household objects pick and place task for AR-601M humanoid robot[C]//International Conference on Interactive Collaborative Robotics, 2019: 139-149.
16. Khusnutdinov K, Sagitov A, Yakupov A, Meshcheryakov R, Hsia K H, Martinez-Garcia E A, Magid E. Development and Implementation of Grasp Algorithm for Humanoid Robot AR-601M[C]//Proceedings of the 16th International Conference on Informatics in Control, Automation and Robotics (ICINCO), 2019, 2: 379-386.
17. Danilov I, Gabbasov B, Afanasyev I, Magid E. Zmp trajectory from human body locomotion dynamics evaluated by kinect-based motion capture system[C]//Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISAPP), 2016, 3:160-166.

Authors Introduction

Mr. Aidar Zagirov

Aidar Zagirov is a fourth-year student of the "Software Engineering" Bachelor degree program at the Institute of Information Technology and Intelligent Systems of the Kazan Federal University. Currently he works as an research assistant at the Laboratory of Intelligent Robotic Systems (LIRS) at Kazan Federal University, Russia

Mr. Artem Apurin

Artem Apurin received his B.Sc. degree at North-Caucasus Federal University in 2021. Now he is a 2-year student of the Master's program in Intelligent Robotics at Kazan Federal University. Currently he works as an engineer at the Laboratory of Intelligent Robotic Systems (LIRS) at Kazan Federal University, Russia

Ms. Elvira Chebotareva

She received her PhD in physics and mathematics from Kazan Federal University. She is currently an assitant professor in Laboratory of Intelligent Robotic Systems (LIRS) at Kazan Federal University, Russia.

Modern Methods of Map Construction Using Optical Sensors Fusion

Ramil Safin, Tatyana Tsoy, Roman Lavrenov, Ilya Afanasyev
Department of Intelligent Robotics, Kazan Federal University, Kazan, Russia

Evgeni Magid
Department of Intelligent Robotics, Kazan Federal University, Kazan, Russia
Higher School of Economics University, Moscow, Russia
E-mail: safin.ramil@it.kfu.ru

Abstract

Map construction, or mapping, plays an important role in robotic applications. Mapping relies on inherently noisy sensor measurements to construct an accurate representation of a surrounding environment. Generally, individual sensors suffer from performance degradation issues under certain conditions in the environment. Sensor fusion allows to obtain statistically more accurate perception and to cope with performance degradation issues by combining data from multiple sensors of different modalities. This article reviews modern sensor fusion methods for map construction applications based on optical sensors, such as cameras and laser range finders. State-of-the-art mapping solutions built upon different mathematical theories and concepts, such as machine learning, are considered.

Keywords: Sensor Fusion, Mapping, SLAM, Machine Learning, Camera, LiDAR

1. Introduction

Mapping is the process of constructing a map of an environment using robot perception. There exist multiple map representations, such as sparse point clouds, topological maps, and dense voxel grids. Mapping could be difficult due to adverse conditions (e.g., low lightning) and presence of dynamic objects.

Each type of sensor has its own limitations. For example, cameras underperform in low-light environments, while laser scanners cannot provide high-resolution data. To obtain reliable maps, it is required to combine strengths of each sensor to cope with their weaknesses. Sensor fusion, or data fusion, is a technique for combining data from multiple sensors in a way that allows to obtain more reliable and accurate information about the system being measured (Fig. 1). Data fusion is used in many robotics and machine vision applications, such as autonomous navigation and localization of mobile robots[1].

This article reviews modern sensor fusion methods for map construction applications based on optical sensors, such as cameras and laser scanners. State-of-the-art mapping solutions built upon different mathematical theories and concepts, such as machine learning, are considered.

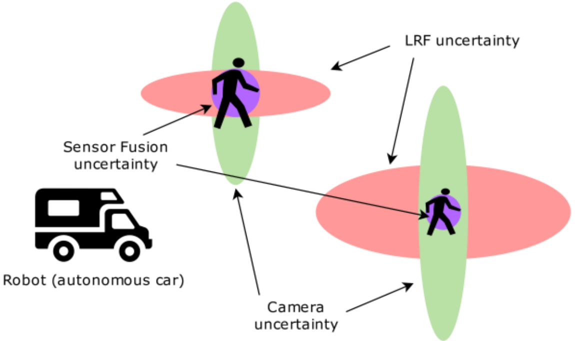


Fig. 1. An illustration of a laser range finder and camera sensor fusion. Uncertainty of the state is reduced due to multiple sources of information.

2. Sensor Fusion

High-level architecture of the sensor fusion system consists of the following components:

1. ***sensors*** that independently measure an observed quantity;
2. mathematical ***models*** that convert the observed quantity into a target value (e.g., calculating a position from distance measurements);
3. an ***inference algorithm*** that calculates the resulting target value by combining data from all sensors.

The inference algorithm combines all available measurements using provided sensor models and calculates an optimal state of a system with respect to specified criteria. The optimality of the solution and the criteria are determined by a cost function (e.g., an absolute error or a mean squared error).

The problem of data fusion can be viewed as combining measurements from multiple sensors to obtain a more optimal assessment (from statistical point of view) of the system's state. There are at least two reasons for data integration: data redundancy and completeness (complementarity). There are three main approaches for data fusion from sensors of various modalities[2]: a high-level (takes place on a decision-making level[3]), a mid-level (uses features of an environment[4]), and a low-level (operates with raw sensor data[5]). Methods can be also categorized into traditional and machine learning approaches[6].

Traditional approaches, such as statistical and probabilistic methods, create a model of a system being measured using various probability distributions. This enables to update an already initialized models when integrating new data. Probabilistic methods are built on Bayesian recursive rules, which enables to calculate and update a probability distribution function of an estimated state given sensor data.

One of the basic Bayesian integration methods is the Probabilistic Occupancy Grid (POG)[7]. The occupancy grid map divides an environment into cells of equal size. Each cell carries information about a probability of its occupancy. The Bayesian rule allows to update the initially constructed map. Computational performance is highly dependent on a map resolution and on the number of degrees of freedom. In practice, 2D maps are often considered due to a lower complexity.

Kalman filter is an analytical solution to the Bayesian approach, provided that a dynamical system is linear, and a state follows a Gaussian distribution. Dynamic systems with non-linear dependencies can be estimated using the Extended Kalman filter (EKF) – the state is linearized at each moment of time by expanding it into Taylor series and calculating the Jacobian matrix. In Unscented Kalman Filter (UKF) the problem is solved by linearization in a certain neighborhood – a group of weighted points (sigma points) is taken in a certain vicinity of a current state. Multi-state constrained Kalman Filter (MSCKF) is adapted to combine IMU data, extracted angles from a laser scanner, and features from a camera[8].

However, Kalman filters (the entire family) are unimodal approaches. Particle filters go further and suggest to model multimodal probability distributions of the state. Methods based on Monte Carlo describe the probability density function in terms of weighted elements of a space of linear and non-linear states. Particle filters are a subclass of the Monte Carlo method that allow calculating an optimal estimate of the state of a nonlinear state space with a Gaussian distribution. Compared to Kalman filters, a complexity of the algorithm grows much faster with a dimension of the state space (an exponential growth).

On the other hand, machine learning methods build a multi-layer network that allows extracting special features from a data stream. One of the most used architectures are convolutional (CNN)[9] and recurrent neural networks (RNN). For example, in[10] the authors propose a CNN-based approach, which combines data from a camera and a laser scanner to improve the object detection algorithm performance.

The problem with the machine learning approach is the quality of data on which a model is trained. The data may not contain all possible real-world scenarios, it may be biased and contain noise. Also, it is noted that machine learning methods are highly susceptible to malicious attacks[11].

Simultaneous localization and camera-based mapping algorithms often use the Bundle Adjustment (BA) approach and its modifications, such as windowed BA[12]. It simultaneously optimizes a pose and a map, can include several data sources by modifying a cost function. Combining visual data with laser scanner data allows obtaining data on a depth of a scene, being reliable under bad weather conditions[13].

Earlier simultaneous localization and mapping (SLAM) algorithms use only one sensor, mostly a camera [14][15] or laser scanner[16]. Currently, SLAM algorithms go through a stage where data is combined from several sensors of different modality. The most popular options are multiple camera configurations and laser scanners.

Deep learning methods are being used in dynamic SLAM problems. In[17], it is proposed to segment dynamic objects in a frame for their further masking and separation from a static scene. This enables to construct a map of an environment without modelling dynamic objects. CNNs are being used to estimate a depth of a scene. The estimated depth can be combined with data from a laser scanner to refine a map of an environment. For example, in[18], a scene depth scene was estimated based on an integration of data from a monocular camera and a laser rangefinder based on CNNs.

In[19], the authors proposed a monocular SLAM algorithm that used a one-dimensional LRF in conjunction with a camera to refine the distance to singular points in the image.

The paper[20] proposed an algorithm for constructing a dense environment map based on combining data from a monocular camera and predicting a scene depth using CNNs. In some works, image features and scanner point clouds are matched using scan matching methods[21], which allows to obtain dense and accurate 3D-maps of an environment. In[22] the author built a voxel environment map to effectively refine poses of key points using data from a laser scanner.

There are mapping algorithms that are based on the idea of associating point clouds from a laser and a camera. For example, in[23], a positioning system was proposed based on combining data from a camera and a laser scanner. The system searches for correlations between point clouds and images. In order to obtain accurate depths of point clouds methods such as segmentation of image areas (planar surfaces, foreground, etc.) are used.

In[24] a method for constructing dense 3D occupancy grid maps was based on sparse data from a laser scanner. To model dynamic objects in the environment, it was proposed to use dynamic occupancy mapping[25]. Recently, there are works of dense surfel mapping systems that can operate on CPU[26].

3. Conclusion

In conclusion, most of the reviewed sensor fusion mapping techniques are based on traditional methods, such as Extended Kalman Filtering and particle filters. In case of SLAM systems, where mapping and localization are done simultaneously, bundle adjustment optimization methods prevail. Most of the time, laser scanners are used to improve the camera depth estimation capabilities. Machine learning solutions enable to learn depth form images but lack accuracy and reliability.

Acknowledgements

The reported study was funded by the Russian Science Foundation (RSF) and the Cabinet of Ministers of the Republic of Tatarstan according to the research project No. 22-21-20033.

References

1. Safin R, Lavrenov R, Martinez-Garcia E A Evaluation of Visual SLAM Methods in USAR Applications Using ROS/Gazebo Simulation[C]//Int. Conf. on Electromechanics and Robotics" Zavalishin's Readings". Springer, 2021: 371-382.
2. Banerjee K et al. Online camera lidar fusion and object detection on hybrid data for autonomous driving[C]//IEEE Int. Vehicles Symposium (IV). IEEE, 2018: 1632–1638.
3. Shahian J B, Tulabandhula T, Cetin S. Real-time hybrid multi-sensor fusion framework for perception in autonomous vehicles[J]. Sensors. MDPI, 2019, 19(20): 4357.
4. Li Y et al. Feature level sensor fusion for target detection in dynamic environments[C]//American Control Conf. (ACC). IEEE, 2015: 2433–2438.
5. Yoo J H et al. 3d-cvf: Generating joint camera and lidar features using cross-view spatial feature fusion for 3d object detection[C]//European Conf. on Computer Vision. Springer, 2020: 720–736.
6. Fayyad J et al. Deep learning sensor fusion for autonomous vehicle perception and localization: A review[J]. Sensors. MDPI, 2020, 20(15): 4220.
7. Yue Y et al. Hierarchical probabilistic fusion framework for matching and merging of 3-d occupancy maps[J]. IEEE Sensors Journal. IEEE, 2018, 18(21): 8933–8949.
8. Xu X et al. A Review of Multi-Sensor Fusion SLAM Systems Based on 3D LIDAR[J]. Remote Sensing. MDPI, 2022, 14(12): 2835.
9. Buyval A, Gavrilenkov M, Magid E. A multithreaded algorithm of UAV visual localization based on a 3D model of environment: implementation with CUDA technology and CNN filtering of minor importance objects[C]//Int. Conf. on Artificial Life and Robotics, 2017: 356–359.
10. Kim J, Kim J, Cho J. An advanced object classification strategy using YOLO through camera and LiDAR sensor fusion[C]//Int. Conf. on Signal Proc. and Comm. Sys. (ICSPCS). IEEE, 2019: 1–5.
11. Pitropakis N et al. A taxonomy and survey of attacks against machine learning[J]. Computer Science Review. Elsevier, 2019, 34: 100199.
12. Zhou L, Kaess M. Windowed bundle adjustment framework for unsupervised learning of monocular depth estimation with u-net extension and clip loss[J]. IEEE Robotics and Automation Letters. IEEE, 2020, 5(2): 3283–3290.

13. Xiao L et al. Hybrid conditional random field based camera-LIDAR fusion for road detection[C]//Inf. Sciences. Elsevier, 2018, 432: 543–558.
14. Abbyasov B et al. Comparative analysis of ROS-based centralized methods for conducting collaborative monocular visual SLAM using a pair of UAVs[C]//Int. Conf. on Climbing and Walking Robots and Support Tech. for Mobile Machines (CLAWAR), 2020: 113–120.
15. Mingachev E et al. Comparison of ros-based monocular visual slam methods: DSO, LDSO, ORB-SLAM2 and DynaSLAM[C]//Int. Conf. on Inter. Collab. Robotics. Springer, 2020: 222–233.
16. Alishev N et al. Network failure detection and autonomous return algorithms for a crawler mobile robot navigation[C]//Int. Conf. on Development in eSystems Eng. (DeSE). IEEE, 2018: 169–174.
17. Xiao L et al. Dynamic-SLAM: Semantic monocular visual localization and mapping based on deep learning in dynamic environment[J]. Robotics and Autonomous Systems. Elsevier, 2019, 117: 1–16.
18. Gao H et al. Object classification using CNN-based fusion of vision and LIDAR in autonomous vehicle environment[J]. IEEE Trans. on Industrial Informatics IEEE, 2018, 14(9): 4224–4231.
19. Zhang Z et al. Scale estimation and correction of the monocular simultaneous localization and mapping (SLAM) based on fusion of 1D laser range finder and vision data[J]. Sensors. MDPI, 2018, 18(6): 1948.
20. Laidlow T, Czarnowski J, Leutenegger S. DeepFusion: Real-time dense 3D reconstruction for monocular SLAM using single-view depth and gradient predictions[C]//Int. Conf. on Robotics and Automation (ICRA). IEEE, 2019: 4068–4074.
21. Khattak S et al. Complementary multi–modal sensor fusion for resilient robot pose estimation in subterranean environments[C]//Int. Conf. on UAS (ICUAS). IEEE, 2020: 1024–1029.
22. Wang P et al. Vanishing Point Aided LiDAR-Visual-Inertial Estimator[C]//IEEE Int. Conf. on Robotics and Automation (ICRA). IEEE, 2021: 13120–13126.
23. Graeter J, Wilczynski A, Lauer M. Limo: Lidar-monocular visual odometry[C]//IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS). IEEE, 2018: 7872–7879.
24. Gan L et al. Bayesian spatial kernel smoothing for scalable dense semantic mapping[J]. IEEE Robotics and Automation Letters. IEEE, 2020, 5(2): 790–797.
25. Min Y, Kim D U, Choi H L. Kernel-based 3-d dynamic occupancy mapping with particle tracking[C]//IEEE Int. Conf. on Robotics and Automation (ICRA). IEEE: 2021: 5268-5274.
26. Wang K, Gao F, Shen S. Real-time scalable dense surfel mapping[C]//Int. Conf. on Robotics and Automation (ICRA). IEEE, 2019: 6919–6925.

Authors Introduction

Mr. Ramil Safin

He is a Research Assistant and Ph.D. student in Kazan Federal University. His research interests are computer vision, mobile robotics, and SLAM.

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.

Assistant Professor Roman Lavrenov

He received his PhD from Kazan Federal University in 2020. Currently, he works as an Assistant Professor at Kazan Federal University, Russia. His research interests are path planning, ROS/Gazebo, and mobile robotics.

Assistant Professor Ilya Afanasyev

He received his PhD from Federal State Unitary Enterprise All-Russian Center “Vavilov State Optical Institute" in 2006. Currently, he works as an Assistant Professor at Kazan Federal University, Russia. He authors over 150 publications.

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 authors over 200 publications.

Omniwheel Chassis' Model and Plugin for Gazebo Simulator

Artem Apurin, Bulat Abbyasov, Alexandra Dobrokvashina
Intelligent Robotics Department, Kazan Federal University, 420111 Kazan, Russian Federation

Yang Bai, Mikhail Svinin
Information Science and Engineering Department, Ritsumeikan University, 1-1-1 Noji-higashi,
Kusatsu, Shiga, 525-8577, Japan

Evgeni Magid
Intelligent Robotics Department, Kazan Federal University, 420111 Kazan, Russian Federation
E-mail: magid@it.kfu.ru
http://robot.kpfu.ru/eng

Abstract

Increasing a mobility of a robot in a limited space or in a presence of a large number of people is an important task. Mecanum wheels could provide the required high flexibility locomotion in any direction. This article presents a virtual model of an omniwheel robot in the Gazebo simulator. Freely rotating rollers were implemented to simulate the robot motion. We developed a four-wheel mecanum mobile robot plugin controls the robot by publishing linear velocity data along X and Y axes, and angular velocity data along Z-axis. The plugin could optionally publish a standard ground-truth odometry of the Gazebo or a calculated in real time wheel odometry. The open source code is extendable for similarly structured platforms and is available for a download via GitLab.

Keywords: Mecanum Wheels, Mobile Platform, Gazebo Simulator

1. Introduction

In the last decade, mobile robots with omnidirectional motion capabilities have been widely applied in various areas of a human life[1]. These mobile platforms are broadly employed in autonomous driving[2], manufacturing[3], industry[4], medical applications[5], social robotics[6], search and rescue[7], agriculture[8], education[9] and others. Omnidirectional movement is the ability to travel in any desired direction at any specified orientation[10]. To perform such locomotion an omnidirectional robot could use a special design of wheels called mecanum[11]. Mecanum wheels gain advantages in terms of maneuverability in narrow and crowded spaces[12]. A robot motion controller is decoupled and could perform translational and rotational motions.

Prior to a construction stage, it is important to model and simulate a new robot behavior in a virtual environment in order to decide on its efficiency and attempt to improve robot’s mechanical and software features. Typically, roboticists prefer using standard simulators (e.g., ROS/Gazebo[13], Webots[14] or CoppeliaSim/V-REP[15]) while using simulators with a high-level of abstraction or developing in-house simulators are already considered to be an out-of-date approaches.

This paper presents a virtual model development of a four-wheeled omniwheel mobile robot LIRS-ArtBul[16] in the Gazebo simulator. Each mecanum wheel model has freely rotating rollers in order to perform omnidirectional motion. A motion control scheme was wrapped into a special four-wheel mecanum ROS plugin, which is fully configurable and can be easily reconfigured for researchers’ tasks.

2. Modeling wheels and rollers

To achieve a high quality of a virtual model, it is important to take into account all the parameters and features of a corresponding real robot when creating its virtual representation for the simulator. The main task was to develop a model of an omniwheel robot with active freely rotating rollers. To create models Blender software was employed. A visual and collision parts of the wheel roller model are shown in Fig.1.

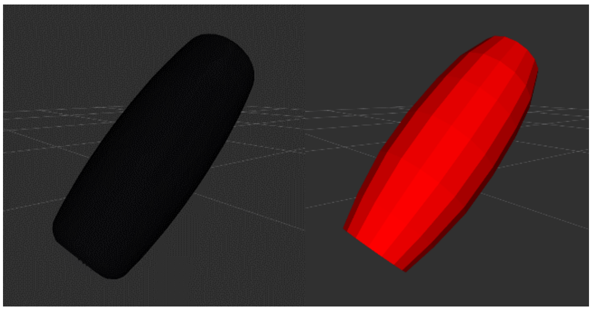


Fig.1.Wheel roller model in RViz: visual appearance (left) and collision (right) parts

To reduce complex collision calculations and increase real time factor (RTF), a low-poly model should be used. Roller's collision mesh with 220 polygons and a roller model for visual part with 7708 polygons allowed increasing RTF to 0.95-0.97. The wheel model in Gazebo is shown in Fig.2, each roller has its own joint and can be freely rotated along the Z-axis.

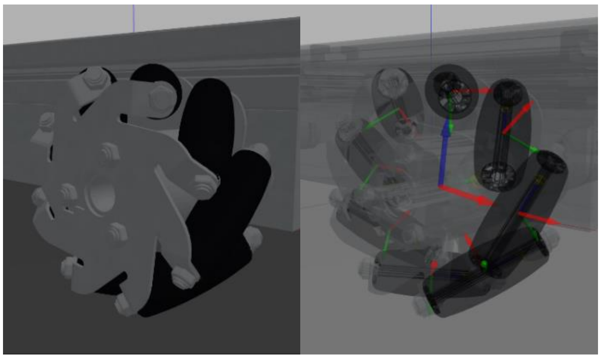


Fig.2.Wheel model in Gazebo

Four controllers of type *joint_velocity_controller* from *effort_controller* plugin were configured, since the real robot has a motor for each wheel. Package *joint_state_publisher* can be used to publish a position and a transformation of moving joints. *joint_state_publisher* connects to all non-fixed joints, but in this case *controller_manager* and *joint_state_publisher* packages publish positions of the duplicated joints to */joint_states* topic and multi-source publisher error occurs. Therefore, rollers' transmission parameters should be configured in URDF file, but controllers should not be created for these joints in YAML file. Then *controller_manager* publishes a position and a transformation of the roller joints and those joints could rotate freely. A mobile chassis with mecanum wheels in Gazebo is shown in Fig.3

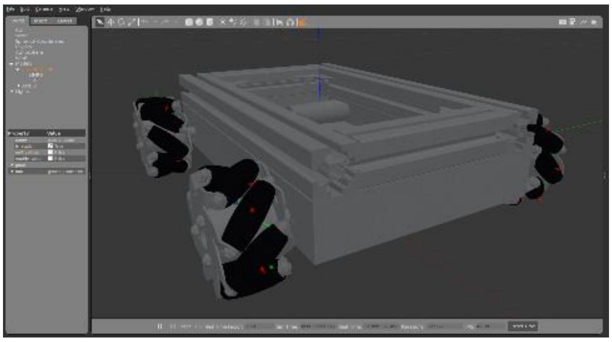


Fig.3.Mobile platform in Gazebo

3. Gazebo plugin for omniwheel robots

A plugin was developed to control the robot in Gazebo. It is suitable for all four-wheel mecanum mobile robots. The robot is controlled by publishing messages with linear velocities along the X and Y axes and an angular velocity along the Z axis to */cmd_vel* topic. From obtained data, the wheel velocities are calculated according to the equations (1) – (4)[17][18]:

$$v_{LF} = \frac{1}{R_{wheel}} \cdot \left(v_x - v_y - \left(\frac{s_{RL}}{2} + \frac{s_{FR}}{2}\right) \cdot v_z\right) \tag{1}$$

$$v_{RF} = \frac{1}{R_{wheel}} \cdot \left(v_x + v_y + \left(\frac{s_{RL}}{2} + \frac{s_{FR}}{2}\right) \cdot v_z\right) \tag{2}$$

$$v_{LR} = \frac{1}{R_{wheel}} \cdot \left(v_x + v_y - \left(\frac{s_{RL}}{2} + \frac{s_{FR}}{2}\right) \cdot v_z\right) \tag{3}$$

$$v_{RR} = \frac{1}{R_{wheel}} \cdot \left(v_x - v_y + \left(\frac{s_{RL}}{2} + \frac{s_{FR}}{2}\right) \cdot v_z\right) \tag{4}$$

where $v_{LF}, v_{RF}, v_{LR}, v_{RR}$ – target rotation velocities of the left front, the right front, the left rear and the right rear wheels respectively in rad/s; R_{wheel} – the wheel radius; v_x is a target linear velocity along X-axis in m/s; v_y is a target linear velocity along Y-axis in m/s; v_z is a target angular velocity along Z-axis in rad/s; s_{RL} is the distance between the right and left wheels; s_{FR} is the distance between the front and rear wheels. Then the target wheel velocity is published to */command* topic of each controller. The plugin allows to select one of two types of published odometry:

1. Ground-truth odometry from Gazebo,
2. Wheel odometry.

In the first case, the position of the robot in the Gazebo is published. In the second case, the plugin searches for the necessary joints in the */joint_states* topic and their current rotation velocity, then calculates the robot velocity using the equations (5) – (7)[19]:

$$v_{xc} = (v_{LFC} + v_{RFC} + v_{LRC} + v_{RRC}) \cdot \frac{R_{wheel}}{4} \qquad (5)$$

$$v_{yc} = (-v_{LFC} + v_{RFC} + v_{LRC} - v_{RRC}) \cdot \frac{R_{wheel}}{4} \qquad (6)$$

$$v_{zc} = (-v_{LFC} + v_{RFC} - v_{LRC} + v_{RRC}) \cdot \frac{R_{wheel}}{2(s_{RL} + s_{FR})} \qquad (7)$$

where v_{xc} is a current linear velocity along X-axis in m/s; v_{yc} is a current linear velocity along Y-axis in m/s; v_{zc} is an current angular velocity along Z-axis in rad/s; v_{LFC}, v_{RFC}, v_{LRC}, v_{RRC} are current rotation velocities of the left front, the right front, the left rear and the right rear wheels respectively; R_{wheel} is the wheel radius; s_{RL} is the distance between the right and left wheels; s_{FR} is the distance between front and rear wheels. Then the current robot position is calculated and the obtained data are published. This allows to choose between the accurate and the realistic odometry. Flowchart of received / sent messages by the plugin is shown in Fig.4

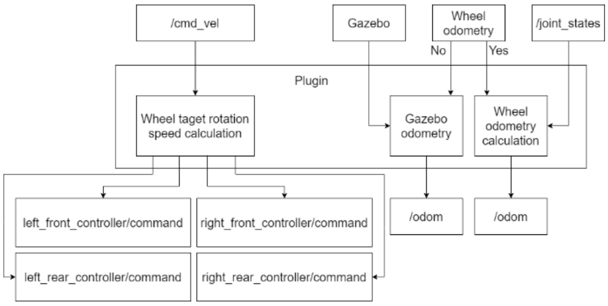

/cmd_vel
Gazebo
Wheel odometry
/joint_states
No
Yes
Plugin
Wheel taget rotation speed calculation
Gazebo odometry
Wheel odometry calculation
left_front_controller/command
right_front_controller/command
/odom
/odom
left_rear_controller/command
right_rear_controller/command

Fig.4.Flowchart of received/sent messages by the plugin

The plugin needs to set a command topic and an odometry topic, an odometry frame, the robot base frame, an update rate, to select the type of odometry, to specify names of the wheel controllers, half the length between the front and rear, right and left wheels and the wheel radius.

4. Conclusion

The article presented the virtual mecanum wheel mobile platform development. A high-poly and detailed model was developed for a visual part of the wheel, and a low-poly model was modeled for the collision part. This allowed to significantly increase the RTF of the Gazebo simulation. The developed plugin controls the model and is suitable for all four wheeled robots with mecanum wheels. The plugin publishes target speeds to wheel controller topics and can publish two types of odometry.

Acknowledgements

This paper has been supported by the Kazan Federal University Strategic Academic Leadership Program ("PRIORITY-2030").

References

1. Rubio F, Valero F, Llopis-Albert C. A review of mobile robots: Concepts, methods, theoretical framework, and applications[J]. Int. J. of Advanced Robotic Systems, 2019.
2. Watanabe K. et al. Autonomous Control for an Omnidirectional Mobile Robot with Feedback Control System[J]. Advances in Intelligent Autonomous Systems, 1999: 289-308.
3. Qian J. et al. The design and development of an omni-directional mobile robot oriented to an intelligent manufacturing system[J]. Sensors, 2017: 2073.
4. Paromtchik I E, Rembold U. A practical approach to motion generation and control for an omnidirectional mobile robot[C]//Int. Conf. on Robotics and Automation, 1994: 2790-2795.
5. Magid E. et al. Automating pandemic mitigation[J]. Advanced Robotics, 2021: 572-589.
6. Senft E, Satake S, Kanda T. Would You Mind Me if I Pass by You? : Socially-Appropriate Behaviour for an Omni-based Social Robot in Narrow Environment[C]//15th ACM/IEEE International Conference on Human-Robot Interaction (HRI). IEEE, 2020: 539-547.
7. Peng H. et al. Stability control for medical rescue robot in unconstructed environment[C]//5th International Conference on Advanced Robotics and Mechatronics (ICARM). IEEE, 2020: 184-188.
8. Iakovlev R, Saveliev A. Approach to implementation of local navigation of mobile robotic systems in agriculture with the aid of radio modules[J]. Telfor Journal, 2020: 92-97.
9. Wu X B. et al. Research on the design of educational robot with four-wheel omni-direction chassis[J]. J. Comput, 2018: 284-294.

10. Watanabe K. Control of an omnidirectional mobile robot[C]//Second Int. Conf. Knowledge-Based Intelligent Electronic Systems, 1998: 51-60.
11. Gfrerrer A. Geometry and kinematics of the Mecanum wheel[J]. Computer Aided Geometric Design, 2008: 784-791.
12. Shabalina K, Sagitov A, Magid E. Comparative analysis of mobile robot wheels design[C]//11th International Conference on Developments in esystems Engineering (dese). IEEE, 2018: 175-179.
13. Shabalina K. et al. Avrora unior car-like robot in gazebo environment[C]//International Conference on Artificial Life and Robotics, 2019: 116-119.
14. Dobrokvashina A. et al. Servosila engineer crawler robot modelling in webots simulator[J]. International Journal of Mechanical Engineering and Robotics Research (IJMERR), 2022: 417-421.
15. Zamfirescu I., Pascal C. Modelling and simulation of an omnidirectional mobile platform with robotic arm in CoppeliaSim[C]//24th International Conference on System Theory, Control and Computing (ICSTCC). IEEE, 2020: 667-672.
16. Apurin A. et al. LIRS-ArtBul: Design, Modelling and Construction of an Omnidirectional Chassis for a Modular Multipurpose Robotic Platform[C]//International Conference on Interactive Collaborative Robotics. Springer, Cham, 2022: 70-80.
17. Taheri H, Qiao B, Ghaeminezhad N. Kinematic model of a four mecanum wheeled mobile robot[J]. Int. J. of computer applications, 2015: 6-9.
18. Dickerson S L, Lapin B D. Control of an omni-directional robotic vehicle with Mecanum wheels[C]//National Telesystems Conf. Proc., 1991: 323-328.
19. Maulana E, Muslim M A, Hendrayawan V. Inverse kinematic implementation of four-wheels mecanum drive mobile robot using stepper motors[C]//Int. Seminar on Intelligent Technology and Its Applications, 2015: 51-56.

Authors Introduction

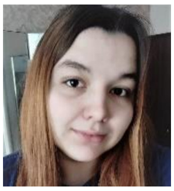
Mr. Artem Apurin

Artem Apurin received his B.Sc. degree at North-Caucasus Federal University in 2021. Now he is a 2-year student of the Master's program in Intelligent Robotics at Kazan Federal University. Currently he works as an engineer at the Laboratory of Intelligent Robotic Systems (LIRS) at Kazan Federal University, Russia

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 first year PhD student in Robotics at ITIS, Kazan Federal University.

Ms. Alexandra Dobrokvashina



She received her 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 Yang Bai

Currently he works as an Assistant Professor at the College of Information Science and Engineering of the Ritsumeikan University. His interests are robotics, non-holonomic systems and motion planning.

Professor Mikhail Svinin

He works as a Professor at the College of Information Science and Engineering of the Ritsumeikan University. His current research interests include robotics, control theory, and analytical mechanics. He is a member of American Society of Mechanical Engineers (ASME), Robotics Society of Japan (RSJ), and the Society of Instrument and Control Engineers (SICE).

Professor Evgeni Magid

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. Previously he worked at University of Bristol; Carnegie Mellon University; University of Tsukuba, Japan; National Institute of Advanced Industrial Science and Technology, Japan. He earned his Ph.D. degree from University of Tsukuba. He authors over 200 publications

Using optical sensors for industrial robot-human interactions in a Gazebo environment.

Elaman Kidiraliev, Roman Lavrenov
Intelligent Robotics Department, Kazan Federal University, 420008, Kazan, Russian Federation
E-mail: EAKidiraliev@stud.kpfu.ru, lavrenov@it.kpfu.ru

Abstract

This paper presents an overview of a robot operating system based architecture for human-industrial robot interactions using peripheral optical sensors for real-time object detection and collision avoidance with an industrial robot in the virtual world of the Gazebo simulator. Machine vision plays a huge role in production automation, and develop a system based on Kuka KR3 industrial robot for detecting and tracking a human and other objects in a working area using optical sensors. The ability to work in low light and crowded conditions, as well as the ability to reconstruct a method of execution of a task, while maintaining control of a robot. This work considers several optical sensors and a comparative analysis.

Keywords: Gazebo, safety, Kuka, HRI, Computer Vision

1. Introduction

Nowadays, while industrial manipulators are effectively used in most modern factories together with human workers[1], many of them are not equipped with touch and collision sensors. Unintended human presence in the work area of a manipulator and incorrect calculations eventually lead to a collision[2]. As a result, standards and regulations are violated.

There are many variants of interaction and collision with robots[3][4][5]. In[6][7] the authors developed an algorithm to classify hazards during human-robot interaction. Cameras were used to ensure safety and calculate distances between a person and a robot. Human interactions with an industrial robot were performed using gestures and commands as follows[8]: the camera detects the hand and recognizes the command[9]. The described method improves the level of human-robot interaction without using a remote control. The authors of the article[10] proposed several methods for face and hand recognition. In the article[11] reviewed the structure of the code, types of scanners and cameras, proposed the calculation and method of human detection.

This article presents an optical sensor which is used for human-industrial robot interaction in a Gazebo environment. In experiment and simulations were used off-the-shelf object recognition methods.

2. System setup

Experiments were conducted using Robot Operating System (ROS), Gazebo, RViZ, MoveIt on Ubuntu Linux 20.04 operating system. Ubuntu is currently considered as a popular distribution based on Debian architecture. ROS is a system used to create robotic applications where the main goal is to provide a powerful robotic application that can be used in other robots. This system has a set of software tools, packages and libraries that make software development for robots easier[12]. With Gazebo we can simulate not only the robot but also simulate a room, sensors and various 3D objects[13]. MoveIt is a set of tools for manipulation tasks, such as picking, moving, placing or simple motion planning using inverse kinematics with the help of ready-made libraries that include a fast inverse kinematics solver. And with RVIZ we can visualize in real time all components of the robot system such as coordinate systems, moving parts, sensor readings, and camera images.

3. Research methods

Research and experiments involving real robots and humans are not the cheapest and safest methods. Mistakes and failures are possible during experiments. As a result, humans can harm themselves or industrial robots can fail due to breakage and damage. Also, during failed experiments, recovery is very expensive and long, which is why using virtual robot and human models in a Gazebo environment is more advantageous[14].

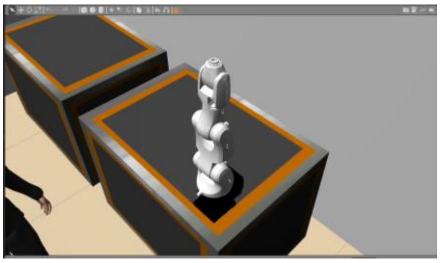


Fig. 1. Kuka KR3 R540 in the Gazebo environment

3.1. Robot control parameters

The described simulations are performed on Gazebo using virtual models. In Gazebo environment we can use 3D model of robots and other objects. Also, due to the capabilities of Gazebo environment, we can create a physical verisimilitude, which is not different from experiments in the real world. During testing, we can use different kinds of models of robots and other objects, and in our experiment, we used a virtual model of Kuka KR3 R540 industrial robot[15] Fig. 1. To control the industrial robot in the Gazebo environment, we use the gazebo ros control plugin with the additional advanced architecture needed to create and provide effective user hardware interfaces between Gazebo and ROS Control.

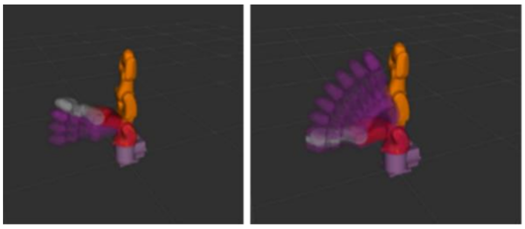


Fig. 2. Example of observation of a manipulator's trajectory

A robot controller is used to make this manipulator model work in reality and perform tasks correctly, which is an important part of this stage, also used as trajectory controller Fig. 2 (a trajectory controller is used for position tracking). To ensure safety and prevent collisions with objects, was developed a system based on the *actionlib* message package, where the communication between the industrial robot and the stereo camera is established. Using such a system provides more accurate calculation of distances and also giving the robot an understanding of its location relative to other objects.

As a result, we know the positions of waypoints and speeds, which gives us a message stream consist of all the information about the final state of the robot. There are also functions needed to convert the desired joint position to a given target. The desired joint position can be obtained from the inverse kinematics and then converted to the ROS goal msg form.

3.2. Optical sensors

A virtual model of the RGB-D-Microsoft Kinect optical sensor is used for experiments. The sensor is an infrared detector (infrared projector with a monochrome CMOS-sensor), which receives an image in three dimensions regardless of lighting conditions, and is also equipped with an RGB camera. The presented virtual model of the optical sensor has the same characteristics as in reality. We get an image from the virtual world of 640 x 480 and depth information, which can be obtained at a maximum frequency of 30 Hz Fig. 3, Fig. 4. The optical sensor is mounted on top of the industrial robot and pointed directly at it.

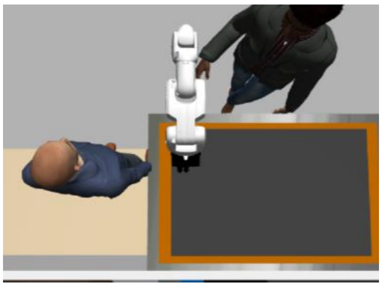


Fig. 3. Image coming from a stereo camera

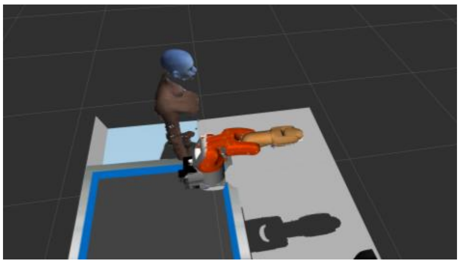


Fig. 4. Point cloud from RVIZ

3.3. Object recognition methods

A system based on the YOLO v3 and MEDIAPIPE algorithms was created to recognize objects from image streams coming from the optical sensor. Used algorithms, unlike others, provide fast and more accurate object detection. The developed Python script integrates the object detection algorithm based on YOLO v3 Fig. 5 and Mediapipe into ROS via a node. The node subscribes to the */depth_camera/image_raw* video stream thread, which receives data from an optical sensor mounted above the industrial robot. The received video stream frames are processed and the downstream module performs industrial robot and human detection.

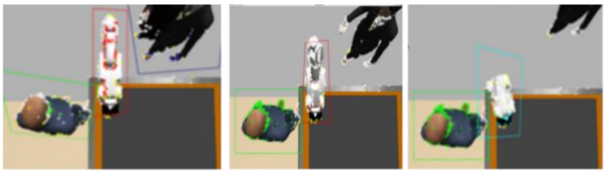


Fig. 5. Detecting objects by stereo camera

3.4. Collision detection methods

After image processing and the detection of objects coming from the image stream, these objects are highlighted and the distances between the object and the industrial robot are calculated. In experiment describes the calculation of the distance between objects[16]. Based on similar works, we developed an algorithm calculating the distance between an industrial robot and objects in the working area of the industrial robot Fig. 6.

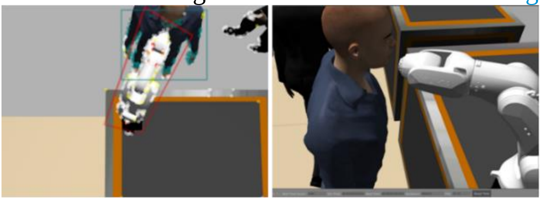


Fig. 6. Detecting a collision between an industrial manipulator and a person

To avoid collisions with objects while controlling an industrial robot, we use two methods. Both methods include autonomous trajectory planning and real-time obstacle detection. The first method is global collision-free trajectory planning of an industrial robot in case the object is static or the route of the moving object in time is known. In the second method, the route of the object is unknown and it is necessary to measure the distance between the object and the industrial robot using a real-time vision system.

3.5. Risk assessment of human-robot interaction

The flowchart shows an algorithm that provides safe interaction without touching an object during a task. The algorithm looks as follows. The industrial robot starts to perform the task if there are no obstacles in the working area Fig. 7. In case that object is detected in the working area, the speed is set to minimum values. Next calculating the distance between the robot and the object function started working. If the distance decreases, the task process is temporarily stopped to find a new route, waiting for the release of the optimal route.

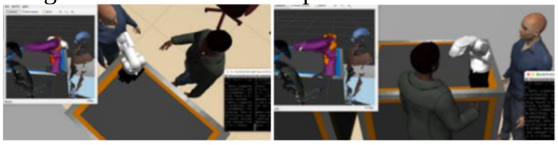


Fig. 7. Calculating the trajectory of the manipulator

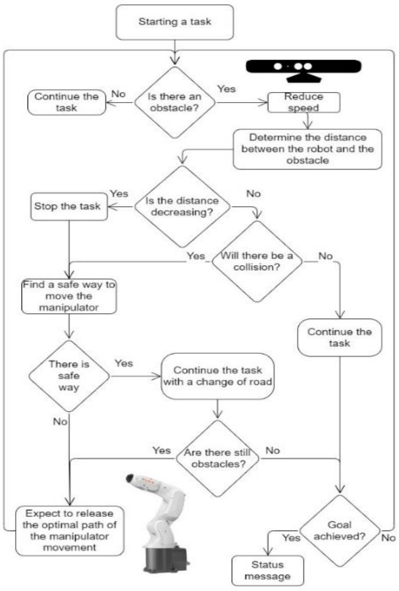

Starting a task
Continue the task
No
Is there an obstacle?
Yes
Reduce speed
Determine the distance between the robot and the obstacle
Stop the task
Yes
Is the distance decreasing?
No
Yes
Will there be a collision?
No
Find a safe way to move the manipulator
Continue the task
There is safe way
Yes
Continue the task with a change of road
No
Yes
Are there still obstacles?
No
Expect to release the optimal path of the manipulator movement
Yes
Goal achieved?
No
Status message

Fig. 8. Block diagram of the collision avoidance method

If the distance does not decrease and does not interfere with the movement of the industrial robot, the execution of the task continues with the set minimum values. In the process of task execution, a report will be issued indicating the status of task execution Fig. 8.

3.6. Results

We conducted a series of experiments simulating human intervention in the process of an industrial robot. In the working area of the robot different positions and animated actions of a virtual human models from an open-source library for the Gazebo[17] were used. This allowed us to test the algorithms of human-industrial robot interaction for safety. All experiments were conducted in the virtual world of Gazebo.

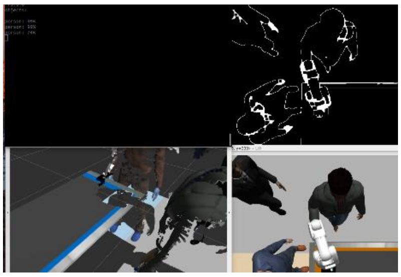


Fig. 9. Human detection using an image stream

Was developed a system which ensures that a human is located safely next to the industrial robot at the moment of performing its tasks Fig. 9.

4. Conclusion

The developed system is quite successful in its tasks, still it needs improvements, because the frequency of image updates is not fast enough to ensure a high efficiency.

Acknowledgements

This paper has been supported by the Kazan Federal University Strategic Academic Leadership Program ("PRIORITY-2030").

References

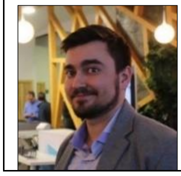
1. Galin R R, Shiroky A A, Magid E A, Mescheriakov E, Mamchenko M V. Effective functioning of a mixed heterogeneous team in a collaborative robotic system[J]. Informatics and Automation, 2021: 1224-1253.
2. Dobrokvashina A, Sulaiman S, Zagirov A, Chebotareva A, Hsia K, Magid E. Human Robot Interaction in Collaborative Manufacturing Scenarios: Prospective Cases[C]//Siberian Conference on Control and Communications, 2022,.
3. Secil S, Ozkan M. Minimum distance calculation using skeletal tracking for safe human-robot interaction[J]. Robotics and Computer-Integrated Manufacturing, 2022: 102-253.
4. Tashtoush T. et al. Human-robot interaction and collaboration (HRI-c) utilizing top-view RGB-d camera system[J]. International Journal of Advanced Computer Science and Applications, 2021: 11-17.
5. Hentout A. et al. Human–robot interaction in industrial collaborative robotics: a literature review of the decade 2008–2017[J]. Advanced Robotics, 2019: 764-799.
6. Vasconez J P, Salvo J. Toward semantic action recognition for avocado harvesting process based on single shot multibox detector[C]//International Conference on Automation/XXIII Congress of the Chilean Association of Automatic Control (ICA-ACCA). IEEE, 2018: 1-6
7. Bdiwi M, Marko P, Sterzing A. A new strategy for ensuring human safety during various levels of interaction with industrial robots[C]//CIRP Annals, 2017 60: 453-456.
8. Mazhar O. et al. A real-time human-robot interaction framework with robust background invariant hand gesture detection[J]. Robotics and Computer-Integrated Manufacturing, 2019: 34-48.
9. Ko D, Lee S, Park J. A study on manufacturing facility safety system using multimedia tools for cyber physical systems[J]. Multimedia Tools and Applications, 2021 80(26): 34553-34570.
10. Plagemann C. et al. Real-time identification and localization of body parts from depth images[C]//Int. Conf. on Robotics and Automation, 2010: 3108-3113
11. Pellegrinelli S. et al. A probabilistic approach to workspace sharing for human–robot cooperation in assembly tasks[C]//CIRP Annals. 2016 65(1): 57-60.
12. Safin R, Lavrenov R, Martínez-García E A. Evaluation of visual slam methods in usar applications using ros/gazebo simulation[C]//Proceedings of 15th International Conference on Electromechanics and Robotics" Zavalishin's Readings". Springer, Singapore, 2021: 371-382.
13. Abbyasov B, Kononov K, Tsoy T, Martinez-Garcia E A, Magid E. Experience in Efficient Real Office Environment Modelling in Gazebo: a Tutorial [C]//International Conference on Artificial Life and Robotics, 2022: 673-677.
14. Gavrilova L, Kotik A, Tsoy T, Martinez-Garcia E A, Svinin M, Magid E. Facilitating a preparatory stage of real-world experiments in a humanoid robot assisted English language teaching using Gazebo simulator[C]//Proceedings of 13th International Conference on Developments in eSystems Engineering (DeSE), 2020: 222-227.
15. Pichkalev M, Lavrenov R, Safin R. Face Drawing by KUKA 6 Axis Robot Manipulator[C]//12th International Conference on Developments in eSystems Engineering (DeSE), 2019: 709-714.
16. Kaonain T. et al. A simulated risk assessment of human-robot interaction in the domestic environment[J]. IAES International Journal of Robotics and Automation, 2020 9(4): 300-310.
17. Tukhtamanov N, Chebotareva E, Tsoy T, Lavrenov R, Svinin M, Magid E. Open Source Library of Human Models for Gazebo Simulator[C]//Siberian Conference on Control and Communications, 2022.

Authors Introduction

Mr. Elaman Kidiraliev

He is a second-year Master degree student at Kazan Federal University. He is interested in industrial manipulator research.

Assistant Professor Roman Lavrenov



He received his PhD from Kazan Federal University in 2020. Currently, he works as an Assistant Professor at Kazan Federal University, Russia. His research interests are path planning, ROS/Gazebo, and mobile robotics.

Predictive Functional Controller Design with Disturbance Observer and Its Application

Syota Yoshida
Graduate School of Advanced Science and Engineering, Hiroshima University, 1-4-1 Kagamiyama Higashihiroshima, Hiroshima, Japan

Shin Wakitani
Graduate School of Advanced Science and Engineering, Hiroshima University, 1-4-1 Kagamiyama Higashihiroshima, Hiroshima, Japan

E-mail: yoshida-syota@hiroshima-u.ac.jp, wakitani@hiroshima-u.ac.jp
www.hiroshima-u.ac.jp

Abstract

This paper discusses a predictive functional controller (PFC) design using a disturbance observer for systems with a dead time. The PFC is a kind of model predictive controller that is an effective method for systems with a dead time. However, in the PFC, the control performance is strongly affected by the accuracy of the designed model. Therefore, this research proposes a method that the PFC maintains good control performance under the modeling error by using the disturbance observer that can estimate the disturbance even if a system has dead time. The effectiveness of the proposed control scheme is evaluated by experimental results.

Keywords: Predictive Functional Control (PFC), Disturbance Observer, Dead Time, Model Predictive Control (MPC)

1. Introduction

Model Predictive Control (MPC) [1], [2] is an effective control method for systems with a dead time. The MPC requires to solve the optimization problem online for each sampling period. So, the application of the MPC is costly. Richalet proposed predictive functional control (PFC) [3], [4] as one of the simplified MPCs**.** The PFC is effective and easy to implement for systems with a dead time. However, the accuracy of the designed model strongly affects the control performance of the PFC. It is difficult to match the designed model perfectly when the PFC is applied to a real system. Therefore, A mechanism to suppress these modeling errors is necessary. A method to suppress modeling errors using a disturbance observer (DOB) has been proposed [5]. However, the existing method using a DOB is ineffective for systems with a dead time. Therefore, this paper discusses an effective disturbance estimation method for systems with a dead time using a DOB and verifies the effectiveness of the method with experiments.

2. Predictive Functional Controller

The PFC applies only to linear time-invariant systems. Low computational cost due to methods that do not require solving optimization problems online. The basic algorithm of the PFC is explained below.

2.1. Control law derivation for PFC

Let T_s is the sampling time, and k is the arbitrary step time. The control target of a first-order system is represented by the following state-space as follows:

$$\begin{aligned} \boldsymbol{x_m}(k+1) &= \boldsymbol{A_m}x(k) + \boldsymbol{B_m}u(k) \\ \boldsymbol{y_m}(k) &= \boldsymbol{C_m}\boldsymbol{x_m}(k) \end{aligned}. \tag{1}$$

Where $\boldsymbol{x_m} \in R^n$ is the state variable vector, $u \in R$ is the control input, and $y_m \in R$ is the model output. The reference trajectory output $y_r(k+i|k)$ at time k is defined as follows:

$$y_r(k+i|k) \coloneqq r(k+i|k) - \lambda^i\{r(k) - y(k)\}. \tag{2}$$

Where $r(k)$ is the target value and $y(k)$ is the system output. where λ is the attenuation rate of the reference trajectory output and can be expressed as follows:

$$\lambda = e^{-\frac{3T_s}{t_{TRBF}}}. \tag{3}$$

t_{TRBF} is called the closed-loop response time taking the response to reach 95% of the target value, and is the main adjustment parameter of the PFC. Define an evaluation function based on the difference between the incremental reference trajectory output $\Delta y_r(k+h_j)$ and the incremental model output $\Delta y_m(k+h_j)$ at the matching point, expressed as follows:

$$J(k) \coloneqq \sum_{j=1}^{n_h} \left\{ \Delta y_m(k+h_j) - \Delta y_r(k+h_j) \right\}^2. \qquad (4)$$

Where $h_j (j = 1,2,\ldots.,n_h)$ are the sample times of the matched points, and $n_h (n_h \geq 1)$ is the number of matched points. The increment of the reference trajectory at the coincidence point, $\Delta y(k+h_j)$, can be expressed from Eq. (2) as follows:

$$\Delta y_r(k+h_j) = [r(k) - y(k)](1 - \lambda^{h_j}) \qquad (5)$$

The PFC expresses the control input $u(k)$, computed at each sampling time, as a weighted sum of polynomial basis functions (step, ramp, parabolic signal, etc.). The control input is expressed as follows:

$$u(k+i) = \sum_{l=1}^{n_B} \mu_l(k) i^{l-1} \qquad (6)$$

Where $\mu_l(k)(l = 1,2,\ldots,n_b)$ are unknown weight coefficients and are optimized so that Eq. (4) is minimized. Only the step term has a value at time k among the two-term basis functions. Therefore, only u(k) is input to the control object and is expressed as follows:

$$u(k) = \sum_{l=1}^{n_B} \mu_l(k) \coloneqq \boldsymbol{U_b}(0)^T \boldsymbol{\mu}(k) \qquad (7)$$

Where $\boldsymbol{U_b}(i)$ and $\boldsymbol{\mu}(k)$ are expressed as follows:

$$\begin{cases} \boldsymbol{U_b}(i) = [1 \quad i \quad \cdots \quad i^{n_B-1}]^T \\ \boldsymbol{\mu}(k) = [\mu_1(k) \quad \mu_2(k) \quad \cdots \quad \mu_{n_B}(k)]^T \end{cases} \qquad (8)$$

In the PFC, the mod output $y_m(k+h_j)$ at the coincident point is expressed sum of the free response $y_l(k+h_j)$ the forced response $y_f(k+h_j)$. The free-response $y_l(k+h_j)$ is expressed as follows from the state-space representation of Eq. (1).

$$y_l(k+h_j) = \boldsymbol{C_m} \boldsymbol{A_m}^{h_j} \boldsymbol{x_m}(k) \qquad (9)$$

Also, the forced response $y_f(k+h_j)$ is expressed as follows:

$$y_f(k+h_j) = \sum_{l=1}^{n_B} \mu_l(k) y_{b_l}(i) \qquad (10)$$

Where $y_{b_l}(i)$ is the response of the system to the base input $u_{b_l}(i) \coloneqq i^{l-1}$. From Eqs. (1), (9), and (10), the incremental model output $\Delta y_m(k+h_j)$at the matching point is expressed as follows:

$$\Delta y_m(k+h_j) = \boldsymbol{C_m} \left(\boldsymbol{A_m}^{h_j} - I \right) \boldsymbol{x_m}(k) + \sum_{l=1}^{n_B} \mu_l(k) y_{b_l}(i) \quad (11)$$

Where $\boldsymbol{y_b}(h_j)$ is expressed as follows:

$$\boldsymbol{y_b}(h_j) = [y_{b_1}(h_j) \quad y_{b_2}(h_j) \quad \cdots \quad y_{b_{nB}}(h_j)]^T \quad (12)$$

The evaluation function in Eq. (4) is expressed using Eqs. (5), (11), and (12) as follows.

$$J(k) = \sum_{j=1}^{n_h} \left\{ \begin{matrix} \boldsymbol{y_b}(h_j)^{\boldsymbol{T}} \boldsymbol{\mu}(k) + \boldsymbol{C_m} \left(\boldsymbol{A_m}^{h_j} - I \right) \boldsymbol{x_m}(k) \\ +(\lambda^{h_j} - 1)(r(k) - y(k)) \end{matrix} \right\} \quad (13)$$

The unknown weight coefficient vector $\boldsymbol{\mu}(k)$ that minimizes the evaluation function in Eq. (13) is obtained from $\partial J(k) / \partial \mu(k) = 0$ as follows:

$$\boldsymbol{\mu}(k) = - \left\{ \sum_{j=1}^{n_h} \boldsymbol{y_b}(h_j) \boldsymbol{y_b}(h_j)^{\boldsymbol{T}} \right\}^{-1}$$

$$\sum_{j=1}^{n_h} \left\{ C_m \left(A_m^{h_j} - I \right) x_m(k) + (\lambda^{h_j} - 1)(r(k) - y(k)) \right\} \quad (14)$$

where the vector v is defined as follows:

$$\boldsymbol{v} \coloneqq \boldsymbol{y_b}(h_j) \left\{ \sum_{j=1}^{n_h} \boldsymbol{y_b}(h_j) \boldsymbol{y_b}(h_j)^{\boldsymbol{T}} \right\}^{-1} \boldsymbol{U_b}(0) \qquad (15)$$

Also, using Eq. (15), the constants $k_0$and the constant vector $\boldsymbol{v_x}$ are defined as follows:

$$\begin{cases} k_0 = \boldsymbol{v}^T [1 - \lambda^{h_1} \quad \cdots \quad 1 - \lambda^{h_{n_h}}]^T \\ \boldsymbol{v_x} = - \left[C_m (A_m^{h_1} - I) \quad \cdots \quad C_m \left(A_m^{h_j} - I \right) \right]^T \boldsymbol{v} \end{cases} \qquad (16)$$

Using Eqs. (7), (14)-(16), the optimal control input is expressed as follows:

$$u(k) = k_0 \{ r(k) - y(k) \} + \boldsymbol{v_x}^T \boldsymbol{x_m}(k) \qquad (17)$$

Where in Eq. (17), the constants k_0 and the constant vector can $\boldsymbol{v_x}$can be computed offline. Therefore, the online calculations at each sampling time can be performed by taking the difference between the target value and the control output and the state variables of the model and multiplying them by a constant or a vector of constants.

2.2. A dead time compensation for the PFC

Let L is the dead time of the control target. The internal model expressed in Eq. (1) has a form that does not include a dead time. Therefore, the control output $y(k+L)$ after a dead time can be expressed using the model output as follows:

$$y(k+L) = y(k) + y_m(k) + y_m(k-L) \qquad (18)$$

Let $\hat{y}(k+L|k) \coloneqq y(k+L)$ be the predicted value, the control input when the control target contains wasted time is as follows:

$$u(k) = k_0 \{ r(k) - \hat{y}(k+L|k) \} + \boldsymbol{v_x}^T \boldsymbol{x_m}(k) \qquad (19)$$

3. Modeling error compensation using DOB

In the PFC, the control performance is strongly affected by the accuracy of the designed model. However, it is difficult to prepare a complete model of an actual control target in advance. Therefore, this paper considers a method to reduce the effects of modeling errors by using a DOB.

3.1. Basic structure of DOB

A compensator with a DOB reduces the effect of modeling error by taking the effect of modeling error as a disturbance and canceling it out. The basic DOB structure is shown in Fig.1. Where $G(s)$ is the transfer function of a controlled object and $G^{-1}(s)$ is the inverse function of the transfer function of the controlled object. To make the observer proper, a low-pass filter $F(s)$ must be used in combination. In this paper, $F(s)$ is designed as a multistage coupling of a first-order system expressed as follows:

$$F(s) = \left(\frac{\omega_c}{s+\omega_c}\right)^{n_s} \tag{20}$$

When the control target includes the wasted time as shown in Fig.1, the estimated input $\hat{U}(s)$ obtained by applying the system output $Y(s)$ to $G^{-1}(s)$ can be expressed as follows:

$$\hat{U}(s) = (U(s)+d)e^{-Ls} \tag{21}$$

When the control input $U(s)$ is subtracted from Eq. (21), the estimated disturbance $\hat{d}$ is expressed as follows:

$$\hat{d} = U(s)(1-e^{-Ls}) + de^{-Ls} \tag{22}$$

Eq. (22) indicates that the disturbance cannot be estimated because of the dead time.

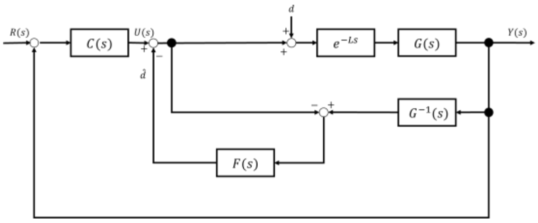


Fig. 1. Basic structure of DOB

3.2. DOB considering a dead time

The structure of the proposed DOB considering the dead time is shown in Fig.2. For the structure shown in Fig. 2, the estimated disturbance $\hat{d}$ is expressed as follows:

$$\hat{d} = U(s)\left(e^{-\hat{L}s} - e^{-Ls}\right) + de^{-Ls} \tag{23}$$

Where $\hat{L}$ represents the estimated dead time of the system. Since $\hat{L} = L$ when a dead time is known, Eq. (23) is expressed as follows:

$$\hat{d} = de^{-Ls} \tag{24}$$

Eq. (24) shows that the proposed DOB configuration is capable of suppressing disturbances in the situation where the system does not have a large amount of dead time.

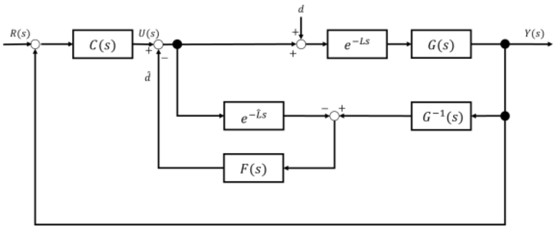


Fig. 2. DOB considering a dead time

4. Numerical example and experimental

The effectiveness of the proposed method is verified by numerical examples and actual experiments. The temperature experiment apparatus used is shown in Fig.3. Approximating the target as a first order + dead time system, the system was identified and $K = 4.27$, $T = 37.6$, $L = 2$ were obtained.

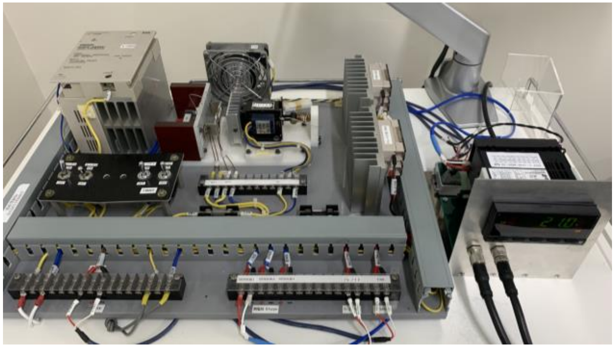


Fig. 3. Temperature experimental device

4.1. Simulation

A numerical example is performed using the obtained model. The system is a first order delay + dead time system with $K \in [2,5], T \in [30,50], L = 2$. Gain and time constants were prepared using random numbers within a range. Conditions for numerical examples are shown in Table 1. The results of applying only the PFC are shown in Fig.4, applying the DOB for the basic structure is shown in Fig.5, and the DOB for the proposed structure is shown in Fig.6. Ten trials were used for the display. When only the PFC is used, the control performance varies due to modeling errors. The response of the basic structure is oscillating when a dead time is included. However, the proposed structure of the DOB appropriately reduces the influence of modeling error and improves controllability.

Table 1. Numerical example conditions

t_{TRBF}	h_1	h_2	h_3	ω_c	u_{max}	u_{min}
30	10	15	30	5	100	0

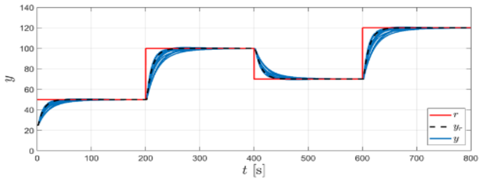


Fig. 4. Simulation result for the PFC

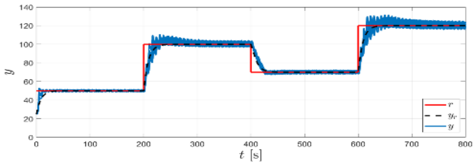


Fig. 5. Simulation result for Basic DOB

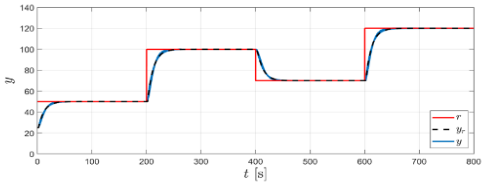


Fig. 6 Simulation result for proposed DOB

4.2. Experimental

The effectiveness of the proposed method is verified by the experiments shown in Fig.3. Experimental conditions are the same as in Table 1. The results of applying only the PFC are shown in Fig.7, and the results of applying the DOB for the proposed structure are shown in Fig.8. The results confirm the effectiveness of the proposed method through experiments.

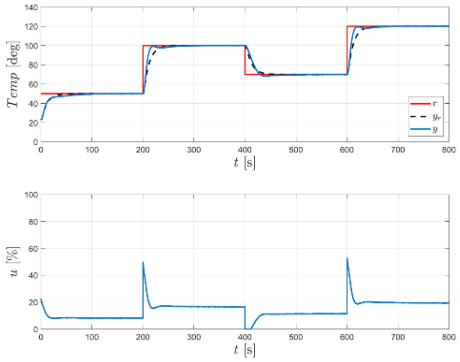


Fig. 7. Experimental result for the PFC

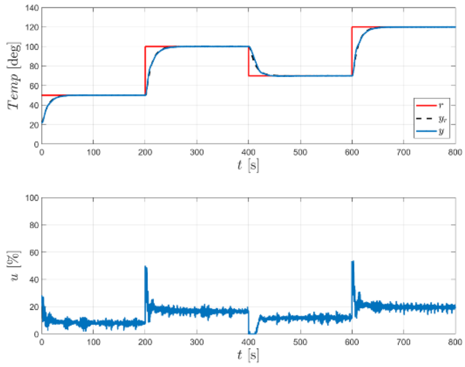


Fig. 8. Experimental result for proposed DOB

5. Conclusion

In this paper, the Eq using the DOB considering the dead time was proposed. The effectiveness of the proposed method was demonstrated through numerical examples and experiments for a system with a dead time. It was confirmed that the proposed structure does not cause oscillations in the response of the system even for systems with a dead time and reduces the influence of modeling errors. In the future, consider designing a data-driven observers that can be applied to larger environmental changes.

References

1. J. M. Maciejowski: *Predictive Control with Constraints*, Prentice-Hall（2002）
2. E. F. Camacho and C. Bordons: *Model Predictive Control (Second Edition)*, Springer-Verlag (2007)

3. J. Richalet, S. Abu el Ata-Doss, C. Arber, H.B. Kuntze, A. Jacubasch and W. Schill: Predictive functional control: application to fast and accurate robot, *Proc. of IFAC 10th World Congress*, 251/258 (1987)
4. J. Richalet, D. O' Donovan: *Predictive Functional Control*, Springer, (2009)
5. T. Sato, K. Kaneko, N. Saito: Improvement of Tracking Performance in Predictive Functional Control Using Disturbance Observer, *Transactions of the Society of Instrument and Control Engineers*, vol. 10, no. 15, pp. 132-139, (2011) (in Japanese)

Authors Introduction

Mr. Syota Yoshida

He received his Diploma from the Department of Electrical and Electronic Engineering, Hiroshima University in 2022. Currently, he is a graduate student at Hiroshima University.

Professor. Shin Wakitani

He received his PhD from the Department of Electrical Engineering, and Master's degree from the Department of Pedagogy, Hiroshima University in 2011 and 2013, respectively. Currently, he is an Associate Professor in Hiroshima University. Her research interest includes control engineering.

Study on a Construction of Velocity Perception Model and Kansei Feedback Control System in Active Behavior

Sota Takemura, Takuya Kinoshita, Toru Yamamoto
Hiroshima University, 1 Chome-3-2 Kagamiyama, Higashihiroshima, Hiroshima 739-8511, Japan

E-mail: Takemura-sota@hiroshima-u.ac.jp

Abstract

In Japanese, there is a word "Kansei" which means "feelings, impulses, and desires stimulated by the senses. In addition, there is a field called Kansei Engineering which is recognized worldwide. In the case of human-operated machines, it is necessary to consider Kansei such as whether the operability is comfortable. Therefore, this paper describes a model that focuses on Kansei and control based on the model. It has been proposed that a human Kansei model based on the Weber-Fechner law. However, the Weber-Fechner law is applied to stimulus such as sound, smell, and light. The model needs to be improved to reduce the error between the velocity assumed in the brain and the actual velocity of the machines. Therefore, this paper proposes a new Kansei model that focuses on the relationship between the actual velocity and the perceived velocity when a human actively operates the machines.

Keywords: PID Control, Database-Driven Control, Kansei

1. Introduction

In recent years, some machines are operated by humans such as cars and construction machines. Since these machines driven by humans, it is necessary that focusing on comfortable operability. In Japanese, there is a word "Kansei" which means "feelings, impulses, and desires stimulated by the senses" [1]. In addition, there is a field called Kansei Engineering which is recognized worldwide [2] [3]. Therefore, this paper describes a model that focuses on Kansei and control based on the model. The control system design that considers human Kansei has been studied [4]. It has been proposed that a human Kansei model based on the Weber-Fechner law. However, Weber-Fechner's law applies to stimulus such as sound, smell, and light. smell, light, and other stimulus. Therefore, it is necessary that further improvement of the model to account for the error between the velocity assumed in the brain and the actual velocity.

This paper proposes a new Kansei model. The model is focus on the relationship between the actual velocity and the perceived velocity [5] when a human actively operates the controlled system.

2. Control System Focusing on the Moment of Inertia

In this paper, it is assumed that a person repeatedly operates the machines. Therefore, the number of trials q is adjusted instead of time t. Fig. 1 shows the control system focusing on the moment of inertia. In the controll system, $J_h(q)$ is the moment of inertia in the human brain, and D is the viscosity. $J_{ref}(q)$ is the adaptively adjusted moment of inertia. $y_{\omega h}(t)$, $r_{\omega h}(t)$ and $e_{\omega h}(t)$ are mean of the actual velocity of the controlled system, the velocity assumed in the brain, the error between $y_{\omega h}(t)$ and $r_{\omega h}(t)$, respectively. The controlled system is most comfortable when $J_{ref}(q)$ matches $J_h(q)$. In this paper, the database-driven control scheme [6] [7] is utilized to adjust $J_{ref}(q)$ so that the Kansei $y(t)$ is improved. Details of the database-driven control are given in the section 4.

3. Construction of a model of human Kansei

3.1. The Conventional Kansei Model

The following conventional model of human Kansei has been proposed using the Weber-Fechner law [4]:

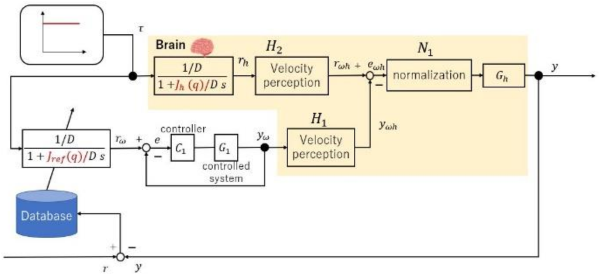


Fig. 1. Schematic figure of the proposed model.

$$y(t)=\frac{1}{1+E(t)\log\left(1+|e_{\omega h}(t)|\right)} \tag{1}$$

$$e_{\omega h}(t)=r_{\omega h}(t)-y_{\omega h}(t), \tag{2}$$

where $E(t)$ is a variable that depends on each person. $r_h(t)$ is the velocity of the controlled system as assumed in a human's brain. It is calculated by the following equation:

$$R_h(s)=\frac{\frac{1}{D}}{1+\frac{J_h(q)}{D}s}T(s). \tag{3}$$

Equation (1) shows that the Kansei $y(t)$ approaches 1 when $e_{\omega h}(t)$ is small. In contrast, the Kansei $y(t)$ is close to 0 when $e_{\omega h}(t)$ is large. However, the Weber-Fechner law used in the above model is valid for simple stimulus such as smell, sound, and light. It is not appropriate for the case of stimulus error assumed in this paper. For example, if the velocity in the brain $r_{\omega h}(t)$ is 50 km/h and the actual velocity of the controlled system $y_{\omega h}(t)$ is 51 km/h. In this case, the error $e_{\omega h}(t)$ is as small as 1 km/h, but the Kansei $y(t)$ is greatly reduced in equation (1). A human would not be so sensitive to small errors. Therefore, the Kansei model in the next section is proposed to improve the model.

3.2. The Proposed Perception Kansei Model

The proposed model focuses on the characteristic of a human actively operating the controlled system [5]. Humans perceive the velocity as fast when the velocity of the controlled system is slow. Conversely, humans perceive the velocity as slow when the one is fast. However, the mathematical formulation has not been proposed in the conventional studies. Therefore, this paper constructs a model for it. The specific model is H_1 in Fig. 1. H_1 is a model that outputs the perceived velocity $y_{\omega h}(t)$ of $y_\omega(t)$. Specifically, the model is given as follows:

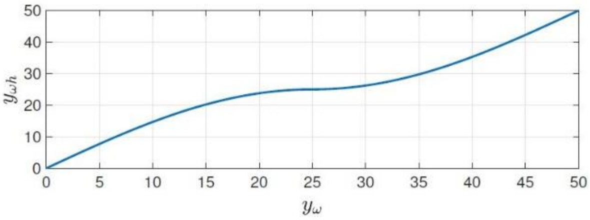


Fig. 2. Relationship between $y_{\omega h}$ and y_ω.

$$y_{\omega h}(t)=\begin{cases}\frac{l}{2}M, & \left(0\le y_\omega(t)\le\frac{l}{2}\right)\\ l\left(1-\frac{M}{2}\right) & \left(\frac{l}{2}<y_\omega(t)\le l\right)\end{cases} \tag{4}$$

$$M=\cos\left\{\left(\frac{y_\omega(t)}{l}-\frac{1}{2}\right)\pi\right\}, \tag{5}$$

where l represents the upper-velocity limit value of the controlled system. In this paper, l is set as l = 50. Fig. 2 shows the relationship between the input and output of H_1. In Fig. 2, the perceived velocity $y_{\omega h}(t)$ is greater than $y_\omega(t)$ when $y_\omega(t)$ < 25. Conversely, the perceived velocity $y_{\omega h}(t)$ is not greater than $y_\omega(t)$ when $y_\omega(t)$ > 25. For the above, the model expresses the actual velocity of controlled system. On the other hand, model H_2 that generates the reference velocity $r_{\omega h}(t)$ in the brain is given in the same way. Calculate the Kansei $y(t)$ from 0 to 1 based on the velocity error $e_{\omega h}(t)$ in model N_1 in Fig. 1. Specifically, it is given by the following equation:

$$y(t)=1-\frac{|e_{\omega h}(t)|}{l} \tag{6}$$

$$e_{\omega h}(t)=r_{\omega h}(t)-y_{\omega h}(t). \tag{7}$$

It is considered that human do not statically judge the Kansei from the current stimulus only but dynamically. In this paper, $G_h(s)$ in Fig. 1 is given by the following first-order system:

$$G_h(s)=\frac{1}{1+2s}. \tag{8}$$

4. Database-Driven Control System Design [6] [7]

4.1. [Step 1] Creation of Initial Database

Controller parameters can be adaptively adjusted by the database-driven control system. This scheme has a database including reference signal, output, and $J_{ref}(q)$. Therefore, control cannot be performed if the database does not exist. Thus, the moment of inertia $J_{ref}(q)$ is set to a fixed value, and an initial database is created consisting of an information vector $\boldsymbol{\phi}(q)$ in the following equation:

$$\boldsymbol{\Phi}_{\mathbf{j}} := [\boldsymbol{\phi}(q_j),\ J_{ref}(q_j)]\ (j = 1,2,\dots,N(0)) \tag{9}$$

$$\boldsymbol{\phi}(q) := [\,\bar{r}(q+1),\ \bar{y}(q),\ \overline{r_\omega}(q),\ \overline{y_\omega}(q)], \tag{10}$$

where j is the jth dataset in the database. $N(0)$ indicates the number of datasets obtained by the initial operation. $\bar{r}(q)$, $\bar{y}(q)$, $\overline{r_\omega}(q)$ and $\overline{y_\omega}(q)$ are the mean value of reference Kansei $r(t)$, Kansei $y(t)$, the reference controlled system velocity $r_\omega(t)$, and output velocity of the controlled system $y_\omega(t)$, respectively.

4.2. [Step 2] Calculation of Distance, Selection of Neighbors

Adjust $J_{ref}(q)$ using the database $\boldsymbol{\Phi}_{\mathbf{j}}$ obtained in [Step 1]. Run the controlled system and obtain the information vector $\boldsymbol{\phi}(q)$ on trial q. Next, calculate the distance $d_j\left(\boldsymbol{\phi}(q),\boldsymbol{\phi}(q_j)\right)$ between $\boldsymbol{\phi}(q)$ and $\boldsymbol{\phi}(q_j)$ in the dataset $\boldsymbol{\Phi}_j$ in database. $d_j\left(\boldsymbol{\phi}(q),\boldsymbol{\phi}(q_j)\right)$ is calculated using the following weighted norm:

$$d_j\left(\boldsymbol{\phi}(q),\boldsymbol{\phi}(q_j)\right) = \sum_{p=1}^{4}\left|\frac{\boldsymbol{\phi}_p(q)-\boldsymbol{\phi}_p(q_j)}{\max\left(\boldsymbol{\phi}_p(m)\right)-min\left(\boldsymbol{\phi}_p(m)\right)}\right|, \tag{11}$$

$$\left(j = 1,2,\dots,N(q)\right)$$

where $N(q)$ represents the number of datasets stored in the database in q trials. Also, $\boldsymbol{\phi}_p(q)$ represents the pth element of $\boldsymbol{\phi}(q)$, and the same is true for $\boldsymbol{\phi}_p(q_j)$. $\max\left(\boldsymbol{\phi}_p(m)\right)$ indicates the largest element among the pth element of all the information vectors $\boldsymbol{\phi}(q_j)$ in the database. Similarly, $\min\left(\boldsymbol{\phi}_p(m)\right)$ indicates the smallest element among the pth element of all the information vectors $\boldsymbol{\phi}(q_j)$ in the database. The top n information vectors are extracted from the ones with the smallest distance d_j as neighbors.

4.3. [Step 3] Calculation of $J_{ref}(q)$

For the neighbors selected in [Step 2] d_j, a local model is constructed using the linearly weighted average shown in below:

$$J_{ref}^{old}(q) = \textstyle\sum_{i=1}^{n}\omega_i J_{ref}(q_i), \tag{12}$$

where ω_i is the weight for $J_{ref}(q_i)$ in the ith selected information vector, which should satisfy $\sum_{i=1}^{n}\omega_i = 1$. The weight ω_i is given by the next equation:

$$\omega_i = \frac{\exp(-d_i)}{\sum_{i=1}^{n}\exp(-d_i)}. \tag{13}$$

The control system is operated using the $J_{ref}^{old}(q)$ calculated by the above scheme.

4.4. [Step 4] Data Correction

The desired control performance is not always achieved when using the $J_{ref}^{old}(q)$ in equation (12). Then, $J_{ref}^{old}(q)$ is updated according to the control error. The corrected data $J_{ref}^{new}(q)$ is stored in a database. The following steepest descent scheme is used for learning:

$$J_{ref}^{new}(q) = J_{ref}^{old}(q) - \eta\frac{\partial L(q+1)}{\partial J_{ref}^{old}(q)}, \tag{14}$$

where η is the learning coefficients and $L(q+1)$ represents the error evaluation norm defined below:

$$L(q) := \frac{1}{2}\bar{\varepsilon}(q)^2 \tag{15}$$

$$\bar{\varepsilon}(q) := \bar{r}(q) - \bar{y}(q). \tag{16}$$

5. Numerical Example

In Fig. 1, the following equation is given for the controlled system:

$$G_1(s) = \frac{25}{1+50s}. \tag{17}$$

In this numerical example, the parameters were set as reference Kansei $r(t) = 1$, viscosity $D = 0.05$, and operating torque $\tau(t) = 1$. The initial value of the

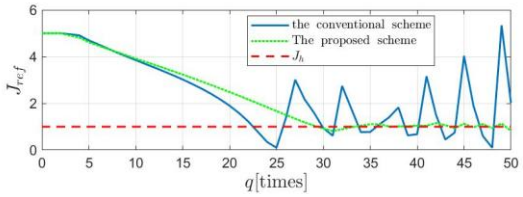


Fig. 3. Moment of inertia of $J_{ref}(q)$.

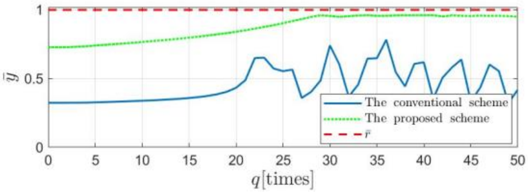


Fig. 4. Average Kansei $\bar{y}(q)$.

moment of inertia were set to $J_{ref}(0)$ = 5. The moment of inertia in brain, $J_h(q)$ = 1. The PI control law is applied for the controller C_1. The initial PI gains were calculated by the FRIT scheme. K_P and K_I are set as K_P = 0.38, K_I = 0.04. The learning coefficient was set to η = 15 for the conventional model, and set η = 40 for the proposed model. The average Kansei $\bar{y}(q)$ and $J_{ref}(q)$ are shown in Fig. 3 and Fig. 4. From Fig. 3 and Fig. 4, the Kansei of the conventional model is significantly reduced worsened by a small velocity error. In contrast, the proposed model confirms that $\bar{y}(q)$ approaches 1 when $J_{ref}(q)$ approaches $J_h(q)$. The proposed scheme adjusts to the desired moment of inertia ($J_{ref} \rightarrow J_h$). The effectiveness of the proposed scheme is confirmed by the fact that the Kansei output is achieved to the reference signal ($\bar{y}$ = 1).

6. Conclusions

In this paper, the new Kansei model has been proposed that consider the active behavior of humans. The proposed model also achieves the desired moment of inertia. In this case, it has been confirmed that the Kansei is close to the reference signal. In the future, it is expected to the effectiveness of the proposed scheme through experiments.

References

1. I Shinmura. Kojien 5th edition. Iwanami, 1998
2. Kansei Engineering International Journal
3. M. Nagamachi, Kansei Engineering: A new ergonomic consumer-oriented technology for product development. International Journal of Industrial Ergonomics 1995, Vol. 15, No. 1: 3-11
4. T Kinsohita, H Ikeda, T Yamamoto, M Machizawa, K Tanaka, Y Yamazaki. Design of a Database Driven Kansei Feed-back-Control System using a Hydraulic Excavators Simulator. Journal of Robotics and Mechatronics 2020, Vol. 32, No. 3: 652-661
5. K Yasunaga, H Kaneko. Caused by own actions Perceived velocity of motion. TVRSJ 2017, Vol.22 No.1: 113-123
6. T Yamamoto, K Takao, T Yamada. Design of a Data-Driven PID Controller. IEEE Trans on control systems Technology 2009, Vol.17 No.1: 29-39
7. T Yamamoto, O Kaneko, S Wakitani, T Kinoshita, Y Oonishi, H Kugemoto, K Koiwai. Data Oriented PID Control. Morikita, 2020

Authors Introduction

Mr. Sota Takemura

He received his B. Eng. from Hiroshima University in Japan in 2022. He belongs Department of Advanced Science and Technology in Hiroshima University for Master. His research interests are database-driven control.

Assistant Professor. Takuya Kinoshita

He received his B. Eng., M. Eng and D. Eng. from Hiroshima University in Japan in 2013, 2015 and 2017, respectively. He was postdoctoral fellow of JSPS (Japan Society for the Promotion of Science) in 2017. His research interests are performance-driven control.

Professor. Toru Yamamoto

He received the B.Eng. and M. Eng. degrees from Tokushima university, Tokushima, Japan, in 1984 and 1987, respectively, and the D.Eng. degree from Osaka University, Osaka, Japan, in 1994. His current research interests are data-driven control, and process control.

Design of a Database-Driven Control System for a Web Conveyor

Atsushi Takatani, Takuya Kinioshita, Toru Yamamoto
Hiroshima University, 1 Chome-3-2 Kagamiyama, Higashihiroshima, Hiroshima 739-8511, Japan

E-mail: takatani-atsushi@hiroshima-u.ac.jp, kinoshita-takuya@hiroshima-u.ac.jp, yama@hiroshima-u.ac.jp

www.cse.hiroshima-u.ac.jp/

Tomohiro Hirakawa, Hiroki Hamamoto, Takashi Ochiiwa, Hideki Tomiyama
Japan Steel Works, Ltd, 1-6-1 Funakoshiminami, Aki-ku, Hiroshima, Hiroshima 736-8602, Japan

E-mail: hiroki_hamamoto@jsw.co.jp, tomohiro_hirakawa@jsw.co.jp, takashi_ochiiwa@jsw.co.jp, hideki_tomiyama@jsw.co.jp

www.jsw.co.jp/en/

Abstract

Web Conveyor used in the processing of plastic films and other materials is a large-scale system consisting of multiple drive rolls, and it is difficult to understand the system characteristics during driving. This paper proposes a database-driven control scheme based on the FRIT scheme to implement motor control of this conveyor without system identification. To reduce the computational cost, the concept of database design based on similarity is introduced. The web conveyor's operating output can quickly achieve a reference signal, because it can be based on a database that holds operational data.

Keywords: PID Controller, Database-Driven Control, Web Conveyor

1. Introduction

A web is a thin, long material such as paper or plastic. It is used in a variety of applications, including paper rolls and wrapping film. Generally, the web is finally shipped as a rolled product. As a specific processing procedure, the web is stretched, printed, and rolled into a rolled product by a web conveyor. Ideally, the web feeding speed and tension should be kept constant during processing to prevent wrinkling and breakage. These are achieved by appropriately controlling the rotation of the multiple drive rolls in the web conveyor.

On the other hand, web characteristics change with temperature and humidity. Furthermore, the web conveyor characteristics are not constant according to the radius and moment of inertia of the drive rolls, which change during take-up. In addition, actual equipment includes uncertain dynamics such as bearing friction. For these reasons, it is difficult to determine appropriate control parameters for schemes that require knowledge of system characteristics.

Web conveyors are large-scale systems that contain many drive rolls. The application of decentralized control is being considered to improve its maintainability [1]. Furthermore, self-tuning of the web tension control system using the FRIT scheme has been proposed to solve the problem of difficulty in determining system characteristics [2]. However, this takes time to reach the reference signal because it requires parameter calculations for each operation.

Therefore, this paper applies database-driven control[3] to improve the control performance of web conveyors. Also, offline adjustment of control parameters using the same scheme is discussed. The database-driven control scheme can store past operational data. Therefore, it is possible to adjust control parameters offline based on the operational data. It is expected that the startup time of the next web conveyor will be reduced.

2. Overview of Web Tension Control System

2.1. *Modeling of a Web Tension Control System*

A schematic diagram of the web conveyor is shown in Fig. 1. The system consists of four drive rolls equipped with servomotors, two of which are equipped with sensors to measure tension. The web is transported from the unwinder, through the lead section and draw rolls to the winder.

In reference [1], the web has viscoelastic properties. Therefore, introduce the Voigt model of equation (1), which considers viscoelasticity as a characteristic of the web:

$$\boldsymbol{P}(\boldsymbol{s}) = \boldsymbol{A}\left(\boldsymbol{\eta_v} + \frac{\boldsymbol{G_v}}{\boldsymbol{s}}\right), \qquad (1)$$

where G_v and η_v are the elastic modulus and viscous modulus, respectively.

The control variables in this paper are tension T_1 and T_3 and transfer speed V_2 and V_3.

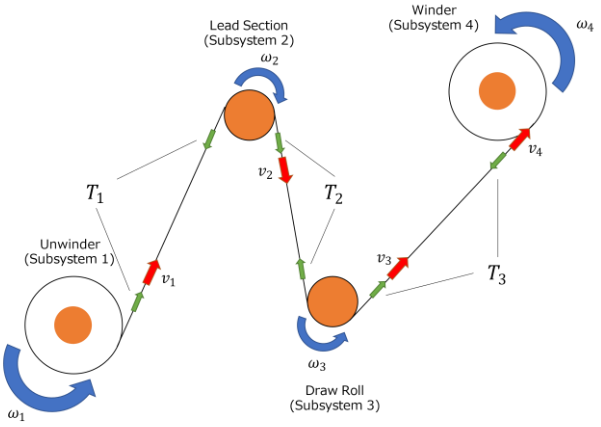


Fig. 1 Schematic diagram of web conveyor.

2.2. *Control System Design*

Fig. 2 shows a block diagram of the web conveyor control system. Subsystems 1 through 4 are assigned to each servo motor. In this paper, the "Controller Section" controls the four servo motors. In addition, a "Sequencer" is added to implement control to reduce instability caused by changes in the characteristics of the controlled object and mutual interference between the subsystems. "Sequencer" uses a proportional-priority PI controller (I-P controller) and applies a database-driven control scheme. The sampling time of the controller is set to 100 [ms] to avoid the upper limit of the computational cost by the PLC, assuming the verification in an actual experiment. In the "Controller section", a PI controller is used. Since the PI gain does not change, the sampling time of the controller is set to 10 [ms].

3. Database-Driven Control Scheme

3.1. *Database-Driven Control Schemes Based on Similarity*

When applying the database, as shown in Reference [4], the PI gains are determined by selecting a certain number of neighbors from the created database in the order of system conditions. Therefore, if the system conditions at the time of application differ significantly from those in the database, it may not be possible to determine the appropriate PI gain, because inappropriate neighbors may be selected. In addition, this controlled system is assumed to operate for a long period of time, up to one hour. Therefore, if the off-line adjustment of control parameters is performed based on the FRIT scheme, the number of databases becomes enormous. Correspondingly, the computational cost of off-line adjustment and determination of PI gains becomes very high. Depending on the calculation cost, the PLC may not be able to complete the calculation within the control cycle.

To address these problems, a database-driven control scheme based on similarity using kernel density estimation has been proposed [5]. The scheme can efficiently create databases because it calculates the similarity between a query and a database. It is also possible to select neighbors in the same way.

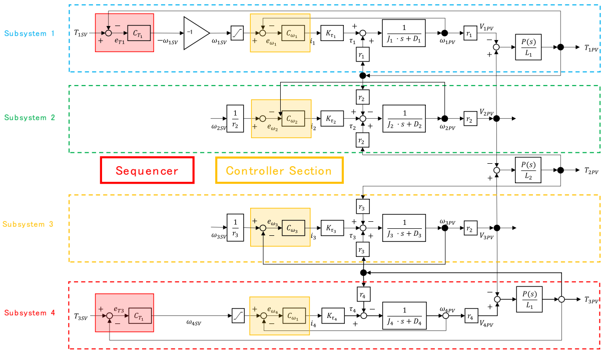


Fig. 2 Block Diagram of Web Tension Control System.

3.2. *Dataset*

The data sets required for the application of the database-driven control scheme are shown in equations (2) through (4) as follows:

$$\boldsymbol{\Phi_j} = [\bar{\boldsymbol{\phi}}(\boldsymbol{k_j}), \boldsymbol{\theta}(\boldsymbol{k_j})], \tag{2}$$

$$\begin{aligned}\bar{\boldsymbol{\phi}}(k_j) = [&r_0(k_j+1), r_0(k_j), y_0(k_j), \dots, y_0(k_j-n_y), \\ &\boldsymbol{u_0}(\boldsymbol{k_j}), \dots, \boldsymbol{u_0}(\boldsymbol{k_j}-\boldsymbol{n_u}), \\ &y_m(k_j), \dots, y_m(k_j-n_y)],\end{aligned} \tag{3}$$

$$\boldsymbol{\theta}(\boldsymbol{k_j}) = [\boldsymbol{K_P}(\boldsymbol{k_j}), \boldsymbol{K_I}(\boldsymbol{k_j})], \tag{4}$$

where j is the jth dataset in the database. Note that $r_0(t), u_0(t), y_0(t), y_m(t)$ are the operation data with the fixed PI gains, which are the reference signal, input, output, and reference output, respectively.

3.3. *Similarity Calculation*

When creating, training, or adapting a database, it is necessary to calculate the similarity between a query and a database for each case.

When adding new data $\bar{\boldsymbol{\phi}}(n)$, the similarity is calculated by the kernel density function-based similarity $S\left(\bar{\boldsymbol{\phi}}(n), \bar{\boldsymbol{\phi}}_A(k_j)\right)$ as follows:

$$\boldsymbol{S}\left(\bar{\boldsymbol{\phi}}(\boldsymbol{n}), \bar{\boldsymbol{\phi}}_{\boldsymbol{A}}(\boldsymbol{k_j})\right) = \prod_{\boldsymbol{i}=\boldsymbol{1}}^{\boldsymbol{n_y}+\boldsymbol{n_u}+\boldsymbol{1}} \frac{\boldsymbol{1}}{\sqrt{\boldsymbol{2\pi}}\boldsymbol{h_i}} \boldsymbol{exp}\left(-\frac{\bar{\boldsymbol{\phi}}(\boldsymbol{n},\boldsymbol{i}) - \left(\bar{\boldsymbol{\phi}}_{\boldsymbol{A}}(\boldsymbol{k_j},\boldsymbol{i})\right)^2}{\boldsymbol{2h_i^2}}\right). \tag{5}$$

$\bar{\boldsymbol{\phi}}_A(k_j)$ is composed of the reference signal, input, and output, which are extracted from the dataset in equation (3) as follows:

$$\begin{aligned}\bar{\boldsymbol{\phi}}_A(k) = [&r_0(k_j+1), r_0(k_j), y_0(k_j), \dots, y_0(k_j-n_y), \\ &\boldsymbol{u_0}(\boldsymbol{k_j}), \dots, \boldsymbol{u_0}(\boldsymbol{k_j}-\boldsymbol{n_u})].\end{aligned} \tag{6}$$

Note that $\bar{\boldsymbol{\phi}}_A(k_j)$ in equation (6) refers to the k_{th} information vector and j element in the dataset $\bar{\boldsymbol{\phi}}_A$, and $\bar{\boldsymbol{\phi}}(n_j)$.

The h_i in equation (5) is the bandwidth and it is computed using the plug-in scheme similar to the one shown in reference [5]. The h_j is computed using the standard deviation σ_i and the mean μ_j when the number of data is N.

$$\boldsymbol{h_i} = \frac{\boldsymbol{1.06\sigma_i}}{\boldsymbol{N}^{(\boldsymbol{1/5})}} \tag{7}$$

$$\boldsymbol{\sigma_i} = \sqrt{\frac{\boldsymbol{1}}{\boldsymbol{N}}\left(\bar{\boldsymbol{\phi}}_{\boldsymbol{A}}(\boldsymbol{k_j},\boldsymbol{i}) - \boldsymbol{\mu_i}\right)} \tag{8}$$

$$\boldsymbol{\mu_i} = \frac{\boldsymbol{1}}{\boldsymbol{N}}\bar{\boldsymbol{\phi}}_{\boldsymbol{A}}(\boldsymbol{i}). \tag{9}$$

In equation (5), the highest similarity is obtained when the information vectors are exactly the same, as follows:

$$\boldsymbol{S}\left(\bar{\boldsymbol{\phi}}(\boldsymbol{n}), \bar{\boldsymbol{\phi}}(\boldsymbol{n})\right) = \prod_{\boldsymbol{i}=\boldsymbol{1}}^{\boldsymbol{n_y}+\boldsymbol{n_u}+\boldsymbol{1}} \frac{\boldsymbol{1}}{\sqrt{\boldsymbol{2\pi}}\boldsymbol{h_i}}. \tag{10}$$

On the other hand, if the similarity is low, $S\left(\bar{\boldsymbol{\phi}}(n), \bar{\boldsymbol{\phi}}_A(k_j)\right) \to 0$. Here, a threshold $\alpha_{th}(0 < \alpha_{th} \leq 1)$ is introduced. Add new data only if the following equation (11) is satisfied for all data in the existing database:

$$\boldsymbol{S}\left(\bar{\boldsymbol{\phi}}(\boldsymbol{n}), \bar{\boldsymbol{\phi}}_{\boldsymbol{A}}(\boldsymbol{k_j})\right) < \boldsymbol{\alpha_{th}} \cdot \prod_{\boldsymbol{i}=\boldsymbol{1}}^{\boldsymbol{n_y}+\boldsymbol{n_u}+\boldsymbol{1}} \frac{\boldsymbol{1}}{\sqrt{\boldsymbol{2\pi}}\boldsymbol{h_i}}. \tag{11}$$

In the same way, consider the case of determining the nearest neighbors. When applying the database, it is necessary to prevent the phenomenon that the number of neighbors tends to become zero when the output improves and differs significantly from the initial output. For this reason, the similarity is calculated from $\bar{\boldsymbol{\phi}}_A(k_j)$, in which the reference signal, output, and reference output are extracted from the data set in equation (8):

$$\begin{aligned}\bar{\boldsymbol{\phi}}_B(k_j) = [&r_0(k_j+1), r_0(k_j), y_0(k_j), \dots, y_0(k_j-n_y), \\ &\boldsymbol{y_m}(\boldsymbol{k_j}), \dots, \boldsymbol{y_m}(\boldsymbol{k_j}-\boldsymbol{n_y})].\end{aligned} \tag{12}$$

The nearest neighbors to the query $\bar{\boldsymbol{\phi}}_i(t)$ at some time t is assumed to satisfy the following equation (13) determined by the threshold $\beta_{th}(0 < \beta_{th} \leq 1)$:

$$\boldsymbol{S}\left(\bar{\boldsymbol{\phi}}_{\boldsymbol{i}}(\boldsymbol{t}), \bar{\boldsymbol{\phi}}_{\boldsymbol{B}}(\boldsymbol{k_j})\right) \geq \boldsymbol{\beta_{th}} \cdot \prod_{\boldsymbol{i}=\boldsymbol{1}}^{\boldsymbol{n_y}+\boldsymbol{n_y}+\boldsymbol{1}} \frac{\boldsymbol{1}}{\sqrt{\boldsymbol{2\pi}}\boldsymbol{h_i}}. \tag{13}$$

4. Numerical Example Verification

4.1. *The threshold for similarity calculation*

The threshold value for initial database creation is $\alpha_{th} = 0.7$, and the threshold value for PI gain training and neighbors determination is $\beta_{th} = 0.1$.

4.2. *Simulation result and discussion*

First, simulations with fixed PI gains were performed to create the data set. The number of data acquired at that time was 1997, which was reduced to 234 and stored in the database. Next, the database was trained and simulations were performed using it. The results are shown in Fig. 3 and Fig. 4, which shows the results for subsystem 1.

Focus on the output result of the tension T_{1PV} in Fig. 3. It takes a very long time for the output to reach the reference signal, when controlled with the fixed PI gains. On the other hand, the control parameters are adjusted to correspond to the system characteristics, when database-driven control is applied. Therefore, it can be confirmed that the rise time is shorter compared to the control with the fixed PI gains.

Fig. 4 shows the results of PI gain transition and the number of neighbors in the sequencer section of subsystem 1. The PI gain fluctuates for about 10 seconds after the start of the simulation to improve the rise time.

The above results confirm that simulation with a database that has been trained improves the startup. In addition, efficient database creation was achieved by incorporating the concept of database-driven control based on similarity. Thus, the effectiveness of the database-driven control scheme for web conveyors was confirmed.

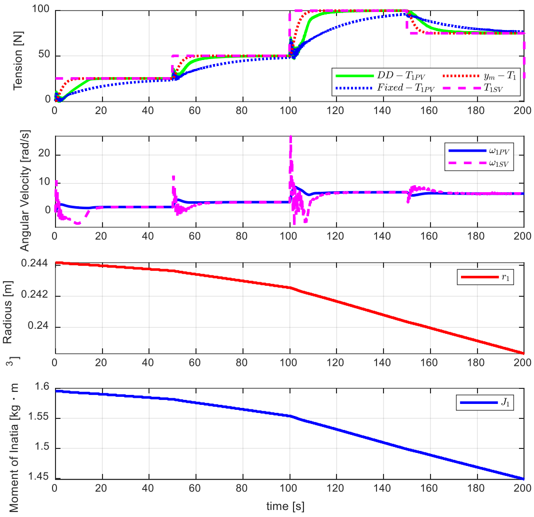


Fig. 4 Control result in subsystem 1 by the proposed scheme.

5. Summary

This paper proposed an offline adjustment scheme of PI gains by applying database-driven control to a web tension control system. The effectiveness of the proposed scheme was verified using numerical examples.

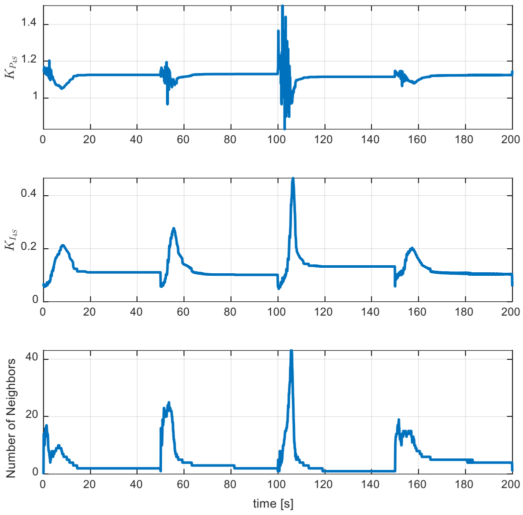


Fig. 3 The trajectories of PI gains and number of neighbors corresponding to Fig. 3.

Plans include the design of a reference signal filter, investigation of soft sensors, and experiments on actual equipment to further improve control performance.

Reference

1. Tetsuzo Sakamoto, Analysis and Control of Web Tension Control System, IEEJ Transactions on Industry Applications, 1997, Volume 117, Issue 3, Pages 274-280
2. Kohei Fukushima, Tetsuzo Sakamoto, Self-tuning of Web Tension Control System via FRIT and OPSO using Linearized Evaluation Function, IEEJ Transactions on Industry Applications, 2018, Volume 138, Issue 2, Pages 128-134
3. Kenji Takao, Toru Yamamoto, Takao Hinamoto, A Design of Generalized Predictive Control Systems Using a Memory-Based System Identification, IEEJ Transactions on Electronics, Information and Systems, 2005, Volume 125, Issue 3, Pages 442-449, Released on J-STAGE June 01, 2005, Online ISSN 1348-8155, Print ISSN 0385-4221
4. Toru Yamamoto (ed.): Data Oriented PID Control, Morikita Publishing (2020)
5. Takuya Kinoshita, Toru Yamamoto, Design of a Database-Driven Control System based on the

Similarity, IEEJ Transactions on Electronics, Information and Systems, 2020, Volume 140, Issue 3, Pages 312-319

Authors Introduction

Mr. Atsushi Takatani

He received his B. Eng. from Hiroshima University in Japan in 2022. He is currently a master's student at Graduate School of Advanced Science and Engineering, Hiroshima University, Japan.

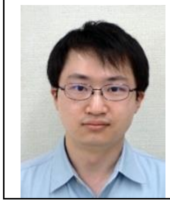
Assistant Professor Takuya Kinoshita

He received his B. Eng., M. Eng and D. Eng. from Hiroshima University in Japan in 2013, 2015 and 2017, respectively. He was postdoctoral fellow of JSPS (Japan Society for the Promotion of Science) in 2017. He is currently an Assistant Professor with the Department of Graduate School of Advanced Science and Engineering, Hiroshima University, Japan. His research interests are performance-driven control.

Professor Toru Yamamoto

He received the B.Eng. and M. Eng. degrees from Tokushima university, Tokushima, Japan, in 1984 and 1987, respectively, and the D.Eng. degree from Osaka University, Osaka, Japan, in 1994. He is currently a Professor with the Graduate School of Advanced Science and Engineering, Hiroshima University, Japan. He was a Visiting Researcher with the Department of Mathematical Engineering and Information Physics, University of Tokyo, Tokyo, Japan, in 1991. He was an Overseas Research Fellow of the Japan Society for Promotion of Science with the University of Alberta for six months in 2006. His current research interests are in the area of data-driven control, and process control. Dr. Yamamoto was the recipient of the Commendation for Science and Technology by the Minister of Education, Culture, Sports and Technology in 2009.

Mr. Tomohiro Hirakawa



He received his B. Eng. and M. Eng. from Hiroshima university, Japan, in 2015 and 2017, respectively. Currently, he works in the Technical Development Department, The Japan Steel Works, LTD., Japan.

Mr. Hiroki Hamamoto

He received his B. Eng. and M. Eng. from Hiroshima university, Japan, in 2018 and 2020, respectively. Currently, he works in the Technical Development Department, The Japan Steel Works, LTD., Japan.

Visiting Associate Professor Takashi Ochiiwa

He received his B. Eng. and M. Eng. from Okayama Prefectural University in Japan, in 1999 and 2001, respectively. He is currently a Group Manager of Numerical Analysis and Control Engineering Group, The Japan Steel Works, LTD., Japan. He is currently a Visiting Associate Professor with the Department of Graduate School of Advanced Science and Engineering, Hiroshima University, Japan.

Visiting Professor Hideki Tomiyama

He received his B. Eng., M. Eng. and D. Eng. from Kyushu Institute of Technology in Japan, in 1995, 1997 and 2001, respectively. He is currently a General Manager of Technical Development Department, The Japan Steel Works, LTD., Japan. He is currently a Visiting Professor with the Department of Graduate School of Advanced Science and Engineering, Hiroshima University, Japan. He is currently a Vice President of the JSPP (Japan Society of Polymer Processing).

Development of IoT self-tuning control device using Wi-Fi

Shinichi Imai
Graduate school, Tokyo Gakugei University,
4-1-1, Nukuikita-machi, Koganei, Tokyo, 184-8501, Japan.

E-mail: shimai@u-gakugei.ac.jp
http://www.u-gakugei.ac.jp/

Abstract

In this paper, development of IoT control device using Wi-Fi. In recent years, IoT has been attracting attention, and there are growing expectations in the industrial world for the utilization of data obtained from many sensors. These data are stored in databases in real time through communication between sensors and the cloud and communicating with the cloud. Meanwhile, digital controllers are widely used in the process industry as general-purpose controllers. However, it is difficult to incorporate AI, machine learning, and databases into general-purpose controllers due to data memory limitations. Therefore, in this paper, we develop an IoT self-tuning controller using Wi-Fi. As a result of experiments, the controller and computer were connected via Wi-Fi, self-tuning was performed on the computer side, and the calculated PID gains could be sent to the controller to achieve control.

Keywords: Wi-Fi, IoT, Control.

1. Introduction

In recent years, the utilization of advanced technologies such as IoT (Internet of Thinks) and AI (Artificial Intelligence) has been remarkable. IoT [1],[2],[3],[4],[5] is attracting attention, and expectations are rising in the industrial world for the utilization of data obtained from the many sensors installed in the system. These data are accumulated in a database in real time through communication between the sensor and the cloud. On the other hand, in the process industry, digital controllers are widely used for general purpose control. Currently, to obtain the desired control performance, it is necessary to appropriately calculate the control parameters according to the characteristics of the controlled object, but most of the actual controlled objects have nonlinearity or fixed control. Therefore, it often happens that sufficient control performance cannot be obtained with that parameter. Therefore, self-tuning methods and data-driven control methods have been proposed, in which control parameters are sequentially changed according to control characteristics. However, because data-driven control determines control parameters by accumulating a large amount of data in a database, it is difficult to implement with the capacity of conventional controllers. Also, from an existing controller. Therefore, in this paper, we will develop an IoT self-tuning control device using Wi-Fi. Specifically, a microcomputer called ESP-WROOM-02 is used instead of the conventional controller. ESP-WROOM-02 is very compact and can replace traditional controllers. It also has built-in Wi-Fi and Bluetooth. Connect the ESP-WROOM-02 and a computer via Wi-Fi, perform calculations for determining control parameters on the computer side, and send the calculated control parameters to the ESP-WROOM-02 controller. This can avoid controller data memory problems.

2. ESP-WROOM-02

ESP-WROOM-02 incorporates Espressif ESP8266EX.

The ESP8266EX integrates a Tensilica L106 32-bit RISC processor for ultra-low power consumption and reaches a maximum clock speed of 160MHz. A real-time operating system (RTOS) and Wi-Fi stack allows 80% of the processing power to be used for programming and developing user applications. The module also integrates antenna switches, RF baluns, power amplifiers, low noise receive amplifiers, filters, and power management modules at the SoC level, resulting in a small size for easy integration into space-constrained devices. increase. The external dimensions of the module are 18 x 20 mm. The type of flash used in this module is SPI flash with package size SOP8-150mil. The antenna applied to this module is a 3DBi PCB onboard antenna. Fig.1 shows ESP-WROOM-02. Table 1 shows the specifications of the ESP-WROOM-02.

3. IoT system

This IoT system consists of a "sensing system" that acquires numerical values with sensors and sends them to a computer, and an "adaptive system" that performs self-tuning based on the input and output data of the sensing system. The calculated parameters are sent to the control controller. Data is transmitted over Wi-Fi between the three systems. ESP-WROOM-02 is used for data transfer. Fig.2 shows a conceptual diagram of the IoT system.

Used by connecting to Arduino to communicate with ESP-WROOM-02. Fig.3 shows the circuit configuration. Arduino works with 5V. ESP-WROOM-02 cannot be directly connected to Arduino because it operates at 3.3V. Therefore, different signal levels are switched by attaching a level conversion IC. The current consumption of ESP-WROOM-02 is about 80mA. Arduino 3.3V terminal can only output 50mA and cannot be used. Therefore, it is necessary to prepare a 5V power supply and a 3.3V power supply in separate systems.

Use the Arduino input port to capture the inputs, outputs, and deviations of your experimental setup. It also sends PI parameters to the controller through the output port. The controller only updates PI parameters. Calculation of PI parameters is performed on a computer via Wi-Fi.

Table 1 Specifications of ESP-WROOM-02.

Supply voltage	3.0-3.6V
Current Consumption	Average 80mA
Supported WiFi Protocols	802.11b/g/n (2.4GHz)
Size	18mm×20mm×3mm
Wi-Fi mode	station softAP SoftAP+station
Security	WPA/WPA2
Encryption	WEP/TKIP/AES

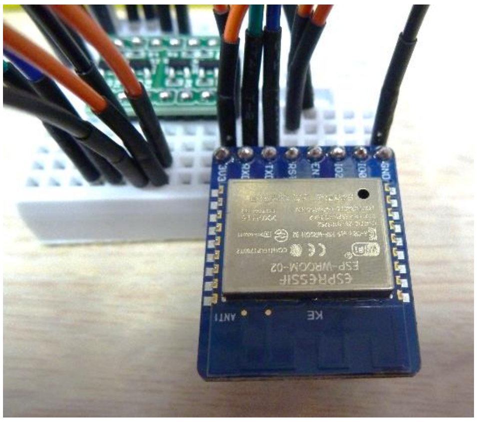


Fig.1 ESP-WROOM-02

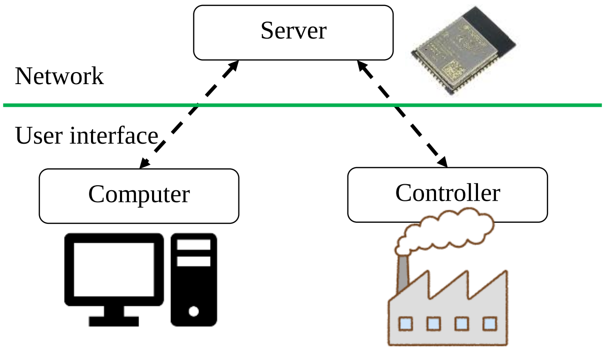


Fig.2 IoT IoT system conceptual diagram

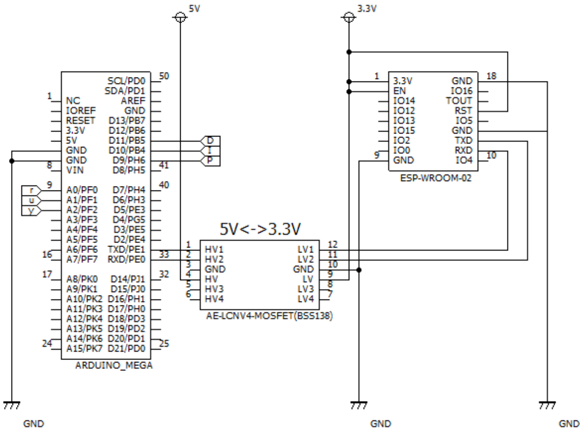

5V
3.3V
ARDUINO_MEGA
5V<->3.3V
AE-LCNV4-MOSFET(BSS138)
ESP-WROOM-02
GND
GND
GND

Fig.3 Circuit configuration

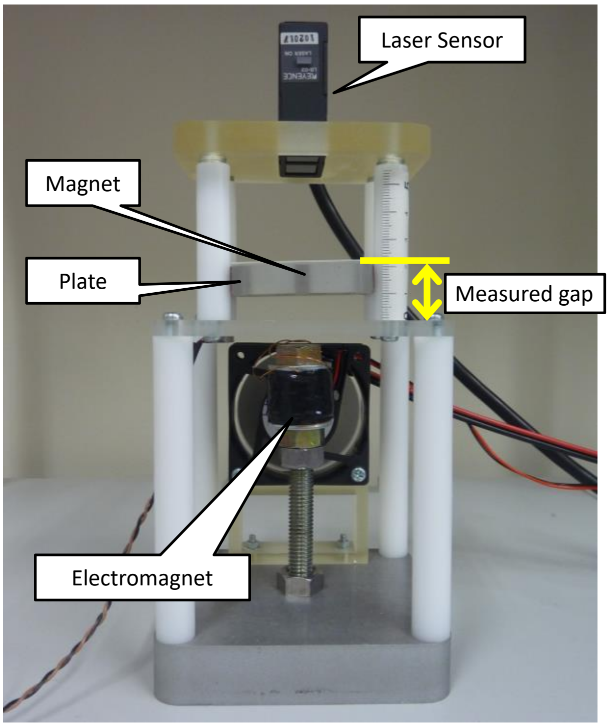

Laser Sensor
Magnet
Plate
Measured gap
Electromagnet

Fig.4 Experimental device

4. Simulation

A self-tuning control is used for the simulation. In addition, a magnetic levitation device is used in the experimental device. Fig.4 shows the magnetic levitation device. A magnetic levitation device levitates a magnetic body by controlling the current flowing through an electromagnet (coil) or the voltage across the coil. In the following, the magnetic body to be levitated is called the levitation body. This experimental device consists of a magnetic levitation device, an electromagnet amplifier, and a control device. Controller uses Arduino to send input/output data to another computer through ESP-WROOM-02. In addition, the displacement of the levitation body is measured using a laser displacement sensor. Fig.5 shows a block diagram of the magnetic levitation device.

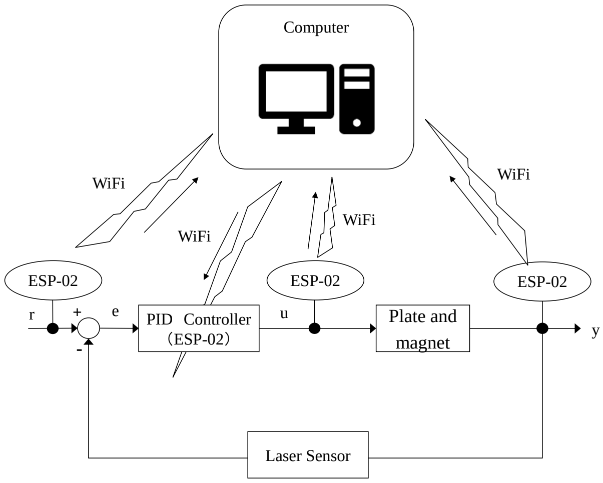

Computer
WiFi
WiFi
WiFi
WiFi
ESP-02
ESP-02
ESP-02
r
+
-
e
PID Controller
(ESP-02)
u
Plate and magnet
y
Laser Sensor

Fig.5 Block diagram

Control input u(t) is the voltage applied to the coil, and control output y(t) is the gap length (the distance between the sensor and the upper surface of the levitation body). In addition, the coil voltage is controlled by a PWM (Pulse Width Modulation) signal with a duty ratio corresponding to the control input. Therefore, the control input u(t) in this experiment is assumed to be the duty ratio (0 to 100%) of the PWM signal. Also, this experimental device can only start from u(0)=100 due to its specifications. In addition, the vibration of about 17.6m/s^2 is given to the levitation body to prevent the levitation body from being caught by the effect of friction.

First, the target value r(t) is given as follows.

$$r(t) = \begin{cases} 11.5\ (0 \leq t < 100) \\ 8.0\ (100 \leq t < 200) \\ 10.0\ (200 \leq t < 300) \\ 7.0\ (300 \leq t < 400) \end{cases}$$

The parameters in the design polynomial and $P(z^{-1})$ re sigma=1, delta=0, and the sampling interval is Ts=0.1[sec].

At this time, $P(z^{-1})$ is obtained as the following equation.

$$P(z^{-1}) = 1 - 1.6375z^{-1} + 0.6703z^{-2}$$

Furthermore, let n_u=1 and k_m=0.

The experimental results are shown in Fig.6. Fig. 7 shows the temporal change of the PI parameter.

From the experimental results, the PI parameters change appropriately according to the characteristics of the system. From this, communication is properly performed through Wi-Fi.

4.1. *Conclusion*

In this paper, we developed an IoT self-tuning control device using Wi-Fi. As a result, he was able to connect the controller and computer with his IoT technology, perform the PID tuning calculations on the computer side, and send his calculated PID gains to the controller. The effect was verified by experiments. In the future, we plan to experiment with delays. We also plan to adapt it to systems with heavy memory loads.

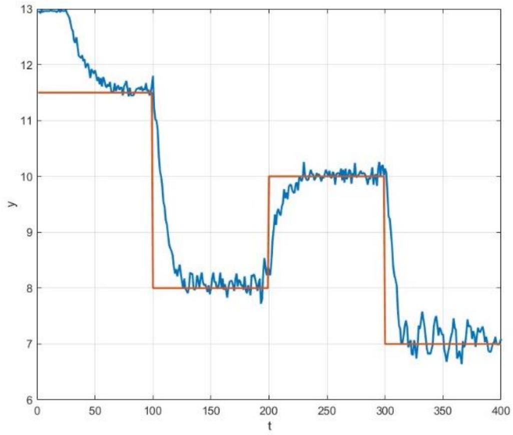


Fig.6 Experimental result

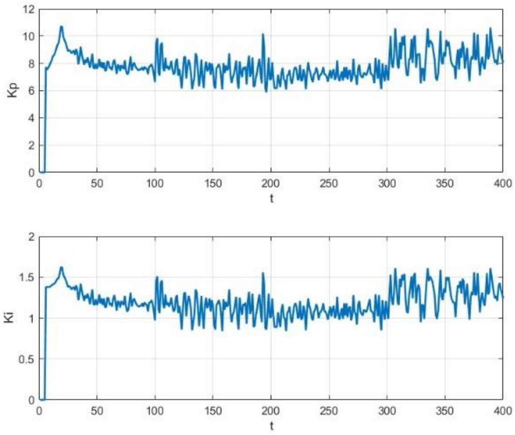


Fig.7 Changes in PI parameters

References

1. S. N. Swamy and S. R. Kota, "An Empirical Study on System Level Aspects of Internet of Things (IoT)," IEEE Access, pp188082-188134, October 2020.
2. T. Lan, J. Duan, B. Zhang, D. Shi, and Z. Wang, "AI-Based Autonomous Line Flow Control via Topology Adjustment for Maximizing Time-Series ATCs," 2020 IEEE Power & Energy Society General Meeting (PESGM), August 2020.
3. A. K. Gupta and R. Johari, "IOT based Electrical Device Surveillance and Control System," 2019 4th International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU), April 2019.
4. Kuan-Chu Lu, I-Hsien Liu, Jia-Wei Liao, Shao-Chun Wu, Zong-Chao Liu, Jung-Shian Li, Chu-Fen Li, "Evaluation and Build to honeypot System about SCADA Security for Large-Scale IoT Devices", Journal of Robotics, Networking and Artificial Life, Vol.6, Issue.3 pp.157-161 (2019)
5. Seyed Mahmood Hashemi, Jingsha He, "LA-Based Approach for IoT Security", Journal of Robotics, Networking and Artificial Life, Vol.3 Issue.4, pp.240-248 (2013)

Authors Introduction

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.

Consensus Control for Dual-rate Multi-agent Systems

Takaya Tanaka, Natsuki Kawaguchi, Takao Sato
Graduate School of Engineering, University of Hyogo, 2167, Shosha, Himeji, Hyogo 671-2280, Japan
E-mail: {Kawaguchi, tsato}@eng.u-hyogo.ac.jp

Abstract

The present study discusses the consensus control multi-agent systems. In such systems, the consensus is achieved through the exchange of information between neighboring agents. In order for mobile systems to consensus with each other, they must be designed with power consumption in mind. Therefore, since communication consumes power, it is important that the communication interval be as long as possible. The present study proposes a design methodology for the consensus achievement of dual-rate multi-agent systems, where the update period of agents is shorter than the communication period.

Keywords: Multi-agent System, Dual-rate, Consensus Control

1. Introduction

With the development of information and communication technology, Internet of Things (IoT), which uses communication to connect devices, is attracting attention. The control method for a group of devices connected to a network can be a centralized management style using a central control terminal or an autonomous decentralized control method. This study investigates the consensus control of multi-agent systems [1] are designed in a dual-rate system, where the communication period differs from the update interval. Networked predictive controller [2], quantized control [3], strict feedback control in a sampled-data system [4] have been reported as studies of dual-rate multi-agent systems. Network topology design [5], state-dependent graph Laplacian [6], and pinning consensus [7] are studies of convergence of multi-agent systems. The present study examines the convergence of dual-rate multi-agent systems.

2. Problem Statement

All agents are assumed to be integrator systems as follows:

$$\dot{x}_i(t) = u_i(t) \tag{1}$$

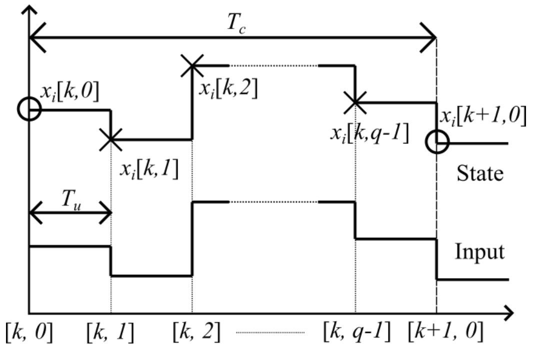


Fig. 1 Communication and input update periods in a dual-rate system; ○: communication with adjacent agents, ×: non-communication

$$(i = 1,2,\cdots,n)$$

where $x_i(t) \in \boldsymbol{R}$ and $u_i(t) \in \boldsymbol{R}$ are the state and input of agent i, respectively, and n is the number of agents. The goal of this study is to achieve the consensus that is defined as follow:

$$x_1(t) = x_2(t) = \cdots = x_n(t). \tag{2}$$

To this end, the control inputs are determined based on the states of its own agent and its neighboring agents that can communicate with its own agent. The following assumption is made about communication between agents.

Assumption 1

As shown in Fig. 1, the communication period of agents, $T_c \in \boldsymbol{R}^+$, is an integer multiple of $T_u \in \boldsymbol{R}^+$, the update period of control input, and

$$T_c = qT_u, \tag{3}$$

where $q \in \boldsymbol{N}^+$. Therefore, the control input is updated q times during the communication interval, and time $t \in \boldsymbol{R}^+$ is described using T_c and T_u as follows:

$$t = kT_c + mT_u, \tag{4}$$

where $k \in \boldsymbol{N}$ is the communication step, and $m \in \boldsymbol{N}$ is the update step ($0 \leq m < q$). Assuming $T_u = 1$, without loss of generality, the dynamics in the discrete time are given as follows:

$$x_i[k, m+1] = x_i[k, m] + u_i[k, m], \tag{5}$$

where $x_i[k, m]$ and $u_i[k, m]$ denote the state and input of the mth step from the communication of the kth step, respectively.

3. System Design

The input of agent i is updated by different control laws for single-rate and dual-rate systems. The control law of the single-rate system is:

$$u_i[k, m] = h \sum_{j=1}^{n} a_{ij}\left(x_j[k, 0] - x_i[k, m]\right), \tag{6}$$

and that of the dual-rate system is:

$$u_i[k, m] = h \sum_{j=1}^{n} a_{ij}\left(x_j[k, m] - x_i[k, m]\right), \tag{7}$$

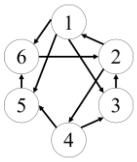


Fig. 2 Directed graph

where $h \in \boldsymbol{R}$ is a controller gain, and $a_{i,j}$ is the i, j element of an adjacency matrix A.

4. Numerical Example

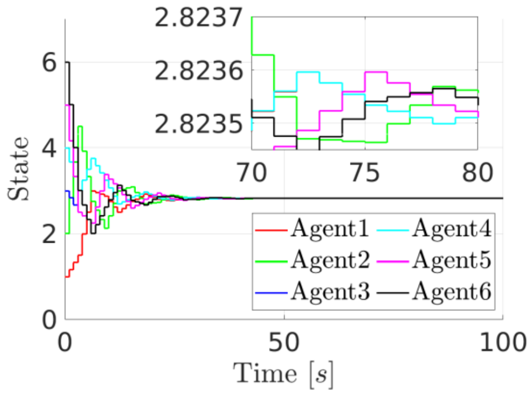
Consider a directed graph as shown in Fig. 2, where the number of agents is 6. The communication period T_c is 3, hence q is 3. The controller gain h is 0.1666 which stabilizes both the single-rate and dual-rate systems. The initial values of the agents are $x_1(0) = 1$, $x_2(0) = 2$, $x_3(0) = 3$, $x_4(0) = 4$, $x_5(0) = 5$, and $x_6(0) = 6$. Fig. 3 and Fig. 4 shows the simulation results of the single-rate and dual-rate systems, respectively. The consensus condition is defined as follows:

$$\max x_i(t) - \min x_i(t) \leq 10^{-4}, \tag{8}$$

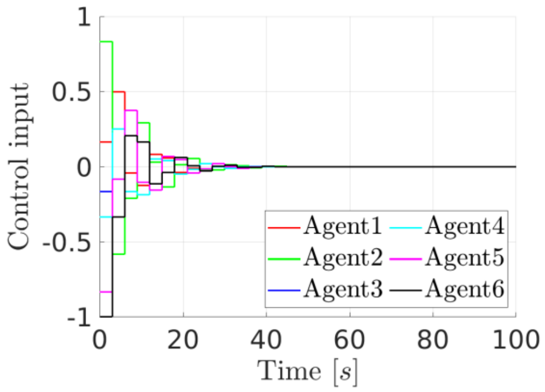
and the consensus time of the single-rate and dual-rate systems are 75s and 57s, respectively. Therefore, the performance of the dual-rate system is superior to that of the single-rate system.

5. Conclusion

The present study has proposed a design method of multi-agent systems in dual-rate system, where the communication period is longer than the input update period.

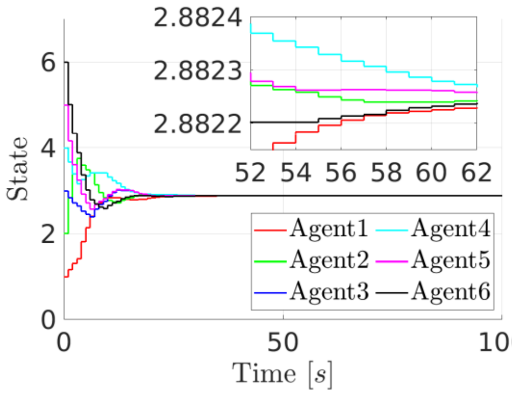


(a) State

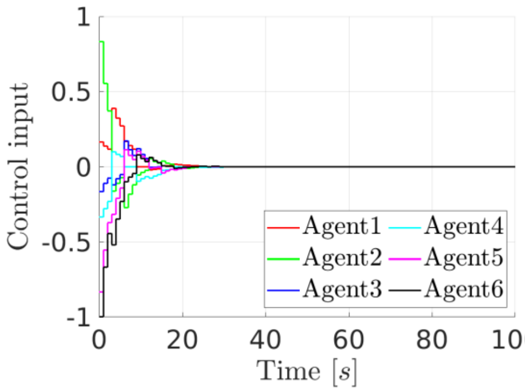


(b) Input

Fig. 4 Simulation results in the single-rate system



(a) State



(b) Input

Fig. 3 Simulation results in the dual-rate system

References

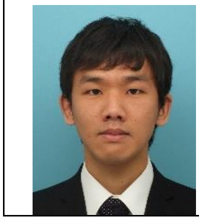
1. S. Azuma and M. Nagahara (Eds.): Control of Multi-agent Systems, CORONA Publishing (2015) (in Japanese).
2. C. J. Li, X. R. Yang, and J. Zhao: Multi-rate Design of Networked Predictive Controller and Its Application for Consensus of Networked Multi-agent Systems with Delays; Proceedings of the 38th Chinese Control Conference, pp. 5474–5479(2019).
3. T. Furusaka, T. Sato, N. Kawaguchi, N. Araki, and Y. Konishi: Consensus Control of Dual-Rate Multi-Agent Systems with Quantized Communication; IEEE Access, Vol. 48, No. 1, pp. 97557–97563(2020).
4. H. Katayama: Multi-rate Design of Consensus Controllers for Multi-rate Sampled-data Strict-Feedback Multi-agent Systems; IFAC-PapersOnLine, Vol. 48, No. 18, pp. 157–162(2015).
5. M. Rafiee and A. M. Bayen: Optimal Network Topology Design in Multi-Agent Systems for Efficient Average Consensus; 49th IEEE Conference on Decision and Control, pp. 3877–3883(2010).
6. Y. Kim and M. Mesbahi: On Maximizing the Second Smallest Eigenvalue of a State-dependent Graph Laplacian; 2005 American Control Conference, pp. 99–103(2005).
7. D. Xu, A. Sakaguchi, T. Ushio: Dynamic Pinning Consensus Control of Discrete-time Multi-agent Systems; Proceedings of the 37th Chinese Control Conference, pp. 6937–6942(2018).

Authors Introduction

Mr. Takaya Tanaka

He received his Bachelor of Engineering from the School of Engineering, University of Hyogo, Japan in 2021 Currently, he is a graduate school student in University of Hyogo, Japan. His research interest includes multi-agent systems, dual-rate systems.

Dr. Natsuki Kawaguchi



He received a D.Eng. degree from University of Hyogo in 2018. He is an assistant professor in the Graduate School of Engineering at University of Hyogo, Japan. His research interests are fault detection and fault tolerant control.

Prof. Takao Sato

He received B.Eng. and M.Eng. degrees from Okayama University in 1997 and 1999, respectively, and a D.Eng. degree from Okayama University in 2002. He is a professor in the Graduate School of Engineering at University of Hyogo. His research interests are PID control, mechanical systems and multi-rate control.

Data-driven Control Experiments of a Quadrotor Drone

Tomonori Yao, Natsuki Kawaguchi, Takao Sato
Graduate School of Engineering, University of Hyogo
2167, Shosha, Himeji, Hyogo 671-2280, Japan
E-mail: es22c037@guh.u-hyogo.ac.jp, {Kawaguchi, tsato}[AT]eng.u-hyogo.ac.jp

Abstract

Multi-rotor unmanned aerial vehicle (UAV)have various advantages due to their high expected performance. Most conventional model-based design methods require the dynamic characteristics of UAV, whereas data-driven design methods allow the controller to be designed directly from flight data without a model. This study reports on the creation and flight experiments of a quadrotor drone with the aim of preparing an environment for implementing data-driven design for UAV. In addition, the usefulness of the controller directly designed from flight data is also reported.

Keywords: Quadrotor Drone, Data-driven Control

1. Introduction

Drones [1], which are expected to play a variety of roles, have been involved in crashes, so further improving their safety is an issue. Since many drones use an internal controller to achieve the movements expected by the operator, the design of the controller is important for safety.

Controller design methods can be broadly divided into model-based design and data-driven design. Since model-based design is based on identified models, control performance depends on the accuracy of model identification. However, obtaining a model with high accuracy requires a large amount of effort. Drones, in particular, are not only prone to deterioration of components such as motors and batteries, but also frequently change their dynamic characteristics due to differences in payload. In contrast to model-based design, data-driven design does not require the model of a controlled process because the controller is designed directly from experimental data.

Virtual Reference Feedback Tuning (VRFT) [2] and Fictitious Reference Iterative Tuning (FRIT) [3] are typical methods for data-driven design. Using these methods, a controller can be designed from a set of input/output data, where there are two types of data used in the design: open-loop data and closed-loop data. Since unstable drones require closed-loop data-based design, an attitude control method using FRIT [4] has been proposed. However, the barometer built into the flight controller used in previous studies [4] cannot accurately measure altitude indoors. Therefore, in this study, a motion capture system will be introduced to accurately measure the altitude, and data-driven flight control design methods are discussed. Finally, the effectiveness of the data-driven method is confirmed through experiments.

2. Drone used in experiments

To allow for flexibility in changing controllers, a research drone [4], shown in Figure 1, was developed, where thelist of partes are summarized in Table 1. The drone is a quad-rotor type, and an ESC (Electronic Speed Controller) is required for each motor, which controls the motor speed by adjusting the current according to the PWM signal calculated by the flight controller. The reference values determined by the pilot are communicated to the flight controller through the radio. The control system developed is equipped with two XBee (wireless communication devices), one for receiving advanced information and the other for transmitting control data, and the two XBee-compatible devices are connected to a PC to send and receive data.

3. Altitude Measurement using an Optical Motion Capture System

Figure 2 shows a motion capture system for drone altitude measurement. In this system, the four cameras are placed in the four corners of the room, and the cameras are adjusted so that the drone is positioned near the center of the shooting range. Five reflective markers are attached to the drone to create a rigid body on the motion capture software. The position coordinates of the rigid bodies measured in real time are captured by software (MATLAB by MathWorks) and sent to the flight controller of the drone via wireless transmitter to control the altitude.

4. Data-driven Controller Design

The design objective is to improve control performance by tuning the controller directly from experimental data of an initial controller. Consider the model reference problem for the single-loop system in Figure 3. In the figure, $C(\Phi)$ denotes the controller, Φ the controller parameters, P the controlled process, r the reference input, u the input, and y the output, and e the error between r and y. M is the reference model, y_M is the reference model output, and G is the following closed-loop transfer function from r to y:

$$G(\Phi) = \frac{PC(\Phi)}{1 + PC(\Phi)}. \quad (1)$$

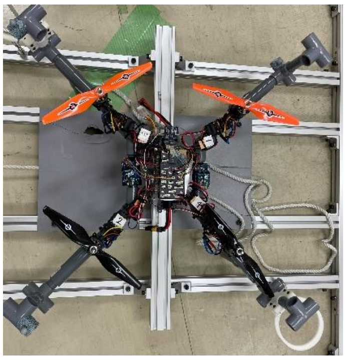


Fig. 1 Picture of the drone

Table 1 Done parts list

Parts name	Product name
Frame	RC Drone Frame Kits XL8 Carbon Fiber FPVRacing Frame Kit for RC Drone
Flight controller	Pixhawk4
Brushless motor	NTM Brushless Motor NTM Prop Drive Series 28-30 1200kv
Electronic speed controller (ESC)	FUTABA CORPORATION ESC Exclusively for brushless motors MC930A
Remote	FUTABA CORPORATION 8-Channel Digital Proportional R/C System T6K-V2
Lithium polymer battery	Zeee 11.1V 50C 3200mAh 3S Lipo Battery
Propeller	MASTER AIRSCREW MR Series 8×4.5 Prop Set Black
Wireless communication device	ZigBee wireless communication module XBee

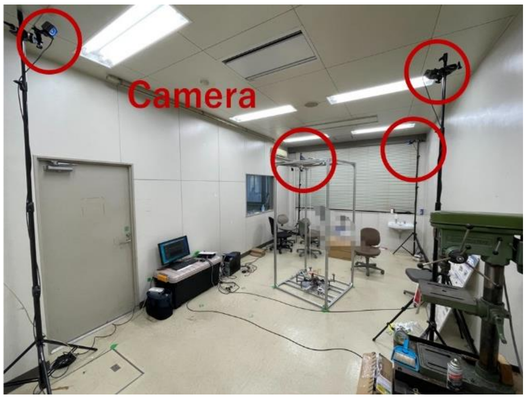


Fig. 2 Environment of experiments

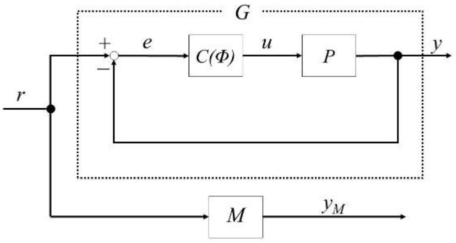


Fig. 3 Model reference problem

In general, the mathematical model of P is not known. To make the closed-loop system $G(\Phi)$ match the reference model M, we define the evaluation function $J(\Phi)$ as follows:

$$J(\Phi) = \|M - G(\Phi)\|_2^2. \qquad (2)$$

Therefore, the controller parameter Φ is determined based on the minimization of this evaluation function. However, since $J(\Phi)$ contains the unknown P, which is it cannot be solved. In the present study, instead of $J(\Phi)$, the following function is used for determining the controller parameters:

$$J_F(\Phi) = \|y_0 - M\tilde{r}(\Phi)\|_2^2, \qquad (3)$$

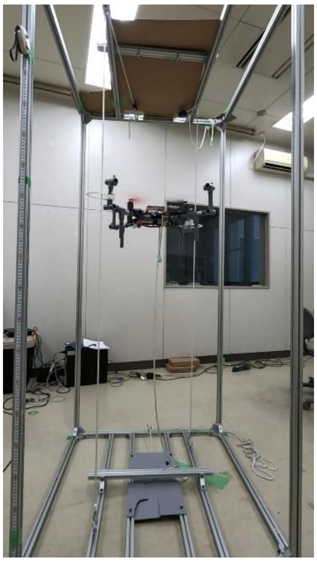
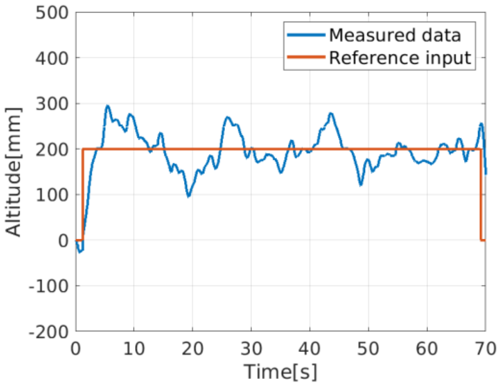


Fig. 4 Scene of a flight experiment

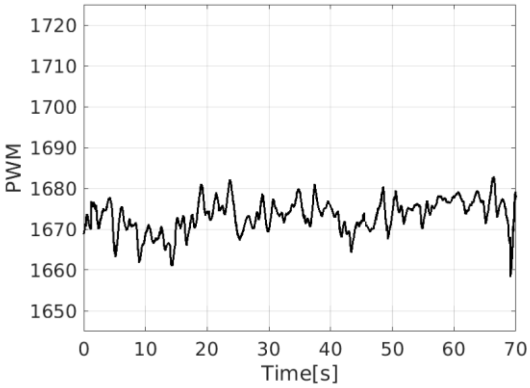
where $\tilde{r}(\Phi)$ is a fictitious reference input given as follows:

$$\tilde{r}(\Phi) = \frac{u_0}{C(\Phi)} + y_0. \qquad (4)$$

u_0 and y_0 are the input and output data collected from an experiment using initial controller parameters. The controller parameters are obtained by solving the above evaluation function.



(a) Output



(b) Input

Fig. 5 Flight result before tuning

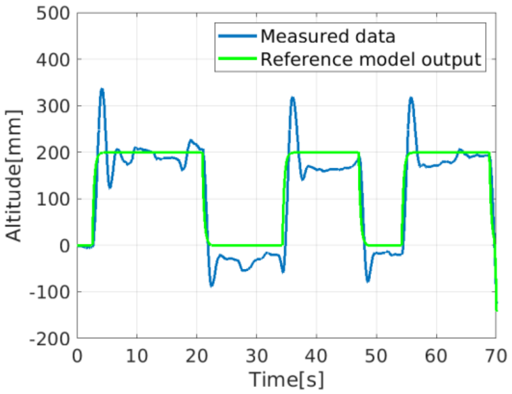
5. Experiment

5.1. Experimental conditions

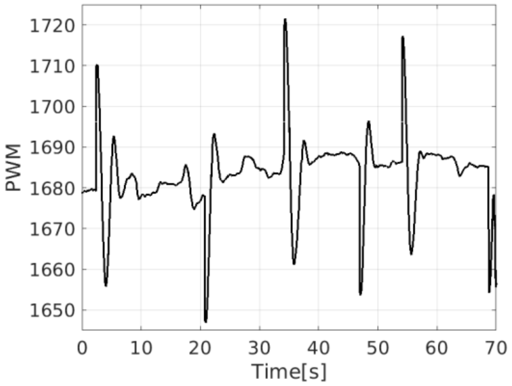
Drone flight experiments were conducted as shown in Figure 4. The altitude of drone is controlled by PI-D control, and the control and sampling periods are 16 [ms]. The drone was allowed to hover at an altitude of approximately 1.25 [m] above the ground, and the altitude at the hovering point was used as the origin.

Table 2 PID gains for experiments

Parameters	P	I	D
Initial	0.05000	4.000×10^{-4}	5.000
Tuned using Eq. (5)	0.1534	1.417×10^{-4}	0.9548
Tuned using Eq. (6)	0.1090	1.866×10^{-4}	3.112



(a) Output



(b) Input

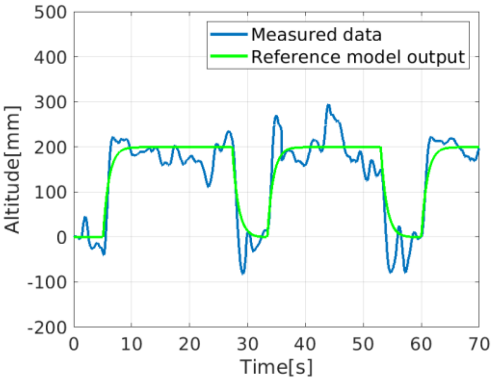
Fig. 7 Flight result after tuning (fast reference model (5))

5.2. Initial experiment and parameter tuning

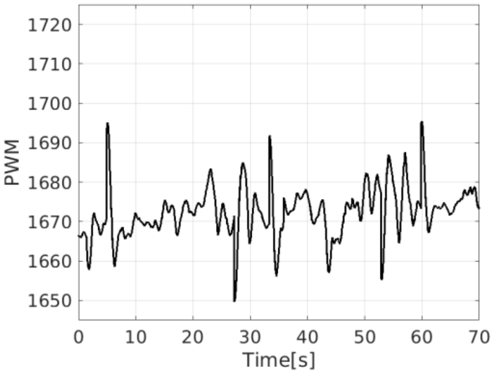
An experiment was conducted with initial parameters shown in Table 2 with a reference altitude of 200 [mm] to obtain data used for controller tuning. The control result is shown in Fig. 5 here the red and blue lines show the reference input and the measured altitude, respectively. Based on the collected data, the PID parameters are tuned using FRIT [3] for the fast and slow reference models (5) and (6), respectively, where CMA-ES [5] is used for optimizing the objective function (3).

$$M(s) = \frac{1}{0.3s + 1} e^{-0.096s} \tag{5}$$

$$M(s) = \frac{1}{0.5s + 1} e^{-0.096s} \tag{6}$$



(a) Output



(b) Input

Fig. 6 Flight result after tuning (slow reference model (6))

The obtained PID parameters are shown in Table 2.

5.3. Experimental results using tuned parameters

The experimental results using the tuned parameters are shown in Fig. 6 and Fig. 7, where the green line shows the reference model output. The control result based on the fast reference model shows that the altitude trajectory tracks the reference model output well in steady state. However, each time the reference input switches, an overshoot occurs. On the other hand, the experimental result for the case based on the slow reference model shows that the overshoot is suppressed, while the altitude oscillates around the reference input. These results show that there is a trade-off between the speed of tracking to the reference input and vibration suppression.

6. Conclusion

This study has reported on the implementation of drone altitude control using a motion capture system. Since the controller for operating the drone was designed based on flight experimental data, the controller is designed directly form the data without using the model of the drone. Finally, the flight experiments confirmed that control performance is improved using the tuned controller parameters.

References

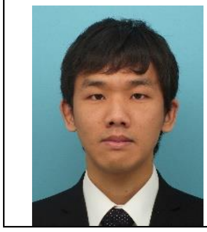
1. K. Nonami: “Introduction to Drone Engineering –Modeling and Control-,” Corona Publishing Co.,Ltd. (2020)
2. M. C. Campi, A. Lecchini and S. M. Savaresi: “Virtual Reference Feedback Tuning (VRFT): A direct method for the design of feedback controllers,” Automatica, Vol. 38, pp. 1337–1346 (2002)
3. S. Soma, O. Kaneko, T. Fuji: ”A New Approach to Parameter Tuning of Controllers by Using One-Shot Experimental Data -A Proposal of Fictitious Reference Iterative Tuning” Vol. 17, No. 12, pp. 528–536 (2004)
4. T. Yao, K. Watanabe, Y. Sakai, N. Kawaguchi, N. Araki, T. Sato: “Flight Control of a Drone Using Data-driven Design,” The Papers of Technical Meeting on “Control”, IEE Japan, pp. 85–88 (2022)
5. N. Hansen: “CMA-ES Source Code,” http://cma.gforge.inria.fr/cmaes_sourcecode_page.html Access 2021.09.1

Authors Introduction

Mr. Tomonori Yao

He received his Bachelor of Engineering from the XSchool of Engineering, University of Hyogo, Japan in 2022 Currently, he is a graduate school student in University of Hyogo, Japan. His research interest includes data-driven control of UAV

Dr. Natsuki Kawaguchi



He received a D.Eng. degree from University of Hyogo in 2018. He is an assistant professor in the Graduate School of Engineering at University of Hyogo, Japan. His research interests are fault detection and fault tolerant control.

Prof. Takao Sato

He received B.Eng. and M.Eng. degrees from Okayama University in 1997 and 1999, respectively, and a D.Eng. degree from Okayama University in 2002. He is a professor in the Graduate School of Engineering at University of Hyogo. His research interests are PID control, mechanical systems and multi-rate control.

Design of Data-driven Multi-agent Systems

Kenta Nagao, Natsuki Kawaguchi, Takao Sato
Graduate School of Engineering, University of Hyogo
2167, Shosha, Himeji, Hyogo 671-2280, Japan
E-mail: {Kawaguchi, tsato}[AT]eng.u-hyogo.ac.jp

Abstract

This study discusses the consensus control of multi-agent systems. The consensus can be achieved when the closed-loop system of multi-agent systems is stable. In conventional model-based methods, since the controller is designed based on the dynamic characteristics of the agents, models of the agents must be used. On the other hand, this study examines data-driven design of multi-agent systems. In the proposed method, the controller of a multi-agent system is designed directly from the control data, where the controller structure is fixed. The usefulness of the proposed method is shown through numerical examples.

Keywords: Multi-agent Systems, Data-driven Control

1. Introduction

In multi-agent systems [1], multiple agents interact with each other and act autonomously to achieve global objectives. Many design methods have been proposed for multi-agent systems, but most of them are model-based design, which requires process models. Therefore, model-free adaptive control has been proposed for designing multi-agent systems with unknown dynamics [2]. In this method, the control system is designed by identifying an unknown model. On the other hand, a data-driven design method for designing a controller directly from data has also been proposed [3]. The data-driven design approach eliminates the need for process models that are required in the model-based approach. Therefore, this study examines a data-driven design method for multi-agent systems that does not require models of agents.

2. Data-driven Design

Consider the model reference problem for agent i shown in Fig. 1, where $P_i(z^{-1})$, $C_i(z^{-1}, \theta_i)$, and $M_i(z^{-1})$ are the process model, controller, and reference model, respectively, $u_i(k)$ and $y_i(k)$ are the process input and process output, respectively, and $r_i(k)$ and $y_{Mi}(k)$ are the reference input and reference model output, respectively. The reference input consists of $r_{ij}(k)$ $(j = 1, \cdots, d_i)$ given by adjacent agents, where d_i denotes the number of adjacent agents of agent i.

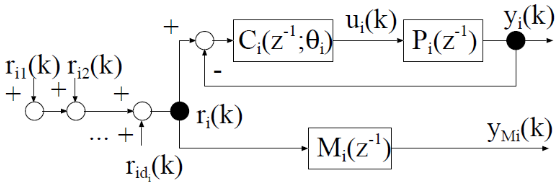


Fig. 1 Block diagram of the model reference problem of a multi-agent system

Assumption

The process model $P_i(z^{-1})$ is unknown.

The controller parameters θ_i is optimized by minimizing the following objective function:

$J_{iMR}(\theta_i) =$

$$\left\|\left(\frac{P_i(z^{-1})C_i(z^{-1},\theta_i)}{1+d_iP_i(z^{-1})C_i(z^{-1},\theta_i)}-M_i(z^{-1})\right)W_i(z^{-1})\right\|_2^2, \quad (1)$$

where $W_i(z^{-1})$ denote a design parameter. Since the objective function involves an unknown process model, it cannot be minimized as is. Instead of the objective function, the controller parameters are determined based on the following function:

$$J_{iVR}(\theta_i) = \frac{1}{N}\sum_{k=1}^{N}\left(L_i(z^{-1})u_i(k) - C_i(z^{-1},\theta_i)L_i(z^{-1})e_i(k)\right)^2 \quad (2)$$

$$e_i(k) = \bar{r}_i(k) - d_iy_i(k) \quad (3)$$

$$\bar{r}_i(k) = \frac{1}{M_i(z^{-1})}y_i(k), \quad (4)$$

where $\bar{r}_i(k)$ is the virtual reference input and $L_i(z^{-1})$ is a filter to be designed. Since $J_{iVR}(\theta_i)$ is convex with respect to θ_i, an optimal solution can be obtained. However, if the equivalent of $J_{iMR}(\theta_i)$ and $J_{iVR}(\theta_i)$ is not guaranteed, the obtained solution may not minimize $J_{iMR}(\theta_i)$. The problem is resolved by using $L_i(z^{-1})$. Comparing $J_{iMR}(\theta_i)$ and $J_{iVR}(\theta_i)$ in the frequency domain, these are equivalent when the next conditions is satisfied:

$$|L_i|^2 = \frac{|M_i|^2|W_i|^2}{|1+d_iP_iC_i(\theta_i)|^2}\frac{1}{\Phi_{u_i}}, \quad (5)$$

where $z^{-1} = e^{-j\omega}$ is omitted. It is assumed that $|1 + d_iP_iC_i(\theta_i)|^2 \cong |1 + d_iP_iC_{i0}|$ when $J_{iMR}(\theta_i)$ is minimized by θ_i, where C_{i0} is an ideal controller that satisfies the next equation:

$$M_i = \frac{P_iC_{i0}}{1+d_iP_iC_{i0}}. \quad (6)$$

This equation is rewritten as follows:

$$1 - d_iM_i = \frac{1}{1+d_iPC_{i0}}. \quad (7)$$

As a result, the filter is designed so as to satisfy the next equation:

$$|L_i|^2 = |1-d_iM_i|^2|M_i|^2|W_i|^2\Phi_{u_i}^{-1}. \quad (8)$$

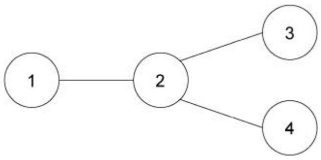


Fig. 2 Graph structure of a multi-agent system

3. Simulation

The graph structure an undirected graph as shown in Fig. 2. The dynamics of the agents are shown as follows:

$$P_1(s) = \frac{3}{s^2+s+2} \quad (9)$$

$$P_2(s) = \frac{3.5}{s^2+1.5s+2.5} \quad (10)$$

$$P_3(s) = \frac{4}{s^2+2s+3} \quad (11)$$

$$P_4(s) = \frac{4.5}{s^2+2.5s+3.5}. \quad (12)$$

The discrete-time period T_s is 1[s], and the control law of agent i is given as follows:

$$u_i(k) = \left(K_{Pi} + K_{Ii}\frac{T_s}{1-z^{-1}} + K_{Di}\frac{1-z^{-1}}{T_s}\right)e_i(k) \quad (13)$$

$$e_i(k) = r_i(k) - y_i(k) \quad (14)$$

$$(i = 1,\cdots,4),$$

where K_{Pi}, K_{Ii}, and K_{Di} are proportional, integral, and derivative gains, respectively and are determined directly from input/output data. Fig. 3 shows the response of agent 1 when white Gaussian noise with variance 1 is applied. White noise is also applied to other agents to obtain response data. Based on the corrected data, the controller parameters are determined by minimizing eq. (2). The obtained controller parameters are shown in Table 1, where the reference model is designed as follows:

$$M_i(s) = \frac{1}{d_i}M(s) \quad (15)$$

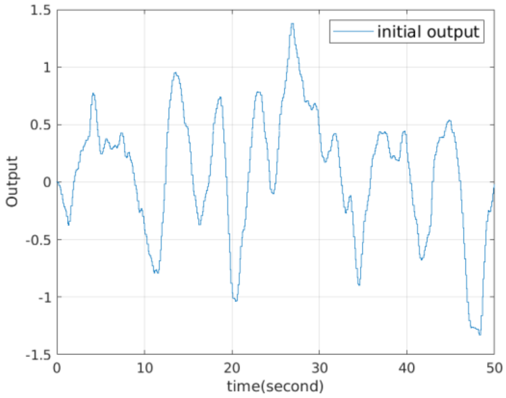
$$M(s) = \frac{1}{s+1}, \qquad (16)$$

where $d_1 = d_3 = d_4 = 1$ and $d_2 = 3$.

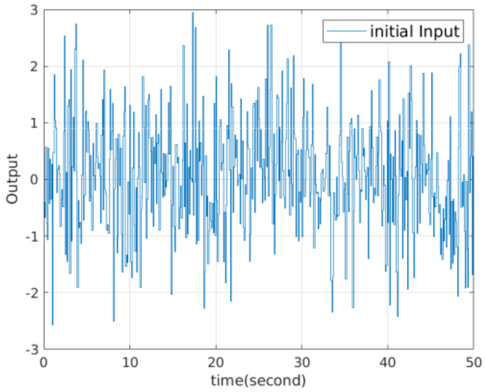
Table 1 PID parameters

	K_{Pi}	K_{Ii}	K_{Di}
θ_1	0.2265	0.6363	0.3204
θ_2	0.1045	0.2271	0.0916
θ_3	0.3727	0.7162	0.2417
θ_4	0.4412	0.7454	0.2154

Fig. 4 shows the simulation result of consensus control using the obtained PID parameters. It can be seen that all agents stably converge to a consensus value.



(a) Output



(b) Input

Fig. 3 Trajectories of initial input/output data

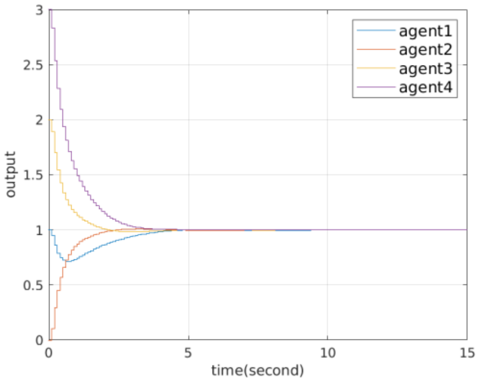


Fig. 4 Trajectories of consensus control

4. Conclusion

The present study has proposed a data-driven design for multi-agent systems. Therefore, even when the dynamic characteristics of agents are unknown, the controller parameters are determined directly from input/output data.

References

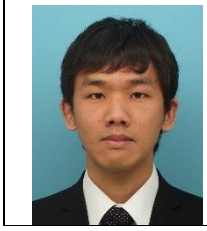
1. S. Azuma and M. Nagahara (Eds.): Control of Multi-agent Systems, CORONA Publishing (2015) (in Japanese).
2. X. Bu, Z. Hou, and H. Zhang: Data-driven multiagent systems consensus tracking using model free adaptive control, IEEE Trans. Neural Netw. Learn. Syst., 2018, 29(5): 1514-1524.
3. M. C. Campi, A. Lecchini and S. M. Savaresi: Virtual Reference Feedback Tuning (VRFT): direct method for the design of feedback controllers, Automatica, 2002, 38(8): 1337-1346.

Authors Introduction

Mr. Kenta Nagao

He received his Bachelor of Engineering from the School of Engineering, University of Hyogo, Japan in 2022 Currently, he is a graduate school student in University of Hyogo, Japan. His research interest includes data-driven control and multi-agent systems.

Dr. Natsuki Kawaguchi



He received a D.Eng. degree from University of Hyogo in 2018. He is an assistant professor in the Graduate School of Engineering at University of Hyogo, Japan. His research interests are fault detection and fault tolerant control.

Prof. Takao Sato

He received B.Eng. and M.Eng. degrees from Okayama University in 1997 and 1999, respectively, and a D.Eng. degree from Okayama University in 2002. He is a professor in the Graduate School of Engineering at University of Hyogo. His research interests are PID control, mechanical systems and multi-rate control.

A Self-triggering Control Based on Adaptive Dynamic Programming for Nonzero-sum Game Systems

Yibo Shi and Chaoli Wang
Department of Control Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China

E-mail: clcwang@usst.edu.cn
https://www.usst.edu.cn/main.htm

Abstract

Recently, for the optimal control problem of nonzero-sum game systems, although it is discussed that these methods are event-triggered, it is still necessary to continuously monitor measurement errors during execution, which is difficult to achieve by hardware. In order to avoid continuous detection measurement errors, a self-triggered control based on adaptive dynamic programming is proposed to solve the optimal control problem for continuous-time nonlinear nonzero-sum game systems with unknown drift dynamics. Firstly, the principle of IRL method is used to avoid the requirement of system drift dynamics in the controller design. Then, to approximate the Nash equilibrium solution, a critic neural network is used to estimate the value function. Furthermore, a self-triggered adaptive control scheme is proposed according to Lyapunov theory to ensure the uniform ultimate boundedness (UUB) of the closed-loop system state. The self-triggered control obtained in this paper can calculate the next trigger point by the information of the current trigger moment.

Keywords: Nonzero-sum games, Integral reinforcement learning, Optimal control, Neural network, Self-triggered

1. Introduction

In recent years, the event-triggered control has attracted many researchers because of its ability to reduce transmission load and computation cost. Unlike time-triggered systems, the control inputs of event-triggered systems are updated only at trigger times determined by appropriately designed trigger conditions. In this way, event-triggered control can reduce network bandwidth and computational burden. For multi-player NZS games, event-triggered ADP has become the most common method used to approximate the control input of each player in [1-4]. Mu, Wang, Ni[1] proposed a dynamic event-triggered method for solving the optimal control problem of a fully known continuous-time nonzero-sum game system. Su, Zhang, Liang, etc. [2] used identifier-critic NN to solve continuous time for partially unknown NZS games. Su, Zhang, Sun, etc.[3] used IRL to solve the optimal control for the partially unknown NZS game. Compared with [2], the introduction of identifier NN was avoided, i.e., the identification error was avoided. Zhao, Sun, Wang, etc.[4] proposed an event-triggered ADP is proposed for NZS games of continuous-time nonlinear systems with completely unknown system dynamics. In order to determine whether event-triggered conditions are met, these papers need a general event-triggered scheme to continuously monitor measurement errors. However, it is hard to realize through the hardware. It may weaken its feasibility of physical implementation.

To overcome such an issue, the self-triggered scheme is presented. By the designed triggered condition, the next trigger time instant can be computed by the current one. Moreover, the self-triggered scheme can also save network resources and reduce the communication burden, the same as the event-triggered scheme. To the best of our knowledge, no similar result has been reported in the literature. Based on the above motivation, we focus on a self-triggered nonzero-sum game system based optimal control problem, which avoids continuous monitoring measurement errors.

2. Preliminaries and Problem Formulation

Consider the general N -player NZS differential games[5]

$$\dot{x}(t) = f(x(t)) + \sum_{j=1}^{N} g_j(x(t))u_j(t) \qquad (1)$$

where $x \in \mathbb{R}^n$ is the system state, $u_j \in \mathbb{R}^{m_j}$ is the control input for player j. $f(x) \in \mathbb{R}^n$ and $g_j(x) \in \mathbb{R}^{n \times m_j}$ are the unknown drift dynamics and the known input dynamics of the system, respectively.

Assumption 1.[6] $f(x)$ and $g_j(x)$ are Lipschitz continuous on a compact set $\Omega \subset \mathbb{R}^n$ with $f(0) = 0$, $f(\cdot) \le b_f \| x \|$, and $\| g_j(x) \| \le b_{g,j}$, where b_f and $b_{g,j}$ are positive constants.

Define the cost function of system (1) as

$$J_i(x(t),u_i,u_{-i}) = \int_t^{\infty} \left(x^{T}(\tau) Q_i x(\tau) + \sum_{j=1}^{N} u_i^{T}(\tau) R_{ij} u_j(\tau) \right) d\tau \qquad (2)$$

where $u_{-i} = \{u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_N\}$, $Q_i = Q_i^{T} \ge 0$, $R_{ii} = R_{ii}^{T} > 0, R_{ij} = R_{ij}^{T} \ge 0$.

For a given set of control policies $\{u_i, u_{-i}\}$, the value function for player i can be written

$$V_i(x(t)) = \int_t^{\infty} \left(x^{T}(\tau) Q_i x(\tau) + \sum_{j=1}^{N} u_i^{T} R_{ij} u_j \right) d\tau, i \in \mathbb{N}. \qquad (3)$$

The optimal control problem is to design a set of control $\{u_1^*, u_2^*, \ldots, u_N^*\}$ to stabilize the closed-loop system while minimizing the value function (3). This control combination $\{u_1^*, u_2^*, \ldots, u_N^*\}$ corresponds to the Nash equilibrium of NZS games.

The value function $V_i(x(t))$ is assumed to be continuously differentiable. By differentiating V_i along the system trajectories (1), we can write Eq. (3) as:

$$0 = U_i\left(x(t),u_i,u_{-i}\right) + \nabla V_i^{T}(x(t)) \left(f(x(t)) + \sum_{j=1}^{N} g_j(x(t))u_j \right) \qquad (4)$$

where

$$U_i\big(x(t),u_i,u_{-i}\big) = x^{T}(t) Q_i x(t) + \sum_{j=1}^{N} u_i^{T} R_{ij} u_j,\ \nabla V_i = \frac{\partial V_i}{\partial x}.$$

The optimal value function V_i^* can be written as

$$V_i^*(x(t)) = \min_{u_i} \int_t^{\infty} \left(x(\tau)^{T} Q_i x(\tau) + \sum_{j=1}^{N} u_j^{T} R_{ij} u_j \right) d\tau \qquad (6)$$

Define the Hamilton-Jacobi-Bellman (HJB) equation as follow:

$$H_i(x, \nabla V_i^*(x(t)), u_i, u_{-i}) = U_i(x(t),u_i,u_{-i}) + (\nabla V_i^*(x(t)))^{T} \left(f(x(t)) + \sum_{j=1}^{N} g_j(x(t))u_j \right) \qquad (5)$$

Using the stationarity conditions $\partial H_i \big/ \partial u_i = 0$ the optimal control input for player i is

$$u_i^*(x(t)) = -\frac{1}{2} R_{ii}^{-1} g_i^{T}(x(t)) \nabla V_i^*(x(t)) \qquad (6)$$

The equivalent transformation of Eq. (6) is

$$V_i(x(t-\Delta t)) = V_i(x(t)) + \int_{t-\Delta t}^{t} U_i(x(\tau),u_i,u_{-i}) d\tau \qquad (7)$$

where $\Delta t > 0$ is a time interval.

According to (7), we have

$$V_i^*(x(t-\Delta t)) - V_i^*(x(t)) = \int_{t-\Delta t}^{t} U_i(x(\tau),u_i^*,u_{-i}^*) d\tau \qquad (8)$$

It's easy to see from Eq. (8) that there are no more unknown dynamics. Therefore, the identification process of unknown dynamics $f(x)$ is no longer needed, that is, identification error is avoided.

3. Design Self-triggered Control and Stability Analysis

3.1. *Design Self-triggered Control*

From the above time-triggered mechanism, it can be seen that the N-tuple control input $\{u_1, \ldots, u_N\}$ is a feedback form of system state updated at each sample time. In self-triggered control, the N-tuple control input $\{u_1, \ldots, u_N\}$ is updated only at the trigger time, and the next trigger time is determined by the current trigger time. In this case, a zero-order holder (ZOH) can be used to ensure that the control input is continuous at the trigger time. Define the triggering instant as τ_k, where $\{\tau_k\}_{k=0}^{\infty}$ is a monotonically increasing sequence of time instants with $\tau_0 = 0$. The trigger error is defined as

$$e_k(t) = \breve{x}_k - x(t), t \in [\tau_k, \tau_{k+1}) \qquad (9)$$

where $\breve{x}_k = x(\tau_k)$ is the trigger state.

In the framework of self-triggered, the optimal control input (6) can be written as

$$u_i^*(\breve{x}_k) = -\frac{1}{2} R_{ii}^{-1} g_i^{T}(\breve{x}_k) \nabla V_i^*(\breve{x}_k) \qquad (10)$$

where $\nabla V_i^*(\breve{x}_k) = \frac{\partial V_i^*}{\partial x} |_{x=\breve{x}_k}$.

The piecewise continuous control signal can be expressed by a ZOH

$$u_i^*(t) = \begin{cases} u_i^*(\breve{x}_k),\ t \in \left[\tau_k, \tau_{k+1}\right) \\ -\frac{1}{2} R_{ii}^{-1} g_i^{\bullet}(\breve{x}_{k+1}) \nabla V_i^*(\breve{x}_{k+1}),\ \ t = \tau_{k+1} \end{cases} \tag{11}$$

In the above analysis, We can get the solution of the optimal control (10) ultimately comes down to the solution of (8), which can be solved by using the critic NN. According to the Weierstrass high-order approximation theorem, we can get

$$\begin{aligned} V_i^*(x) &= \omega_i^{*\bullet} \phi_i(x) + \varepsilon_i(x), \\ \nabla V_i^*(x) &= \nabla \phi_i^{\bullet}(x) \omega_i^* + \nabla \varepsilon_i(x) \end{aligned} \tag{12}$$

where $\omega_i^* \in \square^{K_i}$ is the unknown ideal weight, $\phi_i : \square^n \to \square^{K_i}$ are linearly independent activation functions, K_i denotes the number of neurons, and ε_i is the approximation error.

Assumption 2.[7] (1) The approximation error $\varepsilon_i(x)$ and its gradient $\nabla \varepsilon_i(x)$ are bounded on Ω , i.e., $\| \varepsilon_i(x) \| \le b_{\varepsilon_i}$, and $\| \nabla \varepsilon_i(x) \| \le b_{\nabla \varepsilon_i}$, with $b_{\varepsilon_i}, b_{\nabla \varepsilon_i}$, being positive constants.

(2) The activation function $\phi_i(x)$ and its gradient $\nabla \phi_i(x)$ are bounded on Ω, i.e., $\| \phi_i(x) \| \le b_{\phi_i}$ and $\| \nabla \phi_i(x) \| \le b_{\nabla \phi_i}$, with $b_{\phi_i}, b_{\nabla \phi_i}$, being positive constants.

According to Eq. (8) and Eq. (12), it can be obtained

$$\begin{aligned} e_i(t) = {} & \omega_i^{*\bullet} \left[\phi_i(x(t)) - \phi_i(x(t-\Delta t))\right] \\ & + \int_{t-\Delta t}^{t} U_i(x(\tau), u_i^*(\breve{x}_k), u_{-i}^*(\breve{x}_k)) d\tau \end{aligned} \tag{13}$$

where $e_i(t) = \varepsilon_i(x(t-\Delta t)) - \varepsilon_i(x(t))$ is error from the NN approximation error. According to Assumption 2, $e_i(t)$ is bound on Ω , i.e., $\| e_i(t) \| \le b_{e,imax}$, where $b_{e,imax}$ is a positive constant.

Denote $\hat{\omega}_i$ as the estimations of ω_i^* . Then the value function can be approximated as

$$\hat{V}_i(x) = \hat{\omega}_i^{\bullet} \phi_i(x) \tag{14}$$

Based on (10), the approximate control inputs are

$$\hat{u}_i(\breve{x}_k) = -\frac{1}{2} R_{ii}^{-1} g_i^{\text{晻}}(\breve{x}_k) \nabla \phi_i(\breve{x}_k) \hat{\omega}_i \tag{15}$$

Using $\hat{V}_i(x)$ to replace $V_i^*(x)$ in Eq. (8). Therefore, the Bellman equation (8) can be written

$$\hat{e}_i(t) = \hat{\omega}_i^{\bullet} \rho_i(t) + s_i(t) \tag{16}$$

where

$$\begin{aligned} \rho_i(t) &= \phi_i(x(t)) - \phi_i(x(t-\Delta t)), \\ s_i(t) &= \int_{t-\Delta t}^{t} U_i(x(\tau), \hat{u}_i(\breve{x}_k), \hat{u}_{-i}(\breve{x}_k)) d\tau \end{aligned}$$

As can be seen from Eq. (16), adjusting $\hat{\omega}_i$ can directly affect $\hat{e}_i(t)$. Then the original problem of solving the value function is transformed into minimizing the error $\hat{e}_i(t)$ by adjusting $\hat{\omega}_i$. Consider the objective function

$$E_i(t) = \frac{1}{2} \hat{e}_i^{\bullet}(t) \hat{e}_i(t) \tag{17}$$

According to the gradient descent method, the update rule of $\hat{\omega}_i$ can be obtained as

$$\dot{\hat{\omega}}_i(t) = -\alpha_i \frac{\rho_i(t)}{\left(1 + \rho_i^{\bullet}(t) \rho_i(t)\right)^2} \left(s_i(t) + \rho_i^{\bullet}(t) \hat{\omega}_i(t)\right) \tag{18}$$

Denote $\tilde{\omega}_i = \omega_i^* - \hat{\omega}_i$ is the critic weight estimation error and find that $\dot{\tilde{\omega}}_i = -\dot{\hat{\omega}}_i$.

$$\begin{aligned} \dot{\tilde{\omega}}_i(t) = {} & -\frac{\alpha_i \rho_i(t) \rho_i^{\bullet}(t) \tilde{\omega}_i(t)}{\left(1 + \rho_i^{\bullet}(t) \rho_i(t)\right)^2} \\ & + \alpha_i \frac{\rho_i(t)}{\left(1 + \rho_i^{\bullet}(t) \rho_i(t)\right)^2} (s_i(t) + \rho_i^{\bullet}(t) \omega_i^*(t)) \end{aligned} \tag{19}$$

3.2. *Stability Analysis*

Before we discuss the stability of closed-loop systems, we introduce the following assumptions in [5, 6, 8].

Assumption 3. Let the signals $\bar{\rho}_i$ be persistently exciting over the interval $[t, t+T_1]$, i.e., there exist constants $\beta_1 > 0, \beta_2 > 0$ such that, for all t,

$$\beta_1 I \le \int_t^{t+T_1} \overline{\rho_i}(\tau) \bar{\rho}_i^{\bullet}(\tau) d\tau \le \beta_2 I \tag{20}$$

where $\bar{\rho}_i = \rho_i / (\rho_i^{\bullet} \rho_i + 1),\ i = 1, 2, \ldots, N,$ and I is the identity matrix.

Assumption 4. For $\forall i \in \square$, the control input u_i^* is locally Lipschitz with respect to $e_k(t)$. That is, there exists a constant $L_{u,i} > 0$ satisfying that $\| u_i^*(x) - u_i^*(\breve{x}_k) \|^2 \le L_{u,i} \| e_k(t) \|^2$.

Theorem 1. Suppose that Assumptions (1)-(4) holds. For the system (1), the critic NN is updating by (18) and the following self-triggered condition

$$\tau_{k+1} = inf \left\{ t \mid t > \tau_k \cap \left[\frac{\Lambda}{b_f} (e^{b_f (t-\tau_k)} - 1) \le \sqrt{\frac{\mathrm{U}(\breve{x}_k) + e^{-\ell t}}{\mathrm{L}}} \right] \right\} \tag{21}$$

is adopted. Then, the close-loop system state and the critic NN weight estimation error $\tilde{\omega}_i$ are all UUB, where $\ell > 0$ is the decaying rate,

$$\mathrm{L} = \sum_{i=1}^{N} \left((2\lambda_{max}(R_{ii}) + (N-1) b_{g,i}^2) L_{u,i}^2 \right),$$

$$U(\breve{x}_k)=\sum_{i=1}^{N}\lambda_{min}(R_{ii})\|\,\hat{u}_i(\breve{x}_k)\|^2,$$

$$\Lambda=b_f\|\,\breve{x}_k\|+\frac{1}{2}\sum_{j=1}^{N}b_{g,j}^{2}\,\lambda_{max}(R_{jj}^{-1})b_{\nabla\phi_j}b_{\hat{\omega}_j}.$$

Proof. Due to the limited space, the detailed proof here is omitted. We can send the detailed proof to the readers in need.

4. Conclusion

In this paper, we study the optimal control of non-zero-sum game systems with unknown drift dynamics based on self-triggered control. The IRL method is used to avoid the need of unknown dynamics in the solution process. The solution of Nash equilibrium is obtained by constructing a single layer critic NN. By designing a reasonable self-triggered condition, the calculation and communication burden in the whole control process are reduced and no longer requires continuous detection of measurement errors. The UUB properties of the close-loop state and the critic NN estimation error are proved.

References

1. Mu C X, Wang K, Ni Z. Adaptive learning and sampled-control for nonlinear game systems using dynamic event-triggering strategy[J]. IEEE Transactions on Neural Networks and Learning Systems, 2022.
2. Su H G, Zhang H G, Liang Y L, et al. Online event-triggered adaptive critic design for nonzero-sum games of partially unknown networked systems[J]. Neurocomputing, 2019, 368 (9): 84–98.
3. Su H G, Zhang H G, Sun S X, et al. Integral reinforcement learning-based online adaptive event-triggered control for nonzero-sum games of partially unknown nonlinear systems[J]. Neurocomputing, 2020, 377: 243–255.
4. Zhao Q T, Sun J, Wang G, et al. Event-triggered ADP for nonzero-sum games of unknown nonlinear systems[J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, 33(5): 1905-1913.
5. Vamvoudakis K, Lewis F. Multi-player nonzero-sum games: Online adaptive learning solution of coupled Hamilton-Jacobi equations[J]. Automatica, 2011, 47 (8): 1556–1569.
6. Modares H, Lewis F. Optimal tracking control of nonlinear partially-unknown constrained-input systems using integral reinforcement learning[J]. Automatica. 2014, 50 (7): 1780–1792.
7. Modares H, Lewis F, Sistani N. Integral reinforcement learning and experience replay for adaptive optimal control of partially-unknown constrained-input continuoustime systems[J]. Automatica, 2014, 50 (1): 193–202.
8. Yang X, Wei Q L. Adaptive critic learning for constrained optimal event-triggered control with discounted cost[J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, 32 (1): 91–104.

Authors Introduction

Dr. Yibo Shi

He received the M.Sc. degree in the University of Shanghai for Science and Technology in 2022. He is currently pursuing the Ph.D. degree in control science and engineering at the University of Shanghai for Science and Technology, Shanghai, China. His research interests include adaptive control, optimal control, event-triggered, and game theory.

Dr. Chaoli Wang

He received the B.S. and M.Sc. degrees from Mathematics Department, Lanzhou University, Lanzhou, China, in 1986 and 1992, respectively, and the Ph.D. degree in control theory and engineering from the Beijing University of Aeronautics and Astronautics, Beijing, China, in 1999. He is a Professor with the School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai, China. His current research interests include nonlinear control, robust control, robot dynamic and control, visual servoing feedback control, and pattern identification.

Harmony of Agent System with Heterogeneity

Yunzhong Song[†]
School of Electrical Engineering and Automation, Henan Polytechnic University, 2001 Century Avenue
Jiaozuo, 454003, P.R.China

Weicun Zhang
School of Automation and Electrical Engineering, University of Science and Technology Beijing, 30Xueyuan Road
Beijing, 100083, P.R. China

Fengzhi Dai
School of Electronic Information and Automation, Tianjin University of Science and Technology, 1038 Dagu Nanlu
Tianjin, 300222, P.R. China

Huimin Xiao
School of Computer and Information Engineering, Henan University of Economics and Law, 180 Jinshui Donglu
Zhengzhou, 450046, P.R.China

Shumin Fei
School of Automation, South East University, 2 Sipai Lou
Nanjing, 210096, P.R.China

Jichao Zhao

School of Electronic Information and Automation, Tianjin University of Science and Technology, 1038 Dagu Nanlu
Tianjin, 300222, P.R. China

Abstract

This paper focuses on the integration of heterogeneous agents, also known as harmony of them, where heterogeneity emphasized so often. To advance the idea of heterogeneous agents, harmony of the agents, where homogeneity of the agent system turns into heterogeneity one at start, and then the heterogeneity turns back to homogeneity again after. Initiated from hybrid order agents, heterogeneity like social status, encapsulated agents will be introduced step by step. Finally, Chinese philosophy, which always inspires new ideas of thought, can be a good source of research topic. Conclusion that encapsulated agent is the capital Tao of agent systems was drawn from the paper.

Keywords: heterogeneous agents, homogeneity, harmony, lifting techniques, Chinese philosophy

1. Introduction

Agent systems initiated from engineering field not long ago, had been delved detail during the past several years [1-9]. Among them, the integrator agents, started with first order of them, followed by the second order ones. With the theory of agent came to mature gradually, people began to gaze the necessary of the difference of the agents, for in the real world systems, the individuals are so different that makes the world perfect.

[†] Corresponding author

The following of this note will be arranged as follows: Thrive of the heterogeneous agents will be provided at first, then followed by encapsulated agents which is the integration of the heterogeneous agents, can be the cell of the homogeneous agents, continued with harmony of the agents to advance the development, the last part will be conclusions.

2. Thrive of the heterogeneous agents

Thrive of the heterogeneous agents is driven from the requirement of the engineering.

2.1. *Thrive backed by engineering*

The earliest part of the heterogeneous agents is about hybrid order agents, which is comprised of first order dynamic agents and the second order dynamic agents. The first example of them is from electrical engineering, among electrical engineering, it is well known that we can have first order electric circuits and second order electric circuits. One resistor connected with another capacitor or inductor excited by added alternating circuit could be looked as the first order circuit, for this kind circuit, only one single pole stands there, their dynamics are too simple to be analyzed, the response of the output can be only uprising or downsizing. To be different from them, the second order circuit will be a little more complex, for the second order circuits, they have two energy storage elements, one resistor connected by capacitor and inductor simultaneously, the mathematical model of the second order circuit tells us that they have two different poles. In case of two different poles, oscillation between two of the storage elements can occur. However, protocol of the first order agents and the second order agents function not well to hybrid order agents easily. Take rendezvous as an example, in order to come to the same destination, extra damp argument must be introduced for the second dynamic agents, well for the first order agents, no such constraints is necessary. That is to say, when agents composed by different order agents, their collaborative protocols must also be hybrid.

2.2 *Encapsulated agents*

Homogeneous agent systems are favorite at its early stage of agent system research, and replaced by heterogeneous agent systems then after. Then there comes one question, is it possible to integrate two of them into whole? Is there any bridge between two of them? If the bridge exist, can the bridge bring some new breakthrough to the research of the agent systems?

To answer the urgent questions, we could not stop at homogeneity or heterogeneity, either. We should jump out all of them, from the very high standard to inspect all of them. Inspired by the encapsulated medicine for curing headache, fever, cold, influenza, or the other illness, let us start with Fig. 1, which describe the encapsulated agents.

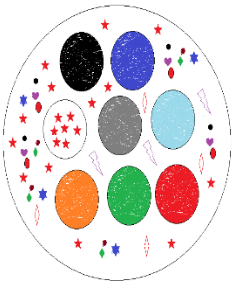


Fig.1 Encapsulated agents

Encapsulated agent is the agent that encapsulated by several hybrid agents, and these agents can be different in shape, in function, in sensing ability and in actuating strength. However, when encapsulated into a single cell, they become one specific agent, which demonstrates whole behavior of itself. To go to further, many of the encapsulated agents can scale up to a larger scale system of them. Fig. 2 gives us an example.

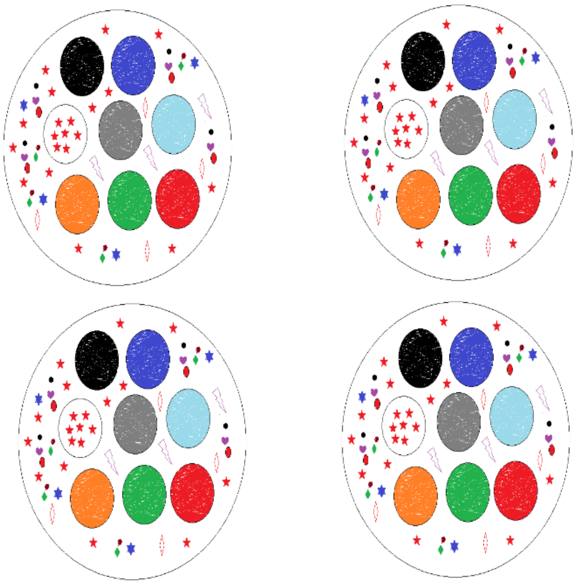


Fig.2 A large scale of the encapsulated system

Particle robotics based on statistical mechanics of loosely coupled components, demonstrated as this kind of example [10].

3 Harmony of the Agents

This part will contribute for harmony of the agents. Chines philosophy as a help, borrowed to understand this important problem.

3.1 Chines philosophy of harmony

As Chinese philosophy has explained: “Besides the tao of every class of things, there is another *Tao* for all things as a whole. In other words, besides the specific multiple tao, there is general unitary Tao which governs the production and transformation of all things. One Yang and one Yin: this is called the Tao. That which ensues from this is goodness, and that which is completed thereby is the nature. This is the Tao of the production of things, and such production is the major achievement of the universe. Everything can in one sense be Yang and in another sense Yin, according to its relation with other things. For instance, a man is Yang in relation to its wife, but Yin in relation to relation to his father. The metaphysical thing Yang which produces all things, however, can only be Yang, and the metaphysical Yin out of which everything is produced can only be Yin; this is called Tao, the Yin and Yang thus spoken of are Yin and Yang in the absolute sense” [11]. The above description of Chinese philosophy of tao and Tao, Yin and Yang can provides us a long time to go to dig out the truth of the harmony of the things in the world. Here, we can try to understand them in a simple way. When heterogeneous agent system was touched upon at first, the homogeneity also existed at the same time, and the reason why people could not see the homogeneity at that time is that the power of homogeneity is weaker than the heterogeneity. With the development of the agent systems, the heterogeneity became stronger and stronger, and the power of heterogeneity transcended the homogeneity, heterogeneous agent systems started to prevail then, almost the same, the second round cycling of heterogeneity homogeneity started again. At first several round of cycling, that can be the concrete examples of them, we can see them, we can touch them, we can decode of them easily. After a long time of cycling, there are so many cores and the cores of them can attract their surroundings, the concrete tao is not enough to deal with the complexity of them, then at this time the Tao must occur to direct the advancement of them.

3.2 *Encapsulated agent is the core of the agents*

From the description of the Chinese philosophy, it is not so arduous for us to get that the encapsulated agent is the core of the agents. We can declare that it is the encapsulated agent, which gives out a landmark for the development of the agent systems. Before appearance of encapsulated agents, we are wandering around heterogeneity and homogeneity of the agent systems, and after that, we feel so comfortable to accept both homogeneity and heterogeneity, for one of them can mean its counterpoint in Chinese philosophy of Yin and Yang, the capital tao of them, also known as Tao of the agent systems.

4 Conclusion

This note mainly contributes to clarify the following two points:

(1) That is the encapsulated agent that forces us to look back, it boils down to that homogeneity and heterogeneity can turn one into another;

(2) Chinese philosophy of Yin and Yang, tao and Tao are a real help for us to understand complex system of agents.

Of course, the Tao induced by time-delay, coupling, fragility, and robust property, is still attractive [12-15].

Acknowledgements

This work is partially supported by NSFC Grant (61340041 and 61374079) and the Project-sponsored by SRF for ROCS, SEM to Yunzhong Song as well as Natural Science Fund of Henan Province (182300410112).

References

1. R. Olfati, R. Murray, Consensus problems in networks of agents with switching topology and time delays, *IEEE Transactions on Automatic Control*, **49**(2004) 1520-1533.
2. Y.Song, F. Dai, H. Xiao and S. Fei, No free lunch principle in agent swarm systems: one Case Study, *Journal of Advances in Artificial Life Robotics*, **1** (2020) 33-37.
3. Y.Song, F. Dai, H. Xiao, *On the complexity of encapsulated agent cell*, Proceedings of 2019 Chinese Intelligent Systems Conference, (2019) 24-30.
4. Y. Song, Z. Fu, F. Wang, Socialized multi-agent system rendezvous via networks of networks, *Journal of Robotics, Networking and Artificial Life*, **2** (2016) 136-139.
5. Y. Song, Z. Fu, F. Wang, Flock guiding of hybrid agents via root block, *Journal of Robotics, Networking and Artificial Life*, **5** (2019) 245-248.

6. Y. Song, Consensus of agents with mixed linear discrete dynamics, *International Journal of Control, Automation and Systems*, **14** (2016) 1139-1143.
7. Y. Song, W. Zhao, Multi-agent system rendezvous via refined social system and individual roles, *WSEAS Transactions on Systems and Control*, **9**(2014),526-532.
8. Y. Hong, J. Hu, L. Gao, Tracking control for multi-agent consensus with an active leader and variable topology, *Automatica*, **42**(2006) 1177-1182.
9. Y. Hong, G. Chen, L. Bushenell, Distributed observers design for leader-following control of multi-agent networks, *Automatica*, **44**(2008) 846-850.
10. S. Li, R. Batra, D. Brown, H. Chang, N. Ranganathan, C. Hoberman, D. Rus & H. Lipson, Particle robotics based on statistical mechanics of loosely coupled components, *Nature*,**567**(2019)361-365.
11. Y. Feng, A short history of Chinese philosophy, *Zhonghua Book Company*, 2017, Peking.
12. Y. Jia, Robust control with decoupling performance for steering and traction of 4WS vehicles under velocity-varying motion, *IEEE Transactions on Control Systems Technology*, **8**(2000) 554-569.
13. Y. Jia, Alternative proofs for improved LMI representations for the analysis and the design of continuous-time systems with polytopic type uncertainty: a predictive approach, *IEEE Transactions on Automatic Control*, **48**(2003) 1413-1416.
14. Y. Jia, General solution to diagonal model matching control of multi-output-delay systems and its applications in adaptive scheme, *Progress in Natural Science*, **19**(2009), 79-90.
15. C. Liu, Y. Song, Distributed economic dispatch strategy of a power system based on load balancing loading, *Power System Protection and Control*, **50**(2022), 139-148 (*In Chinese*).

Authors Introduction

Dr. Yunzhong Song

He received his PhD from Zhejiang University, China in 2006. He is now a full professor in Henan Polytechnic University, and his research interest covers about complex system analysis and control.

Dr. Weicun Zhang

He received his PhD from Tsinghua University, China in1993 and now he is retired from Beijing University of Science and Technology. His research interest covers all most every aspects of adaptive control.

Dr. Fengzhi Dai

He received PhD from Oita University, Japan in 2004. He now works in TianjinUniversity of Science and Technology, China and his main research interests are artificial intelligence and robotics.

Dr. Huimin Xlao

He received his PhD degree in Automatic Control Theory and Its Applications in 1991 from South China University of Technology in China. He is currently a full professor in Henan University of Economics and Law.

Dr. Shumin Fei

He received his PhD degree in Automatic Control Theory and Its Applications in 1995 from Beihang University in China. He is currently a full professor in Southeastern University, China.

Mr. Jichao Zhao

He got his master degree in Tianjin University of Science and Technology in 2022 and now he pursue his PhD in Clemson University, USA.

Apple grading based on IGWO optimized support Vector Machine

Yi Zhao[1]
[1]*School of Electrical Engineering and Automation, Henan Polytechnic University, 2001 Century Avenue, Jiaozuo, (454003), Henan, P.R. China*

Qunpo Liu[1,2], **Yuxi Zhao**[1], **Yueqin Sheng**[1]

[2]*Henan International Joint Laboratory of Direct Drive and Control of Intelligent Equipment，Jiaozuo 454000，P.R. China；*
E-mail: 1535164487@qq.com,lqpny@hpu.edu.cn,, 690430727@qq.com, 212007010029@home.hpu.edu.cn
www.hpu.edu.cn,

Abstract

In order to improve the accuracy of apple external quality classification based on support vector machine, an improved grey wolf optimization algorithm IGWO was proposed by adding Logistic chaos mapping, nonlinear convergence factor and Cauchy variation to the grey wolf optimization algorithm. Firstly, different benchmark functions are used to test the improved IGWO algorithm. The test results show that the IGWO algorithm has improved the convergence speed and accuracy. Secondly, the image processing method is used to extract apple's external features as the data set. The improved grey wolf algorithm was used to optimize the penalty parameters and kernel parameters in support vector machine, and the optimal IGWO-SVM classification model was obtained. Finally, compared with the classification results of SVM and GMO-SVM, the results show that IGWO-SVM has the highest classification accuracy.

Keywords: Apple external quality rating; Improved Grey wolf optimization algorithm IGWO; Support vector machine; Reference function; IGWO-SVM.

1. Introduction

At present, apple classification mainly relies on manual sorting to detect, but this method has strong subjectivity and low classification efficiency, resulting in low accuracy of apple classification. Therefore, it has important research significance and value to use effective science and technology for apple classification.

With the rapid development of image processing technology and machine vision, it has been widely used in the agricultural field.

Li Xianfeng et al. [1] proposed a decision-level multi-feature fusion apple classification method based on D-S evidence theory, and the accuracy rate of apple classification reached 92.5%. Li et al. [2] used image processing technology to extract apple features and BP neural network for classification, and the classification result could reach 92.5%.Xia Qing et al. [3] used the least square support vector machine based on the improved algorithm of particle swarm optimization to classify apples, and the experimental results showed that the accuracy of apple classification was above 96%. Li Xuejun et al.[4] proposed an apple classification algorithm based on the combination of discriminant tree and improved support vector machine decision making. The results showed that the method was feasible and its classification accuracy was above 98%, which could be effectively used for apple classification.

2. Improved grey wolf optimization algorithm

2.1. *Logistic chaos map*

Using Logistic chaotic mapping [5] to substitute random initialization in grey wolf algorithm can improve the uniform distribution of initial population and global search ability of the algorithm. The expression of Logistic chaos mapping is as Eq. (1):

$$X_{n+1} = X_n \times \mu \times (1 - X_n), \mu \in [0,4], X \in [0,1] \qquad (1)$$

2.2. *Improvement of nonlinear convergence factor*

In the improved grey wolf algorithm, in order to improve the search ability of the algorithm, the nonlinear convergence factor such as Eq. (2) is used. The improved convergence factor [6] decreases linearly with the number of iterations from 0 to 2.As shown in Fig.1, at the beginning of iteration, the convergence rate of the improved convergence factor is reduced compared with that of the original algorithm. At the end of iteration, the speed of convergence factor increases, which makes the search for local optimal solution [7] more accurate.

$$a = 2 - 2(\frac{1}{e-1} \times (e^{\frac{t}{t_{\max}}} - 1)) \qquad (2)$$

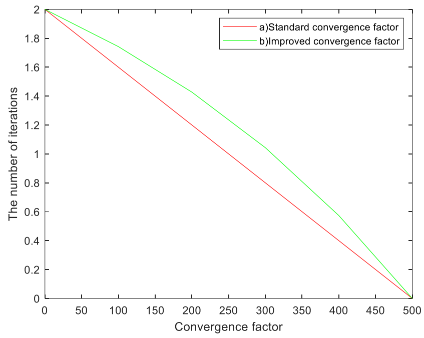


Fig. 1 Nonlinear convergence factor

2.3. *Cauchy variation is introduced*

In order to avoid the algorithm falling into local optimum, Cauchy variation [8] is added to the improved grey wolf algorithm,the variation formula is shown in Eq. (3) :

$$X(\varphi+1) = X_{\text{best}}(\varphi) + C(0,1) \oplus X_{\text{best}}(\varphi) \qquad (3)$$

2.4. *Improved grey Wolf algorithm test*

Tab.1 lists information about standard test functions. Where, F1 and F2 are unimodal functions; F3 and F4 are multimodal functions. Experimental hardware conditions: Intel(R) Core(TM) i5-8300H CPU @ 2.30ghz processor, 16G running memory, software Matlab2018b.

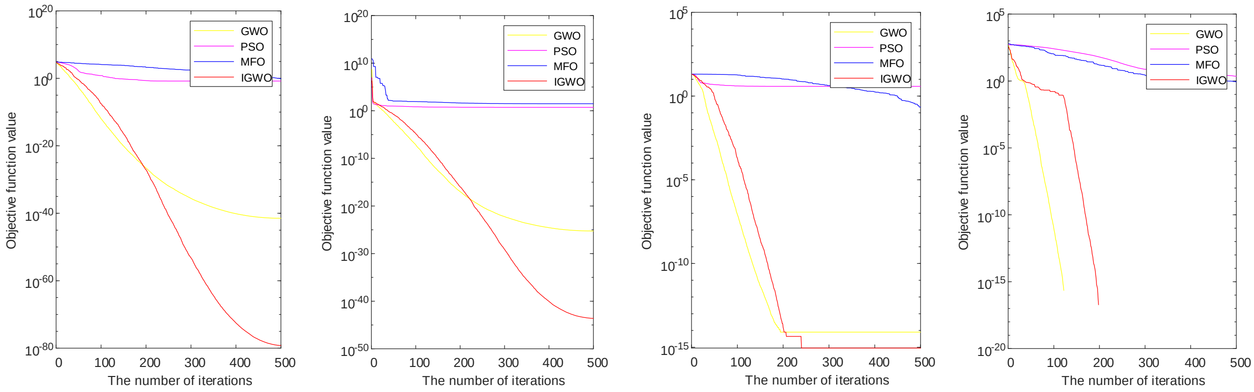
Tab.1 Benchmark function

Function name	Function formula	Dim.	Scope
Sphere	$f_1(x) = \sum_{i=1}^{n} X_i^{\,2}$	30	[-100,100]
Schwefel2.22	$f_2(x) = \sum_{i=1}^{n} \lvert x_i \rvert + \prod_{i=1}^{n} \lvert x_i \rvert$	30	[-10,10]
Ackley	$f_9(x) = -20\exp(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^{\,2}}) - \exp(\frac{1}{n}\sum_{i=1}^{n}\cos(2\pi x_i) + 20 + e$	30	[-32,32]
Griewank	$f_{10}(x) = \frac{1}{4000}\sum_{i=1}^{n} x_i^{\,2} - \prod_{i=1}^{n}\cos\frac{x_i}{\sqrt{i}} + 1$	30	[-600,600]

The experimental results are shown in Tab.2 and the convergence curve is shown in Fig.2. For both single-peak functions F1 and F2 and multi-peak functions F3 and F4, the performance of the improved gray wolf algorithm IGWO is better than that of PSO[9],MFO[10] and GWO.[11]

Tab.2 Comparison of results of different algorithms

Function	PSO		MFO		GWO		IGWO	
	AVE	STD	AVE	STD	AVE	STD	AVE	STD
F1	3.14e-2	9.44e-2	3.57e-3	2.31e-3	4.19e-23	1.02e-22	8.21e-41	2.28e-41
F2	2.00e-3	2.67e-3	1.88e-3	6.85e-3	3.12e-28	2.34e-27	7.85e-47	8.72e-46
F3	2.23e-1	3.02e-2	3.69e-3	5.43e-3	1.74e-16	1.02e-15	4.63e-17	2.53e-16
F4	2.72e+0	1.68e+0	6.18e-2	1.58e-2	2.94e+1	1.12e+0	1.39e-10	7.51e-10



(a) Sphere (b) Schwefel2.22 (c) Ackley (d) Griewank

Fig. 2 Convergence curves of GWO,PSO,MFO and IGWO on test functions

3. Results and Analysis

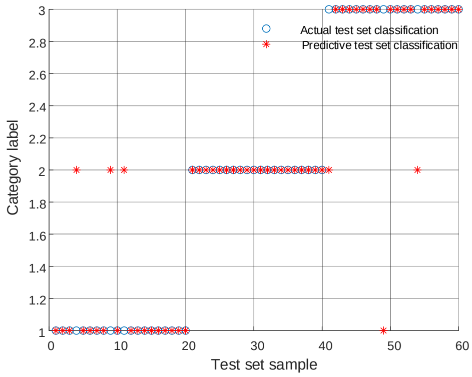


Fig.3 Results of grading using GMO-SVM

In order to verify the classification effect of IGWO proposed in this paper, SVM, [12] GMO-SVM and IGWO-SVM algorithms are successively used for classification experiments. The number of correct classification of SVM algorithm without optimization is 49, and the classification accuracy is 81.66%. As shown in Fig.3, the classification accuracy of the test set has reached 90%, with the classification accuracy of first-class fruit reaching 85%, first-class fruit reaching 100% and second-class fruit reaching 85%. The accuracy of support vector machine optimization using the improved Grey Wolf algorithm reached 98.33%, as shown in Fig.4, which achieved the expected effect of Apple classification.

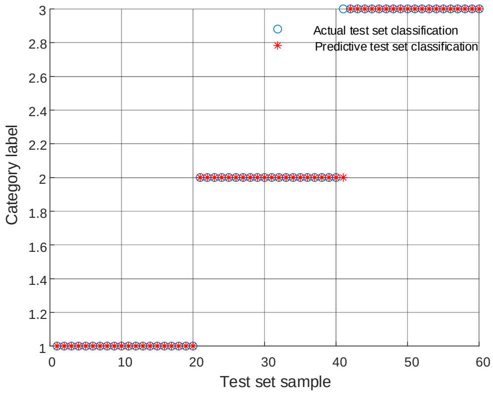


Fig.4 Results of IGWO-SVM classification

4. Conclusion

This paper proposes an improved grey wolf algorithm, namely IGWO algorithm, on the basis of the grey wolf algorithm, adding logistic chaos mapping, nonlinear convergence factor and Cauchy variation. Through the simulation experiments on four standard test functions and compared with PSO,MFO and GWO algorithms, the experimental results show that IGWO algorithm achieves higher optimization accuracy on test functions, and greatly improves the optimization performance in terms of robustness and fast jumping out of local optimum. Secondly, IGWO algorithm is used to optimize the penalty parameters and kernel parameters in support vector machine, and the optimized support vector machine is used to classify apple. The experimental results show that the accuracy rate of IGWO-SVM reaches 98.3%.

References

1. Li Xianfeng, ZHU Weixing, Hua Xiaopeng, KONG Lingdong. Transactions of the Chinese society for agricultural machinery,2011, 42(06):188-192.
2. Xiaoling, L.and Y. Jimin. Detection Level of Apple Based on BP Neural Network. 2015: Atlantis Press.
3. Xia Qing, Li Xianfeng. Apple Classification Detection Based on improved PSO Algorithm and LS-SVM [J]. Computer and Modernization, 2012.
4. Li Xuejun, Cheng Hong. Research on key technology of apple classification detection based on decision fusion. Food & Machinery, 2020, 36(12):5.
5. Yao Y, Ma J. Logical chaotic resonance in a bistable system[J]. International Journal of Bifurcation and Chaos, 2020, 30(13): 2050196.
6. Jarlebring E. Convergence factors of Newton methods for nonlinear eigenvalue problems[J]. Linear algebra and its applications, 2012, 436(10): 3943-3953.
7. Chiang H D, Chu C C. A systematic search method for obtaining multiple local optimal solutions of nonlinear programming problems[J]. IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications, 1996, 43(2): 99-109.
8. Yang J H, Zhao X L, Mei J J, et al. Total variation and high-order total variation adaptive model for restoring blurred images with Cauchy noise[J]. Computers & Mathematics with Applications, 2019, 77(5): 1255-1272.
9. Clerc M. Particle Swarm Optimization[M]. Springer International Publishing, 2016.
10. Li X Y, LI Y. Study on eggshell quality recognition of poultry eggs based on sample entropy and MFO-SVM [J]. Journal of Chinese agricultural mechanization, 2019,40(05):133-139.
11. Mirjalili S, Mirjalili S M, Lewis A. Grey wolf optimizer[J]. Advances in engineering software, 2014, 69: 46-61.
12. Wang H, Hu D. Comparison of SVM and LS-SVM for regression[C]//2005 International conference on neural networks and brain. IEEE, 2005, 1: 279-283.

Authors Introduction

Mrs. Yi Zhao

She graduated from Zhengzhou University of Economics and Business (China) in 2022 with a bachelor's degree in rail transit signal and control. She is currently studying for a master's degree in electronic information at Henan Polytechnic University. She is mainly engaged in research on image processing and behavior recognition.

Dr. Qunpo Liu

He graduated from the Muroran Institute of Technology (Japan) with a Ph.D. in Production Information Systems. He is an associate professor at the School of Electrical Engineering at Henan Polytechnic University (China) and a master's tutor. He is mainly engaged in teaching and research work in robotics, intelligent instruments and machine vision.

Dr. Yuxi Zhao

He graduated from Henan Institute of Engineering with a bachelor's degree in Electrical Engineering and Automation in 2019. Graduate student of Henan Polytechnic University (China), his research interest covers machine vision and image processing.

Mrs. Shengyue Qin

She graduated from Henan Polytechnic University (China) in 2020 with a bachelor's degree in automation. She is currently studying for a master's degree in control science and engineering at Henan Polytechnic University. She is mainly engaged in research on image processing and sign language recognition.

Cartesian Space Coordinated Impedance Control of Redundant Dual-Arm Robots

Yang Zhang
School of Energy and Power 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
E-mail: zhyang19@buaa.edu.cn, ymjia@buaa.edu.cn,
www.buaa.edu.cn

Abstract

This paper presents a cartesian space coordinated impedance control method to achieve coordination when a dual-arm robot operates an object. First, the relative positional and force errors when the two arms operate the object are defined. Then, these relative errors are introduced into the general impedance controller to achieve coordinated impedance control. Compared to the conventional impedance control, this scheme ensures the coordination between the two arms and reduces the contact force error between the end-effectors and the object.

Keywords: Impedance control, Dual-arm robot, Coordinated control, Relative error

1. Introduction

There has been an increasing interest in robotics in recent decades. For additional research and application needs, researchers have developed and studied various types of robots. Among these robots, redundant dual-arm robots are playing an increasingly important role in numerous engineering applications.

When a dual-arm robot operates an object, it can be broadly classified into uncoordinated and coordinated operation according to the constraint relationship between the two arms.[1] In an uncoordinated operation, two robotic arms perform different operation tasks in the same task space, while a coordinated operation means that both robotic arms perform the same or multiple related operation tasks in the same operation space. Compared with the uncoordinated operation, there are stricter motion constraints and force constraints between the robotic manipulators in coordinated operation tasks. The main methods to solve the flexible control of robotic arm by symmetric coordinated control scheme are hybrid force/position control method,[2] impedance control method [3] and intelligent control method.

The impedance control method does not directly control the contact force or the position/velocity but instead satisfies the desired motion characteristics by equating the robot end-operating force-position interactions to a spring-mass-damping model.[4] When a dual-arm robot performs a coordinated operation task, two robotic arms grip the same object and they form a closed-chain motion mechanism.[5] The impedance control of the closed-chain system takes into consideration the distribution of internal forces based on the single-arm impedance control. When controlling a dual-arm robot, the relative error between the arms and the absolute error of the individual arms should be considered. Otherwise, the superimposed error of the two arms will accumulate and then seriously affect the performance of the operation.[6] The problem of error compensation during the flexible operation of a dual-arm robot has yet to be addressed in the existing literature. However, this issue has an essential role in reducing contact force error and ensuring the safety of the operated object.

---

[*] Corresponding author.

This paper proposes a coordinated impedance control method in cartesian space, aiming at the problem of the high requirements for the coordination of the position and force of the arms when a dual-arm robot operates an object. Compared to the traditional impedance control strategy, the proposed method reduces the contact force error between the robotic manipulator and the object, improving the robot's performance when manipulating the object with the dual arm.

2. System Model

The dual-arm robot platform used in this research is shown in Fig. 1.

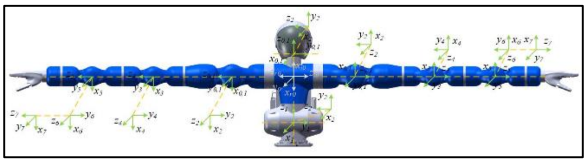


Fig. 1. Redundant dual-arm robot platform

The robot is designed based on humanoid ideas, and each robotics manipulator contains seven rotating joints in a spherical-roll-spherical configuration.

2.1. *Redundant dual-arm robot model*

The dynamics model of the redundant dual-arm robot is shown below:

$$M_D(\theta)\ddot{\theta} + C_D(\theta,\dot{\theta})\dot{\theta} + G_D(\theta) = \tau_a - J_D^T F_D \tag{1}$$

where, $\tau_a \in R^{2n}$ is the joint torque; $F_D \in R^{2m}$ is the contact force from the external environment; $\theta \in R^{2n}$ is the joint angle; $J_D \in R^{2m\times 2n}$ is the Jacobi matrix.

The kinematic model of the redundant dual-arm robot is shown below:

$$\begin{cases} \dot{\chi}_D = J_D(\theta)\dot{\theta} \\ \ddot{\chi}_D = J_D(\theta)\ddot{\theta} + \dot{J}_D(\theta,\dot{\theta})\dot{\theta} \end{cases} \tag{2}$$

where, $\chi_D \in R^{2m}$ is the position and orientation of the dual-arm and satisfies $\chi_D = [\chi_R^T \quad \chi_L^T]^T$.

2.2. *Dual-arm manipulation model*

The schematic diagram of a robot manipulating an object with dual-arm is shown in Fig. 2.

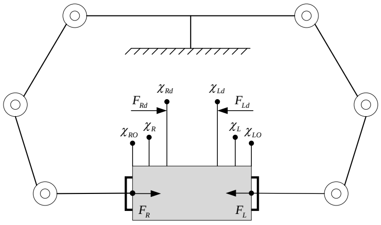


Fig. 2. The analysis of a dual-arm robot manipulating an object

Assume that the two robotic manipulators clamp the object separately and that the end-effectors of the two arms have no relative moving with the object. The mapping between the velocity of the object and the velocity of the end-effector is as follows:

$$\dot{\chi}_D = J_{DO}\dot{y} \tag{3}$$

where, $J_{DO} \in R^{2m\times p}$ is the grasping matrix for the dual-arm. $y \in R^p$ is the position and orientation of the object. The complete kinematic relationship is as follows:

$$\dot{y} = J_{DO}^{\dagger} J_D \dot{\theta} \tag{4}$$

The dynamic equation of the manipulated object is

$$M_Y(y)\ddot{y} + C_Y(y,\dot{y})\dot{y} + G_Y(y)y = F_Y \tag{5}$$

where, F_Y is the combined force that the two robotics manipulators acting on the centroid of the object.

The mapping between the combined force on the centroid of the object and the force at the end of the robotics manipulator is as follows:

$$F_Y = J_{DO}^T F_D \tag{6}$$

3. Coordinated Impedance Controller Design

3.1. *Cartesian space impedance controller*

The cartesian space impedance control is mainly to solve the problem that when the robotic arm is disturbed by the environment, the impedance relationship can still be achieved to ensure the flexibility of the system. The expression of the general impedance model is

$$M\ddot{\tilde{\chi}} + B\dot{\tilde{\chi}} + K\tilde{\chi} = F \tag{7}$$

where, M is the expected inertia matrix; B is the expected damping matrix; K is the expected stiffness matrix; F is the contact force. $\tilde{\chi} \triangleq \chi - \chi_d$ is the position error of end effector.

When grasping an object with dual arms, assume that the outputs of the Cartesian space impedance controller are F_{RI} and F_{LI}, and they satisfy:

$$\begin{cases} F_{RI} = F_R - F_{Rd} \\ F_{LI} = F_L - F_{Ld} \end{cases} \tag{8}$$

Define the output of impedance control as χ_{RI} and χ_{LI}, respectively, and the following relationship can be obtained.

$$\begin{cases} M_R \ddot{\tilde{\chi}}_{RI} + B_R \dot{\tilde{\chi}}_{RI} + K_R \tilde{\chi}_{RI} = F_{RI} \\ M_L \ddot{\tilde{\chi}}_{LI} + B_L \dot{\tilde{\chi}}_{LI} + K_L \tilde{\chi}_{LI} = F_{LI} \end{cases} \tag{9}$$

The tracking trajectory of the general cartesian space impedance control can be obtained as

$$\begin{cases} \chi_R = \chi_{RI} + \chi_{Rd} \\ \chi_L = \chi_{LI} + \chi_{Ld} \end{cases} \tag{10}$$

When the impedance-based control method is used to achieve dual-arm flexible control, the position accuracy of the end-effector is difficult to guarantee, making it more challenging to achieve dual-arm coordination.

3.2. *Definition of relative error*

The difference between each arm's actual and desired pose is defined as the absolute error.

$$\begin{cases} e_R = \chi_{Rd} - \chi_R \\ e_L = \chi_{Ld} - \chi_L \end{cases} \tag{11}$$

Define the product of each arm’s absolute error and the corresponding ratio factor as the relative error.

$$\begin{cases} e_{Rk} \triangleq k_{Rk}(e_R - e_L) \\ e_{Lk} \triangleq k_{Lk}(e_R - e_L) \end{cases} \tag{12}$$

After coordinated control, the end pose of the robotics manipulators satisfies the following relationship.

$$\begin{cases} \chi_{Rk} = \chi_R + e_R + e_{Rk} \\ \chi_{Lk} = \chi_L + e_L + e_{Lk} \end{cases} \tag{13}$$

The controller must compensate for the relative errors for the robot to achieve dual-arm coordination.

3.3. *Coordinated impedance controller*

First, based on the desired trajectory of the object and the external disturbance of the robotics manipulator, we can obtain the impedance control acceleration as

$$\begin{cases} \ddot{\chi}_{RI} = M_R^{-1}\left(F_{RI} - B_R \dot{\tilde{\chi}}_{RI} - K_R \tilde{\chi}_{RI}\right) + \ddot{\chi}_{Rd} \\ \ddot{\chi}_{LI} = M_L^{-1}\left(F_{LI} - B_L \dot{\tilde{\chi}}_{LI} - K_L \tilde{\chi}_{LI}\right) + \ddot{\chi}_{Ld} \end{cases} \tag{14}$$

where, $M_R, B_R, K_R \in R^{m\times m}$ and $M_L, B_L, K_L \in R^{m\times m}$ are the inertia, damping and stiffness of the Cartesian space desired by the right arm and left arm, respectively.

Then, integrating the above equation yields the output pose of the impedance controller. The desired output of the coordinated impedance controller is as follows.

$$\begin{cases} \chi_{RId} = \chi_{RI} + {}^{R}_{L}T\chi_{LI} \\ \chi_{LId} = \chi_{LI} + {}^{L}_{R}T\chi_{RI} \end{cases} \tag{15}$$

After considering the relative error, the desired output of the coordinated impedance controller is as follows.

$$\begin{cases} \chi_{Rk} = e_{Rk} + k_{RI}\chi_{RId} \\ \chi_{Lk} = e_{Lk} + k_{LI}\chi_{LId} \end{cases} \tag{16}$$

By combining the kinematic equations of the two arms, the desired acceleration can be obtained as

$$\ddot{\theta}_d = J_D^{-1}\left[\ddot{\chi}_{Dk} - \dot{J}_D^{-1}\dot{\theta}\right] + N_D^T\lambda \tag{17}$$

where, $\ddot{\chi}_{Dk} = [\ddot{\chi}_{Rk}^T \quad \ddot{\chi}_{Lk}^T]^T$; N_D is the null space matrix for J_D. λ represents an optional vector.

Finally, combining the above equation with the dynamics model yields the desired output torque as

$$\tau_D = M_D J_D^{-1}\left[\ddot{\chi}_{Dk} - \dot{J}_D^{-1}\dot{\theta}\right] + M_D N_D^T \lambda - J_D^T F_D \tag{18}$$

4. Simulation Analysis

This section verified the proposed algorithm using Adams and Matlab simulation platforms. Meanwhile, to demonstrate the superiority of the proposed method, general and coordinated impedance control methods were used in the same task, respectively. The simulation process is that the dual-arm robot holds an object for motion, and the simulation model is shown in Fig. 3.

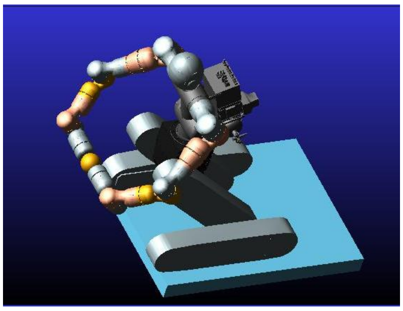


Fig. 3. Coordinated operation of the dual-arm robot

The output force of the robotics manipulators and the total force acting on the object under different controllers are shown in Fig. 4 and Fig. 5.

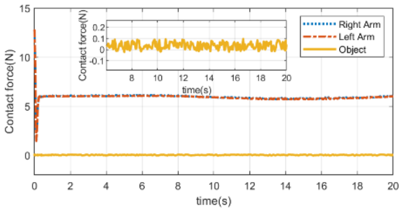


Fig. 4. Variation of contact force based on Cartesian space impedance control method

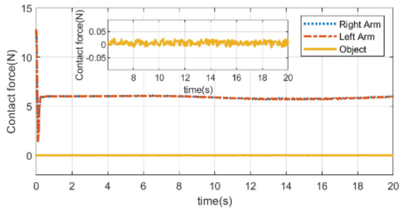


Fig. 5. Variation of contact force based on Cartesian space coordinated impedance control method

Comparing Fig. 4 and Fig. 5, it can be found that the coordination between the arms is significantly improved by using coordinated impedance control, and the fluctuation of the combined force on the object is significantly reduced, which achieves the purpose of regulating the contact force of the object.

5. Conclusion

Aiming at the coordination problem during manipulation of objects by a dual-arm robot, this article proposes a coordinated impedance control method. The contact force error between the dual-arm is compensated by introducing the relative error, which reduces the internal force on the object and improves the system's safety.

Acknowledgements

This work was supported in part by the NSFC under Grant 62133001, and Grant 61520106010, and in part by the National Basic Research Program of China (973 Program) under Grant 2012CB821200 and Grant 2012CB821201.

References

1. C. Smith, et al., "Dual arm manipulation—A survey," Robot. Auton. Syst., vol. 60, no. 10, pp. 1340-1353, 2012.
2. C. Chen, Z. Liu, Y. Zhang and S. Xie, "Coordinated Motion/Force Control of Multiarm Robot with Unknown Sensor Nonlinearity and Manipulated Object's Uncertainty," IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 47, no. 7, pp. 1123-1134, July, 2017.
3. Z. Li, C. Xu, Q. Wei, C. Shi and C. Su, "Human-Inspired Control of Dual-Arm Exoskeleton Robots with Force and Impedance Adaptation," IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 50, no. 12, pp. 5296-5305, December, 2020.
4. G. Xiong, Y. Zhou and J. Yao, "Null-space impedance control of 7-degree-of-freedom redundant manipulators based on the arm angles," Int. J. Adv. Robot. Syst., vol. 17, no. 3, pp. 255688463, May, 2020.
5. J. Lee, P.H. Chang and R.S. Jamisola, "Relative Impedance Control for Dual-Arm Robots Performing Asymmetric Bimanual Tasks," IEEE Trans. Ind. Electron., vol. 61, no. 7, pp. 3786-3796, July, 2014.
6. D. Jinjun, G. Yahui, C. Ming and D. Xianzhong, "Symmetrical adaptive variable admittance control for position/force tracking of dual-arm cooperative manipulators with unknown trajectory deviations," Robot. Comput.-Integr. Manuf., vol. 57, pp. 357-369, 2019.

Authors Introduction

Mr. Yang Zhang

He received the B.S degree in aircraft design and engineering from Northwestern Polytechnical University, Xian, China, in 2016, and the M.S degree in Aerospace Engineering from National University of Defense Technology, Changsha, China, in 2019. He is currently working toward a Ph.D. with the School of Energy and Power Engineering, Beihang University. His main research focusses on motion planning and control of collaborative robot, robust and nonlinear control.

Prof. Yingmin Jia

He received the B.S. degree in control theory from Shandong University, China, in 1982, and the M.S. and PhD degrees both in control theory and applications from Beihang University, China, in 1990 and 1993, respectively. Then, he joined the Seventh Research Division at Beihang University where he is currently Professor of automatic control. His current research interests include robust control, adaptive control and intelligent control, and their applicat-ions in robot systems and distributed parameter systems.

Disturbance Observer-based Anti-unwinding Control for Flexible Spacecrafts

Qian Sun
The Seventh Research Division and the Center for Information and Control, School of Automation Science and Electrical Engineering, Beihang University (BUAA),
Beijing, 100191, P.R.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, P.R.China

Weicun Zhang
School of Automation and Electrical Engineering, University of Science and Technology Beijing,
Beijing 100083, P.R.China
E-mail: mssunqian@buaa.edu.cn, ymjia@buaa.edu.cn, weicunzhang@263.net

Abstract

The anti-unwinding control problem for the six-degrees-of-freedom (6-DOF) motion of the flexible spacecraft is studied in this paper. Firstly, the translation-rotation-vibration coupling motion of the flexible spacecraft is described by dual quaternion. Then, a nonlinear disturbance observer (NDO) is applied to estimate and compensate the lumped disturbances including the flexible vibration and unknown external disturbances. An anti-unwinding controller is designed based on the sliding mode technology. The stability of the closed-loop system is verified via Lyapunov method. Finally, numerical simulations indicate the effectiveness of the designed controller.

Keywords: flexible spacecraft, dual quaternion, disturbance observer, anti-unwinding control

1. Introduction

The flexible spacecraft is a complex dynamic system of rotation-translation-vibration coupling. Many researches have been focused on modeling and control for the 6-DOF motion of the flexible spacecraft. Dual quaternion has been gradually applied to describe the rotational and translational motion of the spacecraft. In this formalism, an actual physical attitude corresponds to two mathematical descriptions. Hence, there are also two equilibrium points $\pm\hat{\mathbf{1}}$, where $\hat{\mathbf{1}} = [1,0,0,0]^{\mathrm{T}} + \varepsilon[0,0,0,0]^{\mathrm{T}}$. Only one of them is considered in most researches, which may lead to the unwinding phenomenon[1]. This phenomenon may cause unnecessary fuel consumptions and should be avoided. There are few relevant investigations about this and it is worthy to be studied.

Vibration suppression is a key issue in flexible spacecraft control. Neural network[2], adaptive control[3] and disturbance observer[4] are three commonly used methods to estimate the flexible vibration which is assumed as disturbances. However, in the initial stage of the neural network approximation, the estimation error is relatively large, which limits its applications. The adaptive control usually estimates the upper bounds of the disturbances. Thus, the disturbance observer is employed in our study.

In this paper, an NDO is firstly proposed to attenuate the influence of unknown external disturbances and the flexible vibration. Then, a sliding mode controller (SMC) is developed, which is of the anti-unwinding performance and guarantees the accuracy and robustness of the flexible spacecraft 6-DOF stability control simultaneously.

---

[*] Corresponding author.

2. Kinematic and Dynamic Model for the Flexible Spacecraft Based on Dual Quaternion

The spacecraft is assumed to be a central rigid body with a flexible appendage, which is shown in Fig. 1.

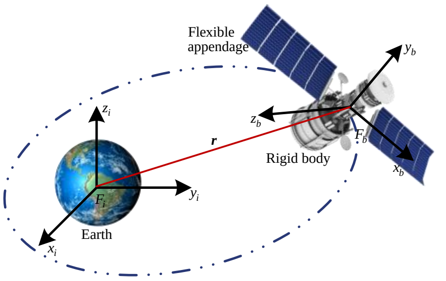


Fig. 1 Coordinate system for the flexible spacecraft

The 6-DOF motion of the flexible spacecraft in F_i can be expressed as $\hat{\boldsymbol{q}} = [\hat{q}_0, \hat{\boldsymbol{q}}_v]^{\mathrm{T}} = \boldsymbol{q} + \varepsilon(1/2\,\boldsymbol{q} \circ \boldsymbol{r})$, where $\boldsymbol{q}$ and $\boldsymbol{r}$ are its rotational and translational motion. Noting that $\hat{\boldsymbol{q}}$ is a unit dual quaternion which satisfies $\|\hat{\boldsymbol{q}}\| = \hat{\mathbf{1}}$ and $q_0 q_0' + \boldsymbol{q}_v^{\mathrm{T}} \boldsymbol{q}_v' = 0$. The kinematic model is given as

$$\dot{\hat{q}}_0 = -\frac{1}{2}\hat{\boldsymbol{q}}_v^{\mathrm{T}}\hat{\boldsymbol{\omega}}, \quad \dot{\hat{\boldsymbol{q}}}_v = \frac{1}{2}\left(\hat{q}_0 \boldsymbol{I} + \hat{\boldsymbol{q}}_v^{\times}\right)\hat{\boldsymbol{\omega}}, \tag{1}$$

where $\hat{\boldsymbol{\omega}} = \boldsymbol{\omega} + \varepsilon \boldsymbol{v} = \boldsymbol{\omega} + \varepsilon\left(\dot{\boldsymbol{r}} + \boldsymbol{\omega} \times \boldsymbol{r}\right)$ is the dual angular velocity. $\boldsymbol{\omega}$ and $\boldsymbol{v}$ are the angular velocity and velocity expressed in F_i, and $\boldsymbol{I} \in \mathbb{R}^{3\times3}$ is the identity matrix. The dynamic model is denoted as

$$\hat{\boldsymbol{M}}\dot{\hat{\boldsymbol{\omega}}} = -\hat{\boldsymbol{\omega}} \times \left(\hat{\boldsymbol{M}}\hat{\boldsymbol{\omega}} + \hat{\boldsymbol{B}}\dot{\hat{\boldsymbol{\eta}}}\right) - \hat{\boldsymbol{B}}\ddot{\hat{\boldsymbol{\eta}}} + \hat{\boldsymbol{g}} + \hat{\boldsymbol{u}} + \hat{\boldsymbol{d}}_0, \tag{2}$$

where $\hat{\boldsymbol{M}}$ is given as $\hat{\boldsymbol{M}} = \left(\mathrm{d}/\mathrm{d}\varepsilon\right) m\boldsymbol{I} + \varepsilon \boldsymbol{J}$. m, $\boldsymbol{J}$ are the mass and inertia matrix. Dual vectors $\hat{\boldsymbol{g}} = \boldsymbol{f}_g + \varepsilon \boldsymbol{\tau}_g$, $\hat{\boldsymbol{u}} = \boldsymbol{f}_u + \varepsilon \boldsymbol{\tau}_u$ and $\hat{\boldsymbol{d}}_0 = \boldsymbol{f}_{d0} + \varepsilon \boldsymbol{\tau}_{d0}$ represent the gravity force, control input and unknown external disturbances, respectively. $\boldsymbol{\eta} \in \mathbb{R}^n$ denotes the modal coordinate vector of the flexible appendage and n is the modal number. The dynamics of the modal coordinate is presented as

$$\ddot{\boldsymbol{\eta}} + 2\xi\boldsymbol{\Omega}\dot{\boldsymbol{\eta}} + \boldsymbol{\Omega}^2\boldsymbol{\eta} + \boldsymbol{B}_r^T\boldsymbol{\omega} + \boldsymbol{B}_t^T\boldsymbol{v} = \boldsymbol{0}, \tag{3}$$

where $\xi = \mathrm{diag}\{\xi_i\}$ $(i = 1, \ldots, n)$ is the damping ratio matrix and $\boldsymbol{\Omega} = \mathrm{diag}\{\Omega_i\}$ is the natural frequency matrix. The dual vector $\hat{\boldsymbol{\eta}}$ holds the definition of $\hat{\boldsymbol{\eta}} = \boldsymbol{\eta} + \varepsilon\boldsymbol{\eta}$. $\hat{\boldsymbol{B}} = \left(\mathrm{d}/\mathrm{d}\varepsilon\right)\boldsymbol{B}_t + \varepsilon \boldsymbol{B}_r$ is the rigid-flexible coupling dual matrix with $\boldsymbol{B}_t$, $\boldsymbol{B}_r \in \mathbb{R}^{3\times n}$.

Then, denote $\hat{\boldsymbol{d}} = \boldsymbol{d} + \varepsilon \boldsymbol{d}' = -\hat{\boldsymbol{\omega}} \times \hat{\boldsymbol{B}}\dot{\hat{\boldsymbol{\eta}}} - \hat{\boldsymbol{B}}\ddot{\hat{\boldsymbol{\eta}}} + \hat{\boldsymbol{d}}_o$ as the lumped disturbances, which are assumed to be bounded and their time derivative is approximated to zero. The dynamic equation (2) can be rewritten as

$$\hat{\boldsymbol{M}}\dot{\hat{\boldsymbol{\omega}}} = -\hat{\boldsymbol{\omega}} \times \hat{\boldsymbol{M}}\hat{\boldsymbol{\omega}} + \hat{\boldsymbol{g}} + \hat{\boldsymbol{u}} + \hat{\boldsymbol{d}}. \tag{4}$$

3. Disturbance Observer-based Anti-unwinding Controller Design

3.1. *Nonlinear Disturbance observer design*

The NDO is design as

$$\begin{cases} \hat{\boldsymbol{D}} = \hat{\boldsymbol{z}}_d + \hat{\boldsymbol{\rho}}\left(\hat{\boldsymbol{\omega}}\right), \\ \dot{\hat{\boldsymbol{z}}}_d = -\hat{\boldsymbol{\lambda}} \square\ \hat{\boldsymbol{z}}_d + \hat{\boldsymbol{\lambda}} \square \left[\hat{\boldsymbol{\omega}} \times \hat{\boldsymbol{M}}\hat{\boldsymbol{\omega}} - \hat{\boldsymbol{g}} - \hat{\boldsymbol{u}} - \hat{\boldsymbol{\rho}}\left(\hat{\boldsymbol{\omega}}\right)\right], \end{cases} \tag{5}$$

where the dual function $\hat{\boldsymbol{\rho}}\left(\hat{\boldsymbol{\omega}}\right) = \hat{\boldsymbol{\lambda}} \square\ (\hat{\boldsymbol{M}}\hat{\boldsymbol{\omega}})$ and the dual matrix $\hat{\boldsymbol{\lambda}} = \boldsymbol{\lambda} + \varepsilon\boldsymbol{\lambda}' = \mathrm{diag}\{\lambda_i\} + \varepsilon\,\mathrm{diag}\{\lambda_i'\}$ $(i = 1,2,3)$ with λ_i, $\lambda_i' > 0$. $\hat{\boldsymbol{D}}$ is the estimation of $\hat{\boldsymbol{d}}$. $\square$ is the corresponding product.

Theorem 1. Considering the dynamic equation (4) of the flexible spacecraft, the estimation error $\hat{\boldsymbol{d}}_e = \hat{\boldsymbol{d}} - \hat{\boldsymbol{D}}$ can converge to the origin with the proposed NDO (5).

Proof. Select the following Lyapunov function candidate

$$V_1 = \frac{1}{2}\left\langle \hat{\boldsymbol{d}}_e, \hat{\boldsymbol{d}}_e \right\rangle, \tag{6}$$

where $\langle \square, \square \rangle$ is the inner product. By (5), we have

$$\dot{V}_1 = -\left\langle \hat{\boldsymbol{d}}_e, \hat{\boldsymbol{\lambda}} \square\ \hat{\boldsymbol{d}}_e \right\rangle = -\sum_{i=1}^{3}\left(\lambda_i d_{ei}^2 + \lambda_i' d_{ei}'^2\right). \tag{7}$$

It can be derived that $\dot{V}_1 < 0$ when $\hat{\boldsymbol{d}}_e \neq \hat{\mathbf{0}}$. By LaSalle's invariance principle, the designed NDO can track the lumped disturbances $\hat{\boldsymbol{d}}$ with asymptotic convergence. □

3.2. *Anti-unwinding Controller Design*

A sliding mode surface $\hat{\boldsymbol{s}}$ is presented as

$$\hat{\boldsymbol{s}} = \boldsymbol{s} + \varepsilon\boldsymbol{s}' = \hat{\boldsymbol{\omega}} + \hat{\boldsymbol{\mu}} \square\ \left(\chi \hat{\boldsymbol{q}}_v\right), \tag{8}$$

where $\hat{\boldsymbol{\mu}} = \mathrm{diag}\{\mu_i\} + \varepsilon\,\mathrm{diag}\{\mu_i'\}$ $(i = 1,2,3)$. μ_i and μ_i' are all positive factors. The parameter χ is defined as

$$\chi = \begin{cases} 1, & q_0(0) \geq 0 \\ -1, & q_0(0) < 0 \end{cases} \tag{9}$$

where $q_0(0)$ is the initial value of the real part of $\hat{q}_0$.

Theorem 2. Under the condition of $\hat{\boldsymbol{s}} = \hat{\mathbf{0}}$, state variables $\hat{\boldsymbol{\omega}}$ and $\hat{\boldsymbol{q}}$ finally converge to the state $\{\hat{\boldsymbol{q}} = \pm\hat{\mathbf{1}},\ \hat{\boldsymbol{\omega}} = \hat{\mathbf{0}}\}$. Meanwhile, the unwinding phenomenon are avoided.

Proof. Select the Lyapunov function candidate

$$V_2 = 2\left(1 - \chi q_0\right) + \frac{1}{4}\boldsymbol{r}^{\mathrm{T}}\boldsymbol{r}. \tag{10}$$

The time derivative of V_2 is given as

$$\dot{V}_2 = -2\chi\dot{q}_0 + \frac{1}{2}\boldsymbol{r}^{\mathrm{T}}\dot{\boldsymbol{r}} = \chi\boldsymbol{q}_v^{\mathrm{T}}\boldsymbol{\omega} + \left(\boldsymbol{q}^* \circ \boldsymbol{q}'\right)^{\mathrm{T}}\boldsymbol{v}. \qquad (11)$$

When $\hat{\boldsymbol{s}} = \hat{\boldsymbol{0}}$, $\hat{\boldsymbol{\omega}}$ can be expressed as $\hat{\boldsymbol{\omega}} = -\hat{\boldsymbol{\mu}} \square \left(\chi\hat{\boldsymbol{q}}_v\right)$. Applying the properties of the unit dual quaternion yields

$$\begin{aligned}\dot{V}_2 &= -|\chi|^2 \boldsymbol{q}_v^{\mathrm{T}}\left(\boldsymbol{\mu}\boldsymbol{q}_v\right) - \left(q_0\boldsymbol{q}_v' - q_0'\boldsymbol{q}_v\right)^{\mathrm{T}}\left(\chi\boldsymbol{\mu}'\boldsymbol{q}_v'\right)\\ &< -\mu_m|\chi|^2\|\boldsymbol{q}_v\|^2 - \mu_m'|q_0|\|\boldsymbol{q}_v'\|^2 - \frac{\mu_m'}{|q_0|}\|\boldsymbol{q}_v\|^2\|\boldsymbol{q}_v'\|^2, \qquad (12)\end{aligned}$$

where $\mu_m = \min\{\mu_i\}$, $\mu_m' = \min\{\mu_i'\}$ $(i = 1, 2, 3)$. It can be obtained that $\dot{V}_2 < 0$ when $\hat{\boldsymbol{q}}_v \neq \hat{\boldsymbol{0}}$. Thus, $\hat{\boldsymbol{\omega}}$ and $\hat{\boldsymbol{q}}$ can converge to the state $\{\hat{\boldsymbol{q}} = \pm\hat{\boldsymbol{1}},\ \hat{\boldsymbol{\omega}} = \hat{\boldsymbol{0}}\}$. In this case, the kinematic model of the rotation can be written as

$$\dot{q}_0 = -\frac{1}{2}\boldsymbol{q}_v^{\mathrm{T}}\boldsymbol{\omega} = \frac{\chi}{2}\boldsymbol{q}_v^{\mathrm{T}}\boldsymbol{\mu}\boldsymbol{q}_v. \qquad (13)$$

Thus, if $q_0(0) \geq 0$, $\chi = 1$ and $\dot{q}_0 \geq 0$ are ensured, the state variables finally reaches the equilibrium point $\hat{\boldsymbol{1}}$. And $q_0(0) < 0$ leads the trajectory reach to the state $-\hat{\boldsymbol{1}}$. Hence, the anti-unwinding property is guaranteed. □

Subsequently, the control law based on the NDO (5) and the sliding mode surface (8) is proposed as

$$\begin{aligned}\hat{\boldsymbol{u}} &= \hat{\boldsymbol{\omega}} \times \hat{\boldsymbol{M}}\hat{\boldsymbol{\omega}} - \frac{1}{2}\chi\hat{\boldsymbol{\mu}} \square \hat{\boldsymbol{M}}\left(\hat{q}_0\boldsymbol{I} + \hat{\boldsymbol{q}}_v^{\times}\right)\hat{\boldsymbol{\omega}} - \hat{\boldsymbol{g}} - \hat{\boldsymbol{D}}\\ &\quad - \hat{\boldsymbol{\kappa}}_1 \square \hat{\boldsymbol{s}}^s - \hat{\boldsymbol{\kappa}}_2 \square \operatorname{sgn}\left(\hat{\boldsymbol{s}}^s\right) \qquad (14)\end{aligned}$$

where $\hat{\boldsymbol{\kappa}}_x = \operatorname{diag}\{\kappa_{xi}\} + \varepsilon\operatorname{diag}\{\kappa_{xi}'\}$ $(x = 1, 2;\ i = 1, 2, 3)$. κ_{xi}, κ_{xi}' are all positive parameters. $\hat{\boldsymbol{s}}^s$ is the swap operation of $\hat{\boldsymbol{s}}$.

Theorem 3. Considering the kinematic and dynamic model of the spacecraft (1) and (4), the trajectory of the closed-loop system will asymptotically converge to the manifold $\hat{\boldsymbol{s}} = \hat{\boldsymbol{0}}$ by the control law (14).

Proof. Consider the Lyapunov candidate function

$$V_3 = \frac{1}{2}\left[\hat{\boldsymbol{s}}\middle|\hat{\boldsymbol{M}}\hat{\boldsymbol{s}}\right] + \frac{1}{2}\left\langle\hat{\boldsymbol{d}}_e, \hat{\boldsymbol{d}}_e\right\rangle, \qquad (15)$$

where $[\square|\square]$ is the switched inner product. Differentiating V_3 yields

$$\dot{V}_3 = \left[\hat{\boldsymbol{s}}\middle|\begin{matrix}-\hat{\boldsymbol{\omega}} \times \hat{\boldsymbol{M}}\hat{\boldsymbol{\omega}} + \hat{\boldsymbol{g}} + \hat{\boldsymbol{u}} + \hat{\boldsymbol{d}}\\ +\frac{1}{2}\chi\hat{\boldsymbol{\mu}} \square \hat{\boldsymbol{M}}\left(\hat{q}_0\boldsymbol{I} + \hat{\boldsymbol{q}}_v^{\times}\right)\hat{\boldsymbol{\omega}}\end{matrix}\right] + \left\langle\hat{\boldsymbol{d}}_e, \dot{\hat{\boldsymbol{d}}}_e\right\rangle, \qquad (16)$$

Substituting (14) into (16), it can be derived that

$$\begin{aligned}\dot{V}_3 &= \sum_{i=1}^{3}\left(d_{ei}'s_i + d_{ei}s_i'\right) - \sum_{i=1}^{3}\left(\kappa_{1i}'s_i^2 + \kappa_{1i}s_i'^2\right)\\ &\quad - \sum_{i=1}^{3}\left(\kappa_{2i}'|s_i| + \kappa_{2i}|s_i'|\right) - \sum_{i=1}^{3}\left(\lambda_i d_{ei}^2 + \lambda_i' d_{ei}'^2\right). \qquad (17)\end{aligned}$$

Moreover, let $a_{1i} = 1/2\sqrt{\lambda_i}$, $a_{1i}' = 1/2\sqrt{\lambda_i'}$, $a_{2i} = 1/4\lambda_i$, and $a_{2i}' = 1/4\lambda_i'$, we have

$$\begin{aligned}\dot{V}_3 &= -\sum_{i=1}^{3}\left[\left(a_{1i}'s_i - \sqrt{\lambda_i'}d_{ei}'\right)^2 + \left(a_{1i}s_i' - \sqrt{\lambda_i}d_{ei}\right)^2\right]\\ &\quad - \sum_{i=1}^{3}\left[\left(\kappa_{1i}' - a_{2i}'\right)s_i^2 + \left(\kappa_{1i} - a_{2i}\right)s_i'^2\right] - \sum_{i=1}^{3}\left(\kappa_{2i}'|s_i|\right.\\ &\quad \left. + \kappa_{2i}|s_i'|\right) \leq 0. \qquad (18)\end{aligned}$$

It is obvious that $V_3 \leq V_3(0)$ and V_3 is bounded. The definition of the boundedness of a dual quaternion is that both its real and dual part are bounded. It is derive that $\hat{\boldsymbol{s}}$, $\hat{\boldsymbol{d}}_e$, $\hat{\boldsymbol{q}}$, $\hat{\boldsymbol{\omega}}$ and $\hat{\boldsymbol{u}}$ are bounded. Then, we can also obtain that $\ddot{V}_3$ is bounded. By Barbalat's Lemma, as time goes to infinity, $\dot{V}_3$ finally converges to zero. Thus, the sliding mode variable $\hat{\boldsymbol{s}}$ converges to $\hat{\boldsymbol{0}}$ asymptotically. □

4. Numerical Simulations

This section simulates the hovering mission. The parameters of the flexible spacecraft are the same as Ref.[5]. The initial pose is set as $\boldsymbol{q}_0 = [-0.6403, -0.5, -0.3, 0.5]^{\mathrm{T}}$ and $\boldsymbol{r}_0 = [5, -5, 5]^{\mathrm{T}}\,\mathrm{m}$. The initial value of the angular velocity and velocity are set as $\boldsymbol{\omega}_0 = [0.1, -0.1, 0.1]^{\mathrm{T}}\,\mathrm{rad/s}$ and $\boldsymbol{v}_0 = [0.1, -0.1, 0.1]^{\mathrm{T}}\,\mathrm{m/s}$. Control parameters are denoted as:

$$\begin{aligned}&\hat{\boldsymbol{\lambda}} = \operatorname{diag}(0.7,\ 0.7,\ 0.7) + \varepsilon\operatorname{diag}(0.7,\ 0.7,\ 0.7),\\ &\hat{\boldsymbol{\mu}} = \operatorname{diag}(0.1,\ 0.1,\ 0.1) + \varepsilon\operatorname{diag}(1,\ 1,\ 1),\\ &\hat{\boldsymbol{\kappa}}_1 = \hat{\boldsymbol{\kappa}}_2 = \operatorname{diag}(0.5,\ 0.5,\ 0.5) + \varepsilon\operatorname{diag}(3,\ 3.5,\ 3).\end{aligned}$$

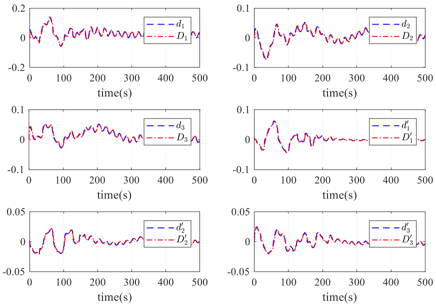


Fig. 2 Time responses of the lumped disturbances and their estimations

Fig. 2 describes the time response of the lumped disturbances and their estimations. The variations of the attitude, angular velocity, position and velocity are mentioned in Fig. 3 and Fig. 4. It can be obtained that when $q_0(0) < 0$, the attitude finally converges to the state $\boldsymbol{q} = [-1, 0, 0, 0]^{\mathrm{T}}$ and the unwinding problem is handled from Fig. 3. The time responses of the control force and torque are shown in Fig. 5. Thus, the accuracy and robustness of the controller (14) can be verified.

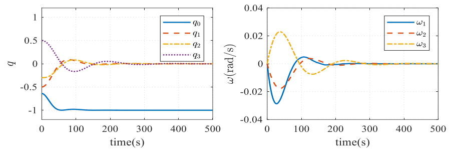


Fig. 3 Time responses of attitude and angular velocity

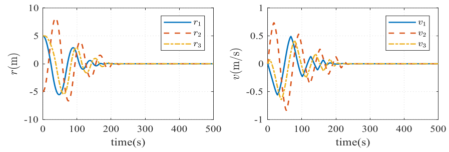


Fig. 4 Time responses of position and velocity

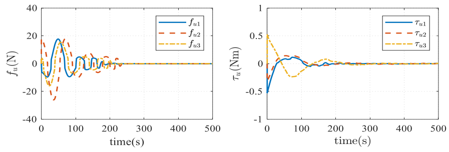


Fig. 5 Time responses of control force and torque

5. Conclusions

This paper considers the stabilization control for the flexible spacecraft's 6-DOF motion. The NDO is constructed to estimate the lumped disturbances. The unwinding phenomenon is solved by designing a SMC, which optimizes the rotational path. Meanwhile, the strong robustness and high precision control can be achieved.

Acknowledgements

This work was supported in part by the NSFC (62133001, 61520106010) and the National Basic Research Program of China (973 Program: 2012CB821200, 2012CB821201).

References

1. R.Q. Dong, A.G. Wu, Y. Zhang. "Anti-unwinding Sliding Mode Attitude Maneuver Control for Rigid Spacecraft", IEEE Transactions on Automatic Control, Vol. 67, No. 2, pp. 978-985, 2021.
2. T. He, Z. Wu. "Neural Network Disturbance Observer with Extended Weight Matrix for Spacecraft Disturbance Attenuation", Aerospace Science and Technology, Vol. 126, pp: 107572.
3. Y. Zhang, Y. Wu, A. Wu, et al. "Adaptive Control for Disturbance Attenuation of Flexible Spacecraft". 2018 37th Chinese Control Conference (CCC), pp. 879-884, 2018.
4. C. Zhang, G. Ma, Y. Sun, et al. "Observer-based Prescribed Performance Attitude Control for Flexible Spacecraft with Actuator Saturation", ISA transactions, Vol. 89, pp. 84-95, 2019.
5. X. Zhu, Z.H. Zhu, J. Chen. "Dual quaternion-based adaptive iterative learning control for flexible spacecraft rendezvous", Acta Astronautica, Vol. 189, pp. 99-118, 2021.

Authors Introduction

Ms. Qian Sun

She received the B.S. degree in Automation from Chongqing University, Chongqing, China, in 2020. She is currently pursuing the PhD degree at Beihang University, Beijing, China. Her research interests include coordinated control of on-orbit services.

Prof. Yingmin Jia

He received the PhD degree in control theory and applications from Beihang University, in 1993. He is currently a Professor with the School of Automation Science and Electrical Engineering, Beihang University. His research interests include robust control, robot control and aerospace control.

Ass. Prof. Weicun Zhang

He received the PhD degree in control theory and applications from Tsinghua University. He is an Associate Professor with the School of Automation and Electrical Engineering, University of Science and Technology Beijing. His current research interests include adaptive control, intelligent control, and their applications.

A Survey of Target Detection Based on Deep Learning

Hucheng Wang*, Fengzhi Dai, Min Zhao

College of Electronic Information and Automation, Tianjin University of Science and Technology,
300222，China

E-mail: *18322744268@163.com
www.tust.edu.cn

Abstract

Object detection is a hot topic in the field of visual detection. Deep learning can greatly compensate for the defect that traditional methods sacrifice real-time for improving accuracy. This paper mainly introduces the main networks and methods of two-stage deep learning algorithm and single-stage deep learning algorithm in the field of target detection. The advantages and disadvantages, usage scenarios and development of each network are described in detail. Finally, the follow-up development in this field is prospected..

Keywords: Machine earning, Deep learning, Object detection, Convolutional neural network

1. Introduction

Object detection is one of the most important topics in the field of computer vision. Its task is to find all interested objects in the detected image and classify them. Because the image quality is affected by weather conditions, lighting, occlusion and other factors, and different objects have different attitudes, shapes and surface roughness, target detection has always been a very hot and difficult problem in the field of computer vision.

Although the traditional method (machine learning) has been very mature in the field of object detection, it still has many problems, such as: (1) How to select the sliding window. If it is a complex practical project application, this problem will consume a lot of time and energy. (2) The robustness of manually calibrated features is poor. How to overcome the above two problems became a hot issue in the field of target detection at that time.

Depth convolution neural network can not only overcome the typical problems of target detection with traditional methods, but also greatly shorten the detection time. this model makes depth learning receive great attention, and gradually become the mainstream of researchers.

2. Two Stage Deep Learning Algorithm

The so-called "two stages" refer to: 1. Extract the object area first. 2. Then classify and identify the area by CNN. The common methods for extracting object regions are: selective search, boundary box, etc. The main feature of this algorithm is high accuracy, but slow speed. Typical algorithms include R-CNN, SPP-Net, fast R-CNN and fast R-CNN.

2.1. Region Convolution Neural Network

The Girshick team [1] first proposed this method in 2014. Its algorithm idea is very simple. First, select several candidate boxes from the original image based on the Selective Search method. Second, scale the image in each candidate box to a fixed scale and send it to the convolutional neural network for feature map. Finally, it is judged by the support vector machine (SVM), Determine whether the images in these candidate boxes are the image type we want to extract or the background.

Compared with traditional methods, the efficiency of this method is greatly improved. The disadvantages are as follows: 1. When the candidate regions are normalized, it is easy to cause image loss. 2. CNN model parameters and classification regression parameters cannot be modified at the same time. 3. It takes a lot of disk space, making the detection time long.

2.2. Spatial Pyramid Pooling Network

In 2014, He Kaiming [2] and others put forward SPP-Net on the basis of R-CNN, which requires input convolutional neural network to extract features of the sub image to be tested with fixed size. SPP Net cancels this restriction. SPP Net has a pyramid space pyramid pool layer, which makes it possible to plan the sub image into a uniform size regardless of the size of the input sub image. Each sub image with the same size is sent to the subsequent network for special extraction, and the extracted features also have the same dimension.

Compared with CNN, SPP Net is more efficient in target recognition.

2.3. Fast Area Convolution Neural Network

In 2015, Girshick team [3] further optimized and improved on the basis of CNN and SPP Net, and proposed convolutional neural network (fast R-CNN), which can simultaneously predict the classification probability and position offset of targets in the network. The network solves two main problems of the two networks mentioned above: 1. There are many training steps and slow training speed. 2. It takes a lot of disk space, resulting in a long training time. Fig.1 is the schematic diagram of fast R-CNN algorithm.

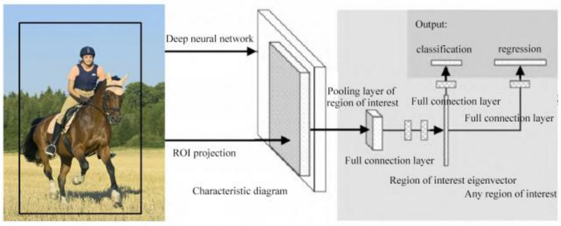

Deep neural network
ROI projection
Characteristic diagram
Output:
classification
regression
Pooling layer of region of interest
Full connection layer
Full connection layer
Full connection layer
Region of interest eigenvector
Any region of interest

Fig.1. the schematic diagram of fast R-CNN algorithm

When using the same scale network and the same data set, fast R-CNN has faster detection speed and higher detection accuracy than R-CNN. Table 1 shows the performance comparison of R-CNN and fast R-CNN training and testing based on VGG-16 convolutional network model on the VOC2007 dataset.

Table 1. Performance comparison between VGG-16 based R-CNN and fast R-CNN algorithm

Algorithm	Training Time (h)	Test time (seconds/frame)	mAP(%)
R-CNN	84	47	66
Fast R-CNN	95	0.32	66.9

2.4. Fast Area Convolution Neural Network

This network was proposed by Ren S Q [4] team in 2015. The biggest feature of this network is that it integrates the four target detection steps into a deep network. The algorithm process can be divided into the following steps: 1. Input the tested image into the network. 2. Use regional bidding network (RPN) (Fig.2 shows the RPN algorithm) and discriminant function to obtain accurate candidate regions. 3. Unify the images to be input to the full connection layer into the same size through pooling layer. 4. Identify the category and position of the tested object in the tested image through the training of the full connection layer.

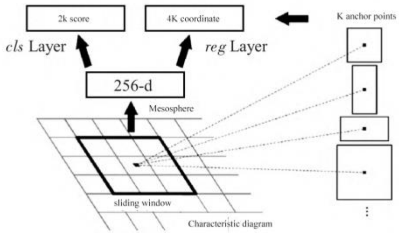

2k score
4K coordinate
K anchor points
cls Layer
reg Layer
256-d
Mesosphere
sliding window
Characteristic diagram

Fig.2. RPN algorithm

3. Single Stage Deep Learning Aigorithm

The single-stage deep learning algorithm refers to the end-to-end method. For a detected image, only one network is used to identify the category and location of the object to be detected, simplifying the complex steps of the two-stage deep learning algorithm, and greatly improving the efficiency. This method has excellent real-time performance. Typical single-stage deep learning algorithms include yolo and SSD.

3.1. YOLO

The two-stage detection algorithm needs to go through two steps: border regression and classification. This will generate a large number of candidate boxes, making the detection time very long despite the high accuracy. It can not meet the requirements when executing some tasks that require high real-time performance. In view of this, Joseph Redmon proposed the yolov1 model. The model cancels the candidate box and directly partitions the image to be measured. One area corresponds to one grid. A neural network is used to traverse the image and predict the border and category information of objects in each grid. Since the prediction of each border is based on the characteristics of the whole picture, the predicted object border combines the information of the whole picture, which not only improves the prediction efficiency but also has better accuracy, However, the prediction of small objects is not effective. Fig.3 is the schematic diagram of YOLOV1 network.

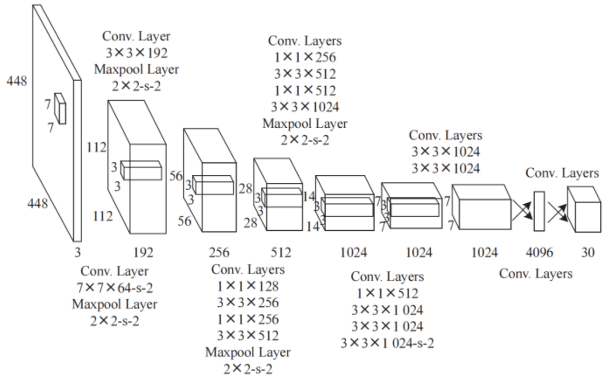

Conv. Layer
3×3×192
Maxpool Layer
2×2-s-2
Conv. Layers
1×1×256
3×3×512
1×1×512
3×3×1024
Maxpool Layer
2×2-s-2
Conv. Layers
3×3×1024
3×3×1024
Conv. Layers
448
7
7
448
112
112
3
3
56
56
3
3
28
28
3
3
14
14
3
3
7
7
3
3
7
7
3
192
256
512
1024
1024
1024
4096
30
Conv. Layer
7×7×64-s-2
Maxpool Layer
2×2-s-2
Conv. Layers
1×1×128
3×3×256
1×1×256
3×3×512
Maxpool Layer
2×2-s-2
Conv. Layers
1×1×512
3×3×1 024
3×3×1 024
3×3×1 024-s-2
Conv. Layers

Fig.3. The schematic diagram of YOLOV1 network

In order to solve the problems such as the poor detection effect of yolov1, the yolov2 version was subsequently iterated. This version introduced the BN algorithm to improve the convergence speed of the model, and took Darknet-19 as the backbone network, instead of using the full connection layer, which simplified the network structure and improved the efficiency.

Yolo algorithm, as a classical algorithm for real-time target detection, has been continuously optimized and iterated since its appearance. So far, the yolov7 version has exceeded all known target detection algorithms both in real-time and accuracy. And it reaches 56.8% AP on COCO dataset. Fig.4 shows the performance comparison of yolo versions.

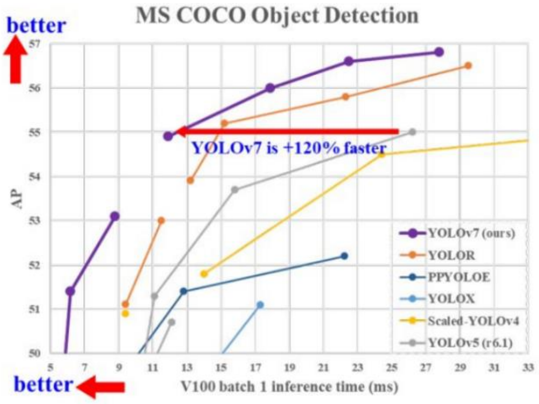


Fig.4. Comparison Chart of Yolo Performance of Different Versions

3.2. SSD

The emergence of SSD is mainly to solve the problems of low positioning accuracy and poor detection effect for small objects existing in the previous version of yolo network. In 2017, Fu C W team [5] further improved the algorithm and proposed the DSSD algorithm, which further improved the detection accuracy of small objects, but also increased the network complexity. Fig.5 is the schematic diagram of SSD network structure.

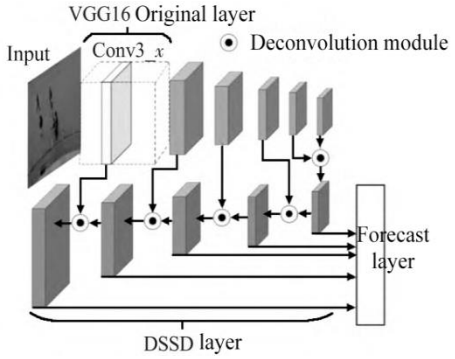

VGG16 Original layer
Input
Conv3_x
Deconvolution module
Forecast layer
DSSD layer

Fig.5. The schematic diagram of SSD network structure

4. Conclusion

Now, target detection is still a hot research direction in the field of artificial intelligence, and it still has huge development potential and broad application prospects. By using DHT11 module, the indoor temperature and humidity data can be real-time transmitted to the remote client.

The current target detection field still has very challenging problems, such as the lack of lightweight network model for mobile terminals; Lack of network model that can accurately detect small objects; Methods that can extract more levels and dimensions.

Acknowledgments

This paper is partly supported by the Education Reform Project (2021-JG-03) from the Teaching Guidance Committee of Electronic Information in Higher Education of the Ministry of Education, China in 2021, and the Graduate Education Reform and Innovation Project (2021YJCB02) from Tianjin University of Science and Technology, China in 2021.

References

1. GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580－587.
2. HE K M, ZHANG X Y, RRN S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition. European Conference on Computer Vision, Springer, Cham, 2014: 346－361.
3. GIRSHICK R. Fast R-CNN. Proceedings of the IEEE International Conference on Computer Vision, 2015: 1440－1448.
4. REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks. Advances in Neural Information Processing Systems, 2015: 91－99.
5. FU C Y, LIU W, RANG A, et a. DSSD: Deconvolutional single shot detecto. arXiv preprint, arXiv: 1701.06659, 2017.

Authors Introduction

Mr. Hucheng Wang

He is a second-year master candidate in Tianjin University of Science and Technology, majoring in machine learning.

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.

Ms. Min Zhao

She is a master student in Tianjin University of Science and Technology. Her research is about automatic and adaptive control.

A Design of New Air Ground Cooperative Unmanned Transportation System

Hucheng Wang*, Min Zhao

College of Electronic Information and Automation, Tianjin University of Science and Technology,
300222，China

*E-mail: *18322744268@163.com*
www.tust.edu.cn

Abstract

With the development of science and technology, the shortage of human resources in the labor market and the increase of human costs, intelligence, intelligence and unmanned logistics have become the mainstream development trend. This paper introduces a new type of space ground coordinated unmanned transportation system, which aims to solve the problems of liberating labor, reducing the operating costs of logistics enterprises, reducing the contact between people, and impeding the spread of the COVID-19 epidemic. This paper introduces the research background and advantages of the system, and describes the selection and working principle of the system in detail.

Keywords: UAV, UGV, Air ground coordination, Transportation

1. Introduction

With the shortage of the labor market and the rising labor costs, the traditional logistics has gradually become unmanned and intelligent. In recent years, UAVs have been widely popularized and used, especially by many large companies for express delivery and emergency response of the last kilometer of the city [1]. During the outbreak of COVID-19 in 2020, JD will use unmanned aerial vehicles to transport drugs and food in areas with serious epidemic. This will replace the traditional logistics blocked by the epidemic. The UAV carries a large number of sensors, making it multi-dimensional and highly flexible. Therefore, the UAV can provide mobile sensing resources with the same characteristics for mobile swarm intelligence [2][3].

Through the existing mobile awareness resources (UAV, UGV) to obtain awareness, and through the deployed mobile Internet to communicate, a new awareness mode - air ground cooperative mobile group intelligence awareness has been formed. Based on this technical background, the air ground cooperative new unmanned transportation system is developed.

This design has the following advantages:

On the technical level, overcome the characteristics of small load of UAV and limited sensing range of intelligent vehicle, and combine the advantages of large sensing range of UAV and strong loading capacity of intelligent vehicle.

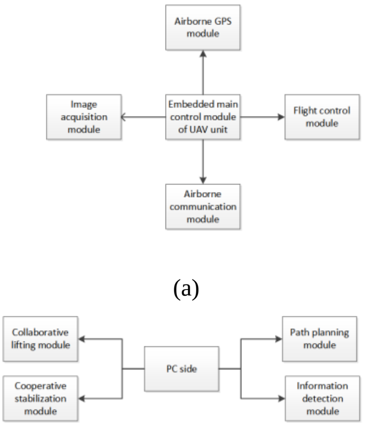
At the social level, the new unmanned transportation system of air ground coordination greatly reduces the participation of people in the distribution link, and reduces the probability of traffic accident risk during transportation while liberating productivity.

In the context of the epidemic, unmanned distribution can change the previous face-to-face logistics terminal distribution scenario, reduce personnel contact, and to some extent, block the spread of the epidemic.

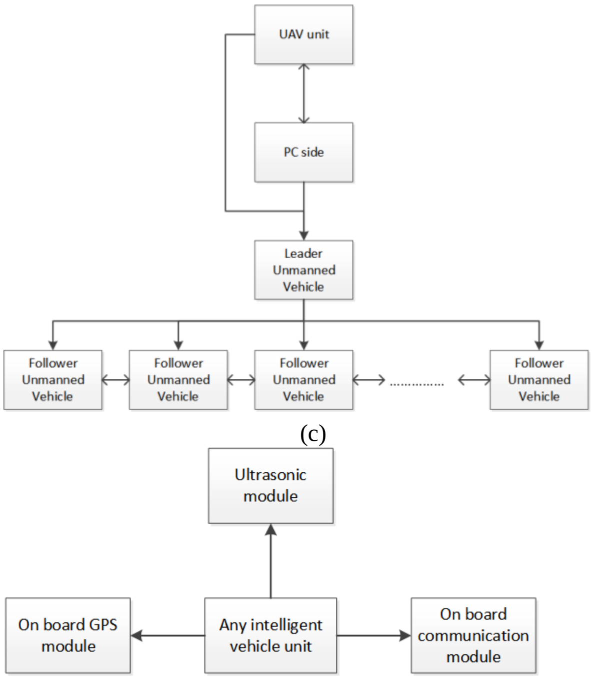
At the economic level, the new unmanned transportation system of air ground coordination can broaden the scope of logistics distribution space, fundamentally improve the efficiency of logistics distribution, control the consumption of human costs during logistics distribution, ensure that the logistics industry can better adapt to the needs of social modernization and improve the comprehensive benefits during logistics distribution.

2. Overall design scheme

The new air ground cooperative unmanned transportation system includes UAV unit, ground station unit and intelligent vehicle group. The UAV unit and the intelligent vehicle unit are connected with the ground station unit through the communication module. The UAV unit can communicate with each other and the UAV can transmit information to the leader intelligent vehicle unit. The UAV unit is mainly responsible for two aspects, one is to detect the position between the transportation target and the UAV, and assist the unmanned vehicle group to carry the transportation target; the other is to be responsible for transmitting the ground information around the unmanned vehicle group to the leader's intelligent vehicle during the transportation, and assist the intelligent vehicle group to perceive the ground information around to ensure safety. The ground station unit is responsible for monitoring the working status of the UAV and the intelligent vehicle group and selecting an optimal transportation route for the UAV and the intelligent vehicle group. The unmanned vehicle group is responsible for transporting goods along the optimal route selected by the ground station. The leader unmanned vehicle receives the transmission information from the ground station unit and the UAV unit to avoid obstacles. The follower unmanned vehicle makes adjustments according to the multi-agent consistency algorithm, so that the entire unmanned vehicle group can stably and quickly transport the target goods (overall structure Fig. 1 (a) (b) (c) (d)).


Airborne GPS module
Image acquisition module
Embedded main control module of UAV unit
Flight control module
Airborne communication module
(a)
Collaborative lifting module
Cooperative stabilization module
PC side
Path planning module
Information detection module

(b)


UAV unit
PC side
Leader Unmanned Vehicle
Follower Unmanned Vehicle
Follower Unmanned Vehicle
Follower Unmanned Vehicle
....
Follower Unmanned Vehicle
(c)
Ultrasonic module
On board GPS module
Any intelligent vehicle unit
On board communication module

(d)

Fig.1. Overall structure

3. Device selection

Device selection is shown in Table 1 and Table 2.

Table 1. Device Selection of UAV

UGV	Pixhawk four rotor aircraft
Battery	5200mah 3S lithium battery
Flight control module	Pixhawk 2.4.8 Flight control module
Camera lens	SJ5000
GPS module	NEO-M8N GPS module
Image processing module	STM32(1)+Opencv3.4.1
The wireless data transmission module is connected with the main control module.	

Table 2. Device Selection of UAV

Control module	STM32F103ZET6 microcontroller
On board communication module	STM32 (2) sets up LAN through router for communication between smart cars; The wireless data transmission module is connected with the ground communication module.
GPS module	NEO-M8N GPS module

4. Technical realization principle

4.1. Target cargo handling

The transportation target and destination are given. The ground station sends path information to the UAV and UAV according to the location of the transportation target, so that the UAV and UAV can fly near the transportation target.

The UAV identifies the transportation target through the image processing module to make the UAV fly above the center of the transportation target.

The UAV sends the image information of the transport target to the ground station through the wireless data transmission module, and the ground station commands the unmanned vehicle so that the midpoint of each side of the transport target is directly below, which is convenient for the subsequent stable transport of the transport target. And set the height..

The ground station sends a command to command the unmanned vehicle group to lift objects together. Fig. 2 is the schematic diagram of lifting scheme and Fig. 3 is the lifting scheme flow chart.

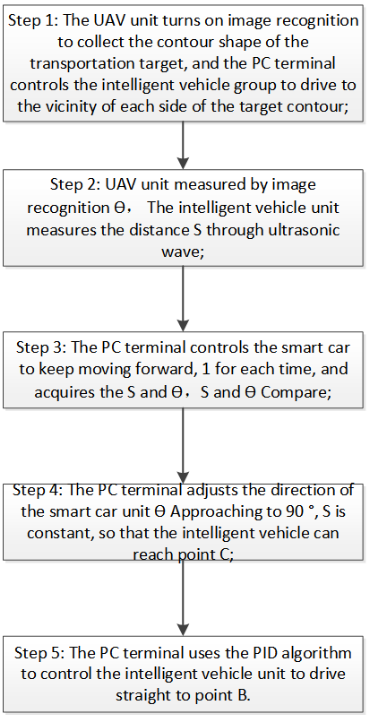

Step 1: The UAV unit turns on image recognition to collect the contour shape of the transportation target, and the PC terminal controls the intelligent vehicle group to drive to the vicinity of each side of the target contour;
Step 2: UAV unit measured by image recognition θ， The intelligent vehicle unit measures the distance S through ultrasonic wave;
Step 3: The PC terminal controls the smart car to keep moving forward, 1 for each time, and acquires the S and θ， S and θ Compare;
Step 4: The PC terminal adjusts the direction of the smart car unit θ Approaching to 90 °, S is constant, so that the intelligent vehicle can reach point C;
Step 5: The PC terminal uses the PID algorithm to control the intelligent vehicle unit to drive straight to point B.

Fig.2. The schematic diagram of lifting scheme

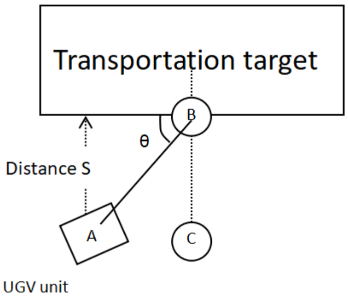

Transportation target
B
θ
Distance S
A
C
UGV unit

Fig.3. The lifting scheme flow chart

4.2. Target cargo transportation

The ground station unit plans an optimal path according to the map, integrating the location information of the transportation target and the location information of the transportation destination, and transmits it to the UAV and the unmanned vehicle group through the wireless data transmission module. The UAV and the unmanned vehicle group travel to the destination according to the optimal path.

Note: (1) In the transportation process, the UAV transmits the ground information to the leader's UAV. The leader's UAV avoids obstacles according to the ground information transmitted by the UAV, and the follower's UAV maintains the formation of the unmanned vehicle group according to the multi-agent consistency algorithm.

(2) The ground station unit monitors the operation status (operation speed, coordinate displacement of the intelligent vehicle set relative to the center of the transportation target, etc.) of the UAV and the intelligent vehicle set in real time, monitors the pressure between the UAV and the target cargo, and can intervene in the control of the UAV and the UAV at any time. The priority of the control of the UAV is higher than that of the UAV. Fig. 4 is the overall schematic diagram of the system.

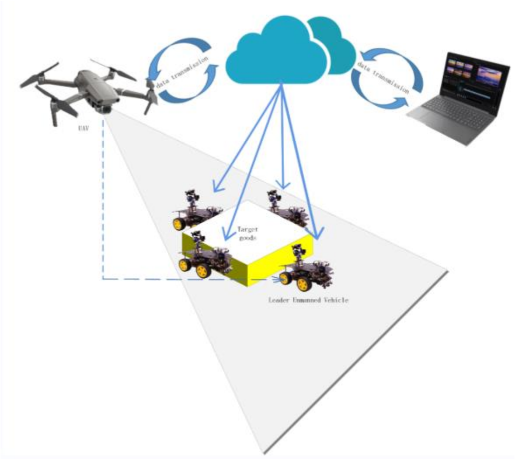


Fig.4. the overall schematic diagram of the system

4.3. Target goods release

When arriving at the destination, the unmanned vehicle group will run to the designated place, unload the goods at the same time, and the UAV and the unmanned vehicle will enter the standby state.

5. Conclusion

This paper designs a new type of air ground unmanned transportation system, and explains its working principle. Most of the current application modes are based on a single agent, and the system designed in this paper can fill the gap of space ground cooperative transportation. Considering the huge development prospect in the field of air ground coordination, I hope this design can provide some reference for technicians in related industries.

Acknowledgments

This paper is partly supported by the Education Reform Project (2021-JG-03) from the Teaching Guidance Committee of Electronic Information in Higher Education of the Ministry of Education, China in 2021, and the Graduate Education Reform and Innovation Project (2021YJCB02) from Tianjin University of Science and Technology, China in 2021.

References

1. DORLING K, HEINRICHS J, MESSIER G G, et al. Vehicle routing problems for drone delivery, IEEE Transactions on System, Man and Cybernetics: Systems, 2016, 47(1):70-85.
2. WANG X H, DUAN L J. Dynamic pricing and capacity allocation of UAV-provided mobile services, Proceedings of IEEE Conference on Computer Communications. IEEE, 2019: 1855-1863.
3. LIU C H, PIAO C Z, TANG J. Energy-efficient UAV crowdsensing with multiple charging stations by deep learning, Proceedings of IEEE Conference on Computer Communications. IEEE, 2020: 199-208.

Authors Introduction

Mr. Hucheng Wang

He is a second-year master candidate in Tianjin University of Science and Technology, majoring in machine learning.

Ms. Min Zhao

She is a master student in Tianjin University of Science and Technology. Her research is about automatic and adaptive control.

Attitude Solution of Quadrotor UAV

Siyuan Liu*, Zhihao Zhao, Haoran Gong
Tianjin University of Science and Technology, Tianjin, China

*E-mail:*lsy22934231542022@163.com*

www.tust.edu.cn

Abstract

In this paper, the attitude solution of a quadcopter UAV is introduced, and the integrated circuit designed by the UAV is verified. The angular velocity data and acceleration data required for this solution are provided by two modules: gyroscope and accelerometer. In this attitude control, three values are used: quaternions, rotation matrices, and Euler angles. Through data algorithm analysis, integrated circuit design and program design, the feasibility of the attitude algorithm is prove. On this basis, the problem of combining two systems: attitude flight and fixed height and fixed point is solved.

Keywords: Attitude solution, quaternion, Euler Angle, rotation matrix, gyroscope

1. Introduction

In order to do the attitude solution, we need to understand the basic idea of the solution. The data required for the calculation is provided by the following modules: Quad copter aircraft, also known as quad rotor aircraft, whose four propellers are simple mechanisms with motors directly connected. The cross-shaped layout allows the aircraft to obtain the force of rotating the fuselage by changing the motor speed, so as to adjust its own attitude.

Generally speaking, attitude means that we stand on the ground to observe the pitch/roll/heading state of the aircraft. The aircraft needs to know its current attitude in real time so that it can control its next actions as needed, such as keeping steady and rolling.

Mathematical model: Attitude is used to describe the angular position relationship between the fixed coordinate system and the reference coordinate system of a rigid body (attitude Angle), there are some mathematical representation methods. The common ones are Euler angles, quaternions, matrices, axis angles.

Applying the above mathematical model to the attitude description of aircraft, the reference coordinate system corresponds to the navigation coordinate system is fixed. We usually use the frame R. The fixed coordinate system corresponds to the body coordinate system, denoted by the coordinate system r. Then we can use Euler Angle, quaternion and other mathematical methods to describe the angular position relationship between r and R. This is the mathematical model for the attitude solution of the vehicle.

The core of the attitude solution of the quad copter is the control rotation. We generally use quaternion to represent the rotation. After obtaining the quaternion, it will be converted into Euler Angle and then input into the attitude control algorithm. In the design process of the quadcopter, we used the gyroscope and other inertia elements. Attitude calculation through this method is also a relatively complicated part of the control quadcopter.

In order to simplify the complexity of the technology, we apply MPU6050, which uses hardware DMP to read the quaternion directly. In other words, MPU6050 has a built-in function module, which can correct and process the original data before output [1].

This paper will study the quadcopter based on STM32 control center carrying MPU6050, and analyze from the aspects of hardware, algorithm and circuit design.

2. Attitude solution

2.1. *Basic idea of solution*

The gyroscope, which measures angular velocity, has high dynamic characteristics. It is a device for measuring angles indirectly. It measures the derivative of the Angle, the angular velocity, and you integrate the angular velocity over time to get the Angle. Due to the influence of noise and other errors, it accumulates continuously under the action of integration, and finally leads to the low frequency interference and drift of the gyroscope.

The direction the accelerometer outputs the current acceleration (including the gravitational acceleration). The high frequency signal sensitivity caused by the measurement principle makes the high frequency interference of the accelerometer in the vibration environment. Therefore, we can conclude that the gyroscope is sensitive to low frequency disturbance and the accelerometer is sensitive to high frequency disturbance. The gyroscope can be trusted in a short time, and the accelerometer needs to be trusted in a long time, so the idea is to hold the attitude calculated by the accelerometer to correct the gyroscope. Namely, complementary fusion, using the acceleration vector and the gyroscope angular velocity vector as the cross product, if there is no change between the two attitudes, the cross product result should be 0, but if there is a change, then the result can be used as the correction of the gyroscope angular velocity. Adding variables Kp and Ki can be used for PI correction to quickly correct the angular velocity value. A relatively accurate attitude value can be obtained by updating the quaternion with the compensated angular velocity.

The fusion of complementary proposed in this paper is shown in Fig.1.

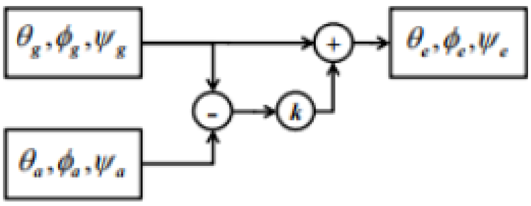

θ_g, ϕ_g, ψ_g
θ_e, ϕ_e, ψ_e
θ_a, ϕ_a, ψ_a
+
-
k

Fig.1 Fusion of complementary

2.2. *Quaternion,rotation matrix and Euler Angle*

In attitude control, we need to know three values: quaternion, rotation matrix and Euler Angle .

The relationship between quaternion, rotation matrix and Euler Angle proposed in this paper is shown in Fig.2.

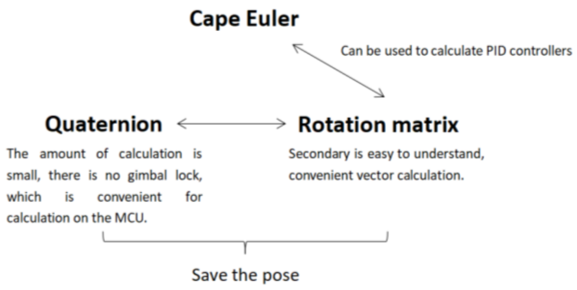

Cape Euler
Can be used to calculate PID controllers
Quaternion
Rotation matrix
The amount of calculation is small, there is no gimbal lock, which is convenient for calculation on the MCU.
Secondary is easy to understand, convenient vector calculation.
Save the pose

Fig.2 Quaternion rotation matrix and Euler Angle

(1) Definition of quaternion [2]

Quaternions were proposed by the Irish mathematician William Rowan Hamilton in 1843. Any rotation of three-dimensional space can be represented by an Angle rotated around an Axis of three-dimensional space, namely the so-called axis-angle representation method. In this representation, Axis can be represented by a three-dimensional vector (x,y,z), θ can be represented by an Angle value, and intuitively, a four-dimensional vector (θ,x,y,z) can represent any rotation in three dimensions. Note that the three-dimensional vector (x,y,z) here is used only to represent the orientation of the axis, so a more compact representation would be to use a unit vector to represent the direction axis and the length of that three-dimensional vector to represent the Angle value θ. Thus, a three-dimensional vector (θ*x,θ*y,θ*z) may be used to represent any rotation in three-dimensional space, provided that (x,y,z) is a unit vector.This is the way the Rotation Vector is represented, which is used extensively in Open CV to represent rotation (see rvec in the Open CV Camera Calibration section) [3].

(2) Euler Angle [4]

Static definition

Three independent coordinate variables are required to describe the configuration of a fixed point rotating rigid body. To describe the configuration of a fixed axis rotating rigid body requires an independent coordinate variable, namely the Angle. The process of fixed point motion is decomposed into three independent angles, namely Euler Angle, which is composed of nutation Angle, precession Angle and rotation Angle. The basic idea of Euler Angle is to decompose angular displacement into a sequence of three rotations about three mutually perpendicular axes. Euler Angle is used to describe the orientation of a rigid body in three-dimensional Euclidean space. For a frame of reference in three dimensions, the orientation of any coordinate system can be represented by three Euler angles. A frame of reference, also known as a laboratory frame of reference, is stationary. The coordinate system is fixed to the rigid body and rotates as the rigid body rotates.

Dynamic definition

There are two different dynamic definitions. One is the composition of three rotations about the coordinate axis fixed to the rigid body. The other is a composite of three rotations around the laboratory reference axis. Note that the XYZ axis is the rotating rigid body axis, while the xyz axis is the stationary laboratory reference axis.

Rotation about the XYZ axis: Initially, the axes of the two coordinate systems xyz and XYZ overlap, starting with a rotation of α about the Z axis, then a rotation of β about X, then a rotation of γ about the Z axis.

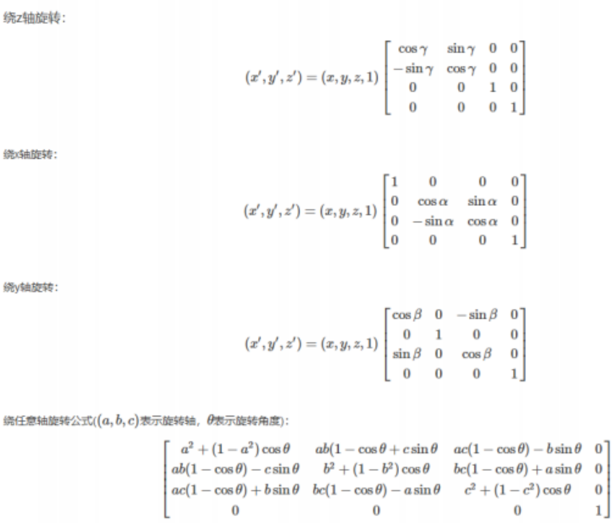
Rotation about the xyz axis: Initially, the axes of both coordinate systems xyz and XYZ overlap, starting with a rotation of gamma Angle about the z axis, then a rotation of alpha Angle about x, and finally a rotation of beta about the z axis.

When using Euler Angle, we need to know another concept: Gimbal Lock, referred to as gimbal lock or gimbal lock, is caused by the fact that two axes of the three gimbal joints overlap and will lose one degree of freedom. Result: Because of the existence of universal lock, Euler Angle cannot realize spherical smoothing interpolation. The solution is to use quaternion spherical linear interpolation [5].

(3)Matrix of rotation [6]

Rotations in three dimensions are more complex than in two dimensions. You need to specify the rotation Angle and axis. If the three coordinate axes xyz of the coordinate system are taken as the rotation axis, then the point actually only makes a two-dimensional transformation on the plane of the vertical coordinate axis, and the three-dimensional transformation matrix can be derived by using a two-dimensional formula. Stipulation: In the right hand coordinate system, the positive direction of rotation is the right hand spiral direction, that is, from the positive half axis of the axis to the origin of the counterclockwise direction.

The rotation matrix relation proposed in this paper is shown in Fig.3.



绕z轴旋转：

$$(x',y',z')=(x,y,z,1)\begin{bmatrix}\cos\gamma & \sin\gamma & 0 & 0\\ -\sin\gamma & \cos\gamma & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}$$

绕x轴旋转：

$$(x',y',z')=(x,y,z,1)\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & \cos\alpha & \sin\alpha & 0\\ 0 & -\sin\alpha & \cos\alpha & 0\\ 0 & 0 & 0 & 1\end{bmatrix}$$

绕y轴旋转：

$$(x',y',z')=(x,y,z,1)\begin{bmatrix}\cos\beta & 0 & -\sin\beta & 0\\ 0 & 1 & 0 & 0\\ \sin\beta & 0 & \cos\beta & 0\\ 0 & 0 & 0 & 1\end{bmatrix}$$

绕任意轴旋转公式((a,b,c)表示旋转轴，θ表示旋转角度)：

$$\begin{bmatrix}a^2+(1-a^2)\cos\theta & ab(1-\cos\theta+c\sin\theta & ac(1-\cos\theta)-b\sin\theta & 0\\ ab(1-\cos\theta)-c\sin\theta & b^2+(1-b^2)\cos\theta & bc(1-\cos\theta)+a\sin\theta & 0\\ ac(1-\cos\theta)+b\sin\theta & bc(1-\cos\theta)-a\sin\theta & c^2+(1-c^2)\cos\theta & 0\\ 0 & 0 & 0 & 1\end{bmatrix}$$

Fig.3 Rotation matrix relation

Rotation matrix, Euler Angle, quaternion comparison

Rotation matrix, Euler Angle, quaternion are mainly used for: vector rotation, conversion between coordinate systems, angular displacement calculation, azimuth smooth interpolation calculation.

Different azimuth representations are suitable for different situations:

Euler angles are the easiest to use. Euler angles can greatly simplify human-computer interaction when it comes to specifying the orientation of objects in the world, including typing the orientation directly from the keyboard, specifying the orientation in code (such as setting the camera for rendering), and testing during debugging.

If you need to convert vectors between coordinate systems, choose the matrix form. Another approach is to use Euler angles as a "master copy" of the azimuth, but maintain a rotation matrix that is updated whenever Euler angles change.

Euler angles or quaternons are used when large amounts of azimuth data (such as animations) need to be stored. Euler angles take up 25% less memory, but are slower to convert to the matrix. If animation data requires nested connections between coordinate systems, quaternons may be the best choice. Smooth interpolation can only use quaternions. In other forms, you must go to the quaternion, and then go back after interpolation.

2.3. *Attitude update frequency and practical application*

In the calculation, we are actually updating the attitude, and then put the result of Euler Angle into the controller to calculate the control result of the motor.

The process involved in this article is shown in Fig.4.

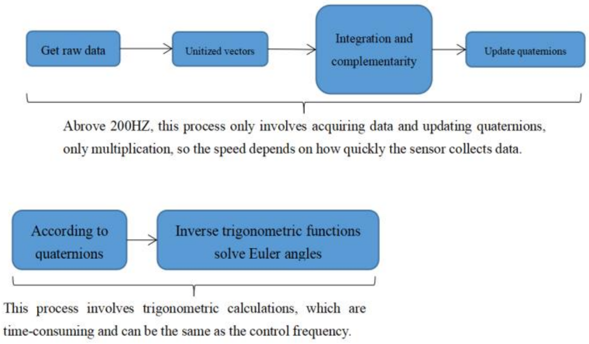

Get raw data
Unitized vectors
Integration and complementarity
Update quaternions
Abrove 200HZ, this process only involves acquiring data and updating quaternions, only multiplication, so the speed depends on how quickly the sensor collects data.
According to quaternions
Inverse trigonometric functions solve Euler angles
This process involves trigonometric calculations, which are time-consuming and can be the same as the control frequency.

Fig.4 Actual application process

3. UAV programming

3.1. *Complete Framework Introduction*

For the complete framework of the entire flight control source code, we can get the following information according to the following figure:

The remote control input rudder information into the controller, at the same time, the controller can also give itself control rudder, in line control and other programs often use this function;

The sensor can input the attitude values of the accelerometer, barometer and gyroscope into the IMU solving unit, so that we can get a relatively accurate position of the UAV in space;

After the IMU is solved, the controller will control the motor according to the settlement result returned by it.

The flow chart covered in this paragraph is shown in Fig.5.

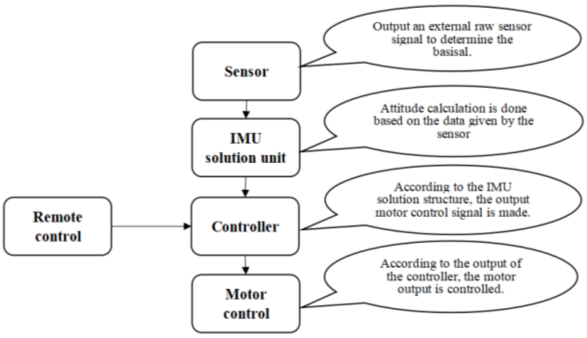

Sensor
Output an external raw sensor signal to determine the basisal.
IMU solution unit
Attitude calculation is done based on the data given by the sensor
Remote control
Controller
According to the IMU solution structure, the output motor control signal is made.
Motor control
According to the output of the controller, the motor output is controlled.

Fig.5 flow chart

3.2. *Remote control module frame*

According to Fig.6, we can learn the framework information about the remote control module.

In the attitude mode, the engine rudder directly controls the speed of the motor and flies high as much as it pushes.

In the fixed-point mode, the engine rudder only controls the flight altitude information.

The Pitch rudder controls the pitch Angle;

The Roll rudder controls the roll Angle;

The Yaw rudder controls the direction of the drone.

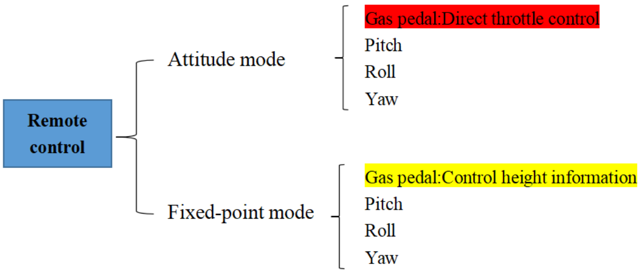

Remote control
Attitude mode
Gas pedal:Direct throttle control
Pitch
Roll
Yaw
Fixed-point mode
Gas pedal:Control height information
Pitch
Roll
Yaw

Fig.6 The frame information of the remote control module

3.3. *Sensor module frame*

According to Fig.7, we can see the information of the sensor module framework.

The attitude sensor used in IMU calculation is MPU6050;

The height sensor uses infrared laser sensor (measuring range of about 2m, recommended flight 1m~1.5m) and barometer.Data are prone to inaccuracies.

The horizontal sensor uses the optical flow sensor to measure the velocity value of the UAV in all directions, which can be combined with the IMU solution.

The unit measures more accurate horizontal velocity values;

The module used by the line sensor is OpenMV (more functional than other similar modules on the market).

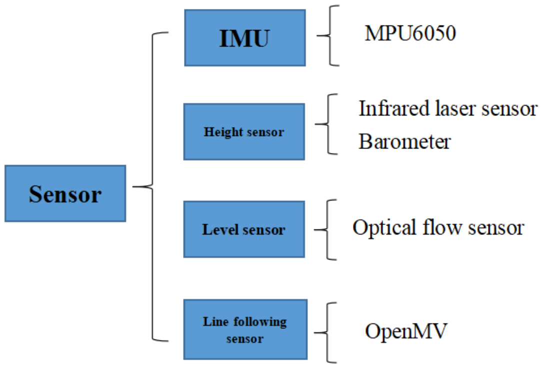

Sensor
IMU
MPU6050
Height sensor
Infrared laser sensor
Barometer
Level sensor
Optical flow sensor
Line following sensor
OpenMV

Fig.7 Sensor module frame information

3.4. *Introduction to IMU solution unit framework*

According to Fig.8, we can obtain the following information about IMU solving unit:

Acceleration information is used to correct gyroscope;

Gyroscope information is used to be calibrated;

The temperature information is used for thermostatic control.

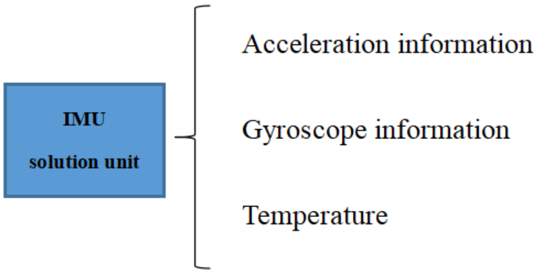

IMU
solution unit
Acceleration information
Gyroscope information
Temperature

Fig.8 IMU solves the information of the unit

3.5. *Controller module framework*

According to Fig.9, information about the controller can be obtained as follows:

Attitude controller is the most basic controller to ensure the UAV has a relatively stable flight. It is a controller with two ring PID.

The height controller controls the height stability of the UAV when the height is fixed.

The horizontal controller controls the stability of the UAV in the horizontal direction;

The line controller uses the data from OpenMV to control the flight of the line.

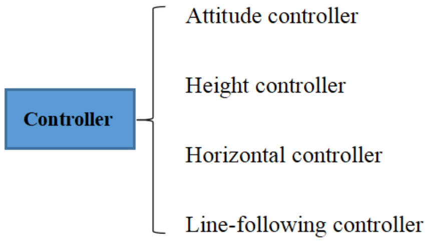

Controller
Attitude controller
Height controller
Horizontal controller
Line-following controller

Fig.9 Controller Information

3.6. *Motor control module frame*

According to Fig.10, we can get the following information about motor control:

In UAV control, we use 400HZ PWM wave to control the motor (because the solution speed is not less than 100HZ, not more than 200HZ).

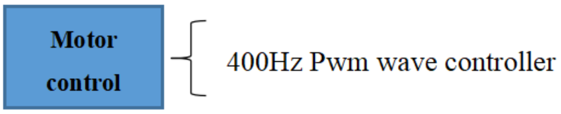

Motor
control
400Hz Pwm wave controller

Fig.10 Information on motor control

4. Algorithm analysis [7]

In the design, the MCU timer peripheral output PWM signal to the electric modulation, the electric modulation control motor rotation to produce lift, so that you can fly. If the motor and the frame are completely symmetric without any error, it is a linear system. The diagonal of the motor rotates in the same direction, and the main and auxiliary diagonal lines will generate torsion forces in opposite directions to each other. It is assumed that the quadrotor is a linear system. You increase the voltage and you increase the lift and you fly. In theory, only a pair of diagonal motors can fly. However, the vehicle will rotate in one direction. Then the movement of the quadrotor is very simple, increase the speed of the motor 12, produce pitching motion, pitch forward, increase the speed of the motor 14, produce roll motion, roll to the right, increase the speed of the motor 13, is to produce yaw motion, clockwise yaw. Quadrotor aircraft can not be a linear system, there will be deviation, so we need to use the sensor to measure the deviation, as the system observation quantity, and then through the control algorithm of the flight control code to achieve control.

The difficulty in achieving attitude flight and altitude fixing is actually the combination of these two systems, which requires fitting together the offset data of these sensors on the XYZ axis. The fixed height point corresponds to three axes in the navigation coordinate system. The data of the data accelerometer, optical flow, gyroscope, accelerometer and gyroscope on the X-axis and Y-axis are mapped to the navigation coordinate system. After optical flow rotation, the fused data on the XY axis can be obtained by fitting the data of the accelerometer. Similarly, on the Z-axis, the laser ranging module can simply obtain the height information through the differential pressure method. After fitting with the Z-axis data of the accelerometer, the Z-axis data after fusion can be obtained. The fusion uses a third order complementary filtering method. Take the z-axis of the accelerometer as an example. If the complementary filtering method is not used, the system is an open-loop system. The position information can be obtained after two integrations of the Z-axis accelerometer. The poles under pure integration are zero, and the system is divergent and cannot be stabilized. The system block diagram of the third-order complementary filtering shown in the figure below is the position, velocity and acceleration respectively, with three appropriate values. And here are three third-order complementary systems, the XYZ axis.

The above flow chart is shown in Fig.11.

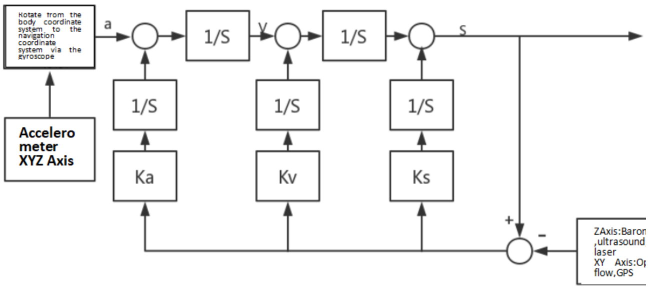

Rotate from the body coordinate system to the navigation coordinate system via the gyroscope
Accelero meter XYZ Axis
a
1/S
v
1/S
s
1/S
1/S
1/S
Ka
Kv
Ks
+
-
ZAxis:Baron ,ultrasound laser XY Axis:Op flow,GPS

Fig.11 Flow chart

The controller used in this design is a cascade PID controller, the input is the ideal flight height, the output is PWM, drive motor, because the quadrotor is not a linear system, in order to minimize the possible deviation of the aircraft, using the current height minus the expected height is the deviation, the deviation is fed back to the input and mapped to the PWM output, forming a closed-loop system. Adjust these three parameters, you can achieve a stable state. We can use MATLAB simulation debugging, in fact, is to make the system transfer function stability of the optimal solution. We understand the unipolar system, then the cascade system is very simple, is the output of a PID controller as the input of another PID controller, the final output control quantity to the controlled object, the PID controller inside belongs to the inner ring, the outer controller belongs to the outer ring, that is, the output of the outer ring as the input of the inner ring, the final output of the inner ring to the controlled object. In a cascade system, the parameters of the inner ring are more important than those of the outer ring. If the parameters of the inner ring are properly adjusted, the parameters of the outer ring are relatively easy to solve.

The above flow chart is shown in Fig.12.

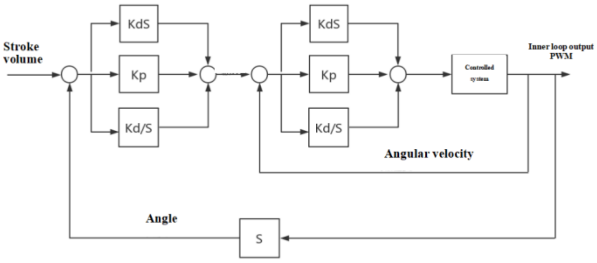

Stroke volume
KdS
Kp
Kd/S
KdS
Kp
Kd/S
Controlled system
Inner loop output PWM
Angular velocity
Angle
S

Fig.12 Flow chart

The following is the program design block diagram of this research. We adopt the timer time division multiplexing method, which is called in different time threads according to the update time required by different peripherals, different algorithms and different sensors.

The above flow chart is shown in Fig.13.

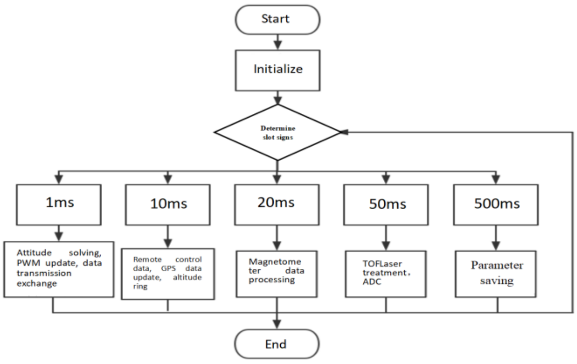

Start
Initialize
Determine slot signs
1ms
10ms
20ms
50ms
500ms
Attitude solving, PWM update, data transmission exchange
Remote control data, GPS data update, altitude ring
Magnetometer data processing
TOFLaser treatment, ADC
Parameter saving
End

Fig.13 Flow chart

Acknowledgments

This paper is partly supported by the Education Reform Project (2021-JG-03) from the Teaching Guidance Committee of Electronic Information in Higher Education of the Ministry of Education, China in 2021, and the Graduate Education Reform and Innovation

Project (2021YJCB02) from Tianjin University of Science and Technology, China in 2021.

References

1. CHEN Guo-ding，ZHOU Peng-hao，HU Zhen-hao，etal． Four-axis hardware attitude calculation based on MPU6050[J]. Journal of Mechanical & Electrical Engineering，2017，35(1) : 95 — 100.
2. Liao Kun male. Four rotor unmanned aerial vehicle (uav) attitude algorithm and the control method research [D]. Southwest university of science and technology, 2022. The DOI: 10.27415 /, dc nki. GXNGC. 2022.000468.
3. Lu Zhe-jun. Based on the four rotor unmanned aerial vehicle (uav) attitude algorithm research [D]. Shanghai university of engineering science, 2021. The DOI: 10.27715 /, dc nki. GSHGJ. 2021.000632.
4. Dong Changjun. Research on Attitude Solution Method and Flight Control System Implementation of Quadrotor Aircraft [D]. Suzhou university, 2021. DOI: 10.27351 /, dc nki. Gszhu. 2021.001239.
5. Zhang yan. Four rotor unmanned aerial vehicles (uavs) based on multi-sensor attitude algorithm [D]. Harbin university of science and technology, 2022. The DOI: 10.27063 /, dc nki. Ghlgu. 2022.000379.
6. WU Xiaochu. Research on Flight Control System of Quadrotor UAV [D]. Dalian University of Technology,2018.
7. Yuan Lao Hu, Lian Dongshan, Lu Lianjie, Liu Yi. Research on Attitude Solution Algorithm of quadrotor [J]. Journal of Shenyang Aerospace University,20,37(03):1-11.

Authors Introduction

Mr. Siyuan Liu

He will receive his bachelor's degree from Tianjin University of Science and Technology in China in 2024. His research interests include robot dynamics and attitude calculation of unmanned aerial vehicles.

Mr. Zhihao Zhao

His bachelor's degree is in intelligent science and advanced manufacturing, School of Electronic Information and Automation, Tianjin University of Science and Technology, majoring in artificial intelligence, proficient in C language, proficient in using Solid works, CAD and other software, very interested in machine vision research.

Mr. Haoran Gong

He participated in the American College Mathematical Contest in Modeling as a modeler and data analyst in early 2022 and won the s Award. During his freshman year, he published one international CPCI paper. In 2022, I will attend the "Hi cool" Global Entrepreneur Summit with my mentor.

A Four-dimensional Conservative Chaotic System and Its Application in Image Encryption

Wei Li, Jingwen Liu[#]**, Hongyan Jia**[*]

College of Electronic Information and Automation, Tianjin University of Science and Technology,
300222, China

E-mail: jiahy@tust.edu.cn

[#]*These authors contributed equally to this work.*

www.tust.edu.cn

Abstract

In this paper, a four-dimensional conservative chaotic system is firstly analyzed and investigated. It is found that the four-dimensional conservative chaotic system shows some complex dynamics, such as multi-stability and strong pseudo-randomness. Secondly, based on pseudo-random sequences and two-dimensional discrete wavelet transform, an image encryption algorithm is realized. Finally, all the experiment results and the security analysis show that the algorithm show good encryption characteristics, which further prove the image encryption algorithm.

Keywords: Hamiltonian chaotic system, Pseudo-randomness, Image encryption, Wavelet transform

1. Introduction

With the rapid development of computer and network technology, multimedia information occupies an increasing proportion in the whole human digital communication. The necessary encryption and protection of private digital information has become a growing concern. Compared with text information, image data has unique characteristics such as large amount of data, high redundancy and strong correlation between adjacent pixels. These characteristics make the traditional encryption algorithms used for text encryption no longer suitable for image encryption[1,2]. Therefore, it is urgent to find a fast and secure encryption algorithm for image encryption[3-7].

Chaotic system has good pseudo-randomness, ergonomic, unpredictability and sensitivity to initial values and parameters and other unique characteristics[8-13]. Therefore, these characteristics make chaotic system very suitable for image encryption. At first, J. Matthews and A. Robert[14] put forward the concept of "chaos cipher" in 1989. Subsequently, many cryptography schemes are proposed based on chaos[15-30]. For example, Jeri Fridrich used reversible two-dimensional chaotic maps on a torus or square to create an image encryption algorithm for new symmetric blocks[15]. Wang et al. proposed an image encryption algorithm based on fractional-order one-dimensional chaotic mapping with large chaotic space[7]. Chai et al. proposed a block-obfuscated image encryption algorithm based on three-dimensional Brownian motion, using Logistic Tent system to generate the direction of motion of particles, and introduce block-obfuscated image based on position sequence group[19]. Wang et al. proposed an image encryption algorithm based on hidden attractor chaotic system and Knuth Seinfeld algorithm[30]. However, the proposed scheme is mainly designed for some grayscale images, which require that the color images and multimedia data must first be converted to the same mode as the grayscale images, and then the scheme can be used for encryption. All above image encryption algorithms are based on dissipate chaotic system (DCS), few image encryption algorithms based on conservative chaotic system (CCS) have been proposed. Although DCS may have good pseudo-randomness, it produces singular attractors in the fractal dimension. Therefore, it is easy to be attacked by reconstructing the attractor. At the same time, most of the orbit around the attractor is unreachable, so the ergodic property of DCS is poor. Compared with DCS, CCS does not generate attractors, which ensures that the attacker cannot reconstruct the attractor and crack the encryption scheme. Meanwhile, the dimension of CCS is consistent with that of the system, and it has better ergodicity than DCS.

In this paper, a conservative chaotic system is firstly analyzed. Subsequently, based on the conservative chaotic system and two-dimensional discrete wavelet transform, a new image encryption algorithm is proposed. Finally, the new image encryption algorithm is verified, and all the results are analyzed from the aspects of statistical analysis, difference analysis and speed. It shows that the algorithm has good security performance and operation speed, and significantly improves the key space.

2. A Four-dimensional Conservative Chaotic System

By studying the method of constructing four-dimensional conservative chaotic system by Qi et al., a new four-dimensional conservative chaotic system is constructed on the basis of the method, which can be described as:

$$\begin{cases}\dot{x}_1=(\Pi_4-\Pi_2)x_2x_4+(\Pi_4-\Pi_3)x_3x_4\\ \dot{x}_2=(\Pi_1-\Pi_4)x_1x_4+c\Pi_3x_3\\ \dot{x}_3=(\Pi_1-\Pi_4)x_1x_4-c\Pi_2x_2\\ \dot{x}_4=(\Pi_2-\Pi_1)x_1x_2+(\Pi_3-\Pi_1)x_1x_3\end{cases}\tag{1}$$

where x_1, x_2, x_3, x_4 are state variables, Π_1, Π_2, Π_3, and c are system variables. When letting $c=0$, $(\Pi_1,\Pi_2,\Pi_3,\Pi_4)=(5,6,7,8)$, and choosing different initial values and system parameters, the system can both rich periodic dynamics and chaotic dynamics. In order to further study the reasons for chaotic dynamics of the system (1) from the perspective of energy, it is first convert into Kolmogorov form as follows:

$$\dot{x}=J(x)\nabla H(x)=\begin{bmatrix}0 & -x_4 & -x_4 & x_2+x_3\\ x_4 & 0 & c & -x_1\\ x_4 & -c & 0 & -x_1\\ -x_2-x_3 & x_1 & x_1 & 0\end{bmatrix}\begin{bmatrix}\Pi_1x_1\\ \Pi_2x_2\\ \Pi_3x_3\\ \Pi_4x_4\end{bmatrix}\tag{2}$$

Where, $J(x)=\begin{bmatrix}0 & -x_4 & -x_4 & x_2+x_3\\ x_4 & 0 & c & -x_1\\ x_4 & -c & 0 & -x_1\\ -x_2-x_3 & x_1 & x_1 & 0\end{bmatrix}$, $H(x)=\frac{1}{2}(\Pi_1x_1^2+\Pi_2x_2^2+\Pi_3x_3^2+\Pi_4x_4^2)$. It is found that the Jacoby matrix of system (1) is an antisymmetric matrix.

Then keep all the state variables and system variables same as the above except for $c=3$, when changing initial value x_1, the Lyapunov exponent diagram of system (1) is shown in Fig. 1. It can be found when $x_1\in(-1.8,3.7)$, the system (1) shows a periodic or counter-periodic dynamics; when $x_1\in(-20,-1.8)\cup(3.7,20)$, the system (1) shows a chaotic dynamics. Next, set system variables $(\Pi_1,\Pi_2,\Pi_3,\Pi_4,c)=(5,6,7,8,3)$, when selecting initial values $(0,5,-5,-3)$, system (1) shows a quasi-periodic attractor, when selecting initial values $(10,5,-5,-3)$ system (1) shows a chaotic attractor, respectively, as shown in Fig. 2.

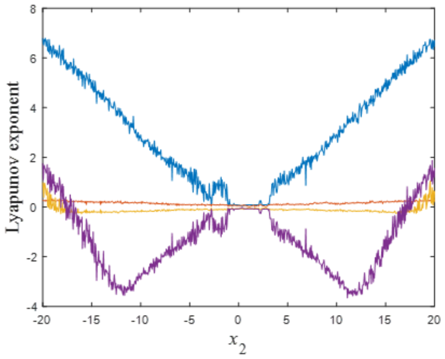
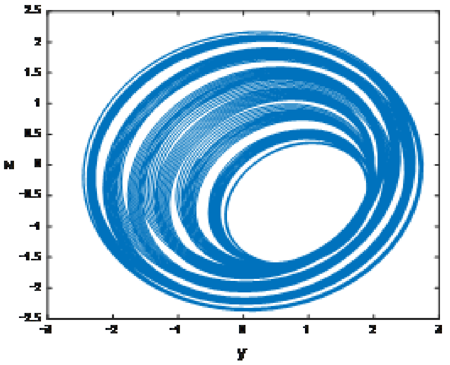
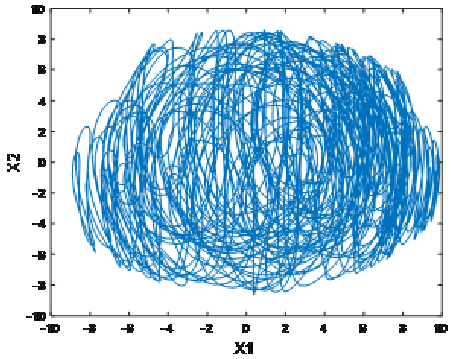


Fig. 1 Lyapunov exponent diagram of system (1)



(a) quasi-periodic attractor



(b) chaotic attractor

Fig. 2 Phase diagrams of system (1) when $c=3$

3. Image Encryption Algorithm

3.1 Encryption process

In this paper, the initial parameters of the system (1) are used as the encryption algorithm keys, and the plain image can be scrambled and spread by pseudo-random sequences generated by the system (1). In addition, during the scrambling operation, two-dimensional discrete wavelet is used to transform and extract the low-frequency part of the image. Here, only the low-frequency part is scrambled to improve the calculation speed of the encryption algorithm, as shown in Fig. 3.

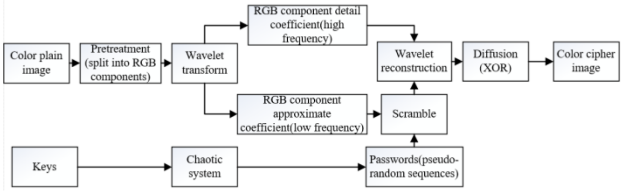

Color plain image
Pretreatment (split into RGB components)
Wavelet transform
RGB component detail coefficient(high frequency)
RGB component approximate coefficient(low frequency)
Wavelet reconstruction
Scramble
Diffusion (XOR)
Color cipher image
Keys
Chaotic system
Passwords(pseudo-random sequences)

Fig. 3 Encryption algorithm flowchart

The specific steps are as follows:

Step 1: To ensure the validity of the sequence, delete the first 2000 values of pseudo-random sequences $x-value$, $y-value$, $z-value$, $w-value$ generated by the system (1), and start counting from the 2001th value;

Step 2: Decompose the color picture I into three components, R, G, and B, denoted as I_R, I_G, I_B respectively;

Step 3: Extract the approximate coefficients and detail coefficients (horizontal coefficient ch1, vertical coefficient cv1, diagonal coefficient cd1) using the "db1" wavelet basis functions for the three components of I_R, I_G, I_B. Taking the R component as an example, the extracted two-dimensional array is marked as $R-ca1(i,j)$ $(i=1,2,...,\frac{m}{2}, j=1,2,...,\frac{n}{2})$, and the extracted two-dimensional array is converted into a one-dimensional column vector $R-ca1(k)$ $(k=1,2,...,\frac{m}{2}?\frac{n}{2})$ by column;

Step 4: Calculate from the 2001th value of $w-value$, and select $\frac{m}{2}\acute{} \frac{n}{2}$ values in turn, denoted as $w-value(p)$ $(p=1,2,...,\frac{m}{2}?\frac{n}{2})$. Sort $w-value(p)$ from small to large, and record the sorted row vector as $w_1-value(p\text{¢})$ $(p\text{¢}=1,2,...,\frac{m}{2}?\frac{n}{2})$. If $p=k$, $R-ca1\text{ïi}(p)=R-ca1(k)$, forming a new column vector $R-ca1\text{ïi}(p)$. Finally, it is transformed into $\frac{m}{2}\acute{} \frac{n}{2}$ matrix through the "reshape" function, so as to realize the scrambling of the approximate coefficients ca1 of the R, G, and B components;

Step 5: Perform wavelet reconstruction on the approximate coefficients of the R, G, and B components after scrambling in Step 4 and the detail coefficients of the image before encryption, and finally get the R, G, and B components after scrambling, denoted as I_r, I_g, I_b;

Step 6: Calculate from the 2001th value of $x-value$, $y-value$, $z-value$, $w-value$, select $m\acute{}\ n$ values in turn, mark them as $x-value(p)$, $y-value(p)$, $z-value(p)$ $(p=1,2,...,m?n)$, and perform the operation of equation (3-1), so that all elements in these three vectors are in the range of $[0,255]$.

$$y_n = \mathrm{mod}(10000?y_n, 256) \tag{2}$$

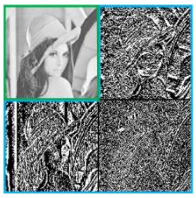
Step 7: Perform a bit wise XOR operation for $x-value(p)$, $y-value(p)$, $z-value(p)$ generated in Step 6 with the elements in the I_r, I_g, I_b matrix, and finally get the encrypted R, G, B components, denoted as $I\text{¢}_r, I\text{¢}_g, I\text{¢}_b$; ;

Step 8: Reconstruct $I\text{¢}_r, I\text{¢}_g, I\text{¢}_b$ component generated in step 7, and finally get the color cipher text image I'.

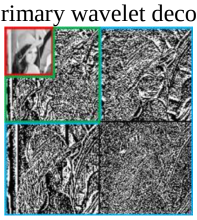
3.2 2D discrete wavelet transform

In this paper, 2D discrete wavelet transform can be used to optimize the speed of the image algorithm. By 2D discrete wavelet transform, the image can be divided into low frequency components I_{LL}, high frequency components I_{HH}, vertical components I_{LH}, diagonal components I_{HL}. The R component of Lena after secondary discrete wavelet transform is shown in Fig. 4. It can be seen that most information of the image is concentrated in the upper left corner, i.e., I_{LL},, while the other parts contain little image information, which can be ignored. Therefore, in image encryption, only I_{LL}, of the image can be encrypted, greatly improving the speed of encryption.

(a) before wavelet transform



(b) after primary wavelet decomposition

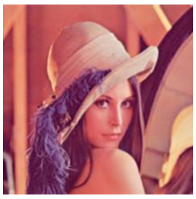


(c) after secondary wavelet decomposition

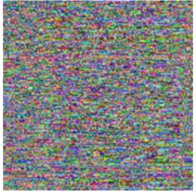
Fig. 4 *R* component of Lena image before and after wavelet transform

3.1 Experiment results

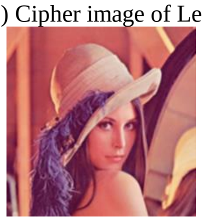
In this subsection, an experiment for the image encryption algorithm is done by using some frequently-used color images whose size are 512×512, such as Lena, as shown in Fig. 5 (a). Set key as $(\Pi_1, \Pi_2, \Pi_3, \Pi_4, c) = (5,6,7,8,3)$ and $(x_1, x_2, x_3, x_4) = (3,5,-5,-3)$, respectively, and run the image encryption algorithm, the cipher image is obtained, as shown in Fig. 5 (b). Use the same keys, the decrypt images are also obtained by running decryption algorithm, as shown in Fig. 5 (c). It can be found that the decrypt images are consistent with the plain images, which shows the image encryption algorithm proposed in this paper is effective.



(a) Lena



(b) Cipher image of Lena



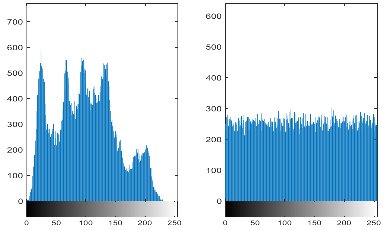
(c) Decrypt image of Lena

Fig. 5 Experiment results

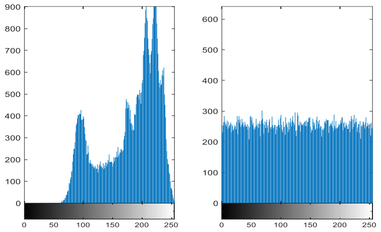
4. Security Analysis of Encryption Algorithm

4.1 Histogram analysis

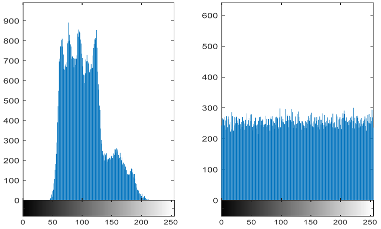
The image histogram represents the distribution of pixel intensity values in the image. When the histogram of the image is flat and there is no fluctuation trend, it indicates that the encryption algorithm can resist the statistical attack well. Fig.6 shows the histogram of each component of the original image before and after encryption. It can be seen that after the encryption process, the original image with uneven histogram distribution is transformed into a cartographic image with uniform histogram distribution. The results also prove that the image encryption algorithm can resist statistical attacks very well.



(a) R component



(b) G component



（c） B component

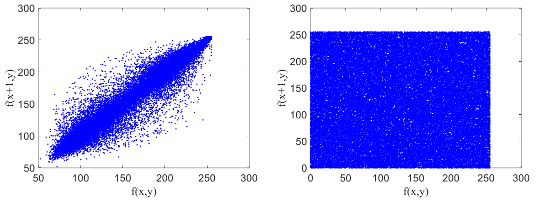
Fig. 6 Histogram of the plain image and the cipher image

4.2 Correlation analysis of adjacent pixels

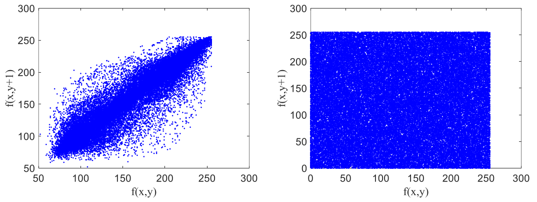
The correlation between adjacent pixels is one of the important indicators to judge the security of the encryption algorithm. In this paper, a row and a column of pixels are selected from the original image and the encrypted image respectively, and the correlation coefficients between adjacent pixels are shown in Table 1. As can be seen from Table 1, in the plain image, pixels in the vertical, horizontal and diagonal directions are strongly correlated, and the correlation coefficient is basically close to 1. However, the correlation between adjacent pixels in the cipher image is relatively small, and the correlation coefficient is almost close to 0, indicating that weak correlation between adjacent pixels. In addition, correlation of the R component in the plain image and the cipher image is also given to further show difference between them, as shown in Fig. 7. It can be found that the distribution of adjacent pixels in the plain image is highly concentrated, which means a strong correlation. Whereas the distribution of adjacent pixels in the cipher image is random, it means a weak correlation.

Table 1 Corresponding correlation coefficients of plain image and cipher image

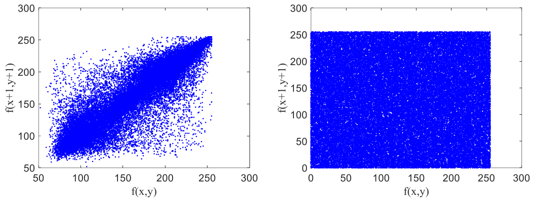
image	Plain image			Cipher image		
	R	G	B	R	G	B
Horizontal	0.97757	0.97161	0.95435	0.000756	0.007529	0.004363
Vertical	0.95550	0.94387	0.92538	0.002360	0.002484	0.001545
Diagonal	0.93215	0.91972	0.89641	0.004159	0.002781	0.000881



(a) Horizontal direction of R component before and after encryption



(b) Vertical direction of R component before and after encryption



(c) Diagonal direction of R component before and after encryption

Fig. 7 Correlation between adjacent pixels of R component

4.3 Information entropy analysis

Information entropy is used to characterize the strength of randomness of the system. For an image encryption algorithm with superior performance, the information entropy of the cipher image should be very close to 8. Table 2 shows the calculation results of the information entropy of the plain image and the cipher image. It can be seen that after encryption, the information entropy of the cipher image is closer to 8. Therefore, the algorithm has better information entropy characteristics, strong randomness, and good security.

Table 2 Information entropy of original and encrypted image

Test images	R component	R component	R component
Original image	7.1520	7.2598	6.9110
Cipher image	7.9982	7.9995	7.9986

4.4 Differential attack analysis

For encryption algorithms, NPCR (pixel change rate) and UACI (uniform average change intensity) are usually used to evaluate if they can resist differential attacks well. Generally speaking, the proposed image encryption algorithm can resist differential attacks well, when NPCR is close to 1 and UACI is close to 0.334. NPCR and UACI calculated based on the proposed encryption algorithm are shown in Table 3. It can be seen NPCR of each component is close to 1, and UACI is close to 0.334. Therefore, the algorithm can effectively resist differential attacks.

Table 3 Test results of NPCR and UACI

Index	R component	R component	R component
UACI	0.3365	0.3345	0.3336
NPCR	0.9978	0.9975	0.9967

5. Conclusion

This paper proposes a new image encryption scheme based on a four-dimensional conservative chaotic system and two-dimensional discrete wavelet transform. The four-dimensional conservative chaotic system is used to provide keys in the scrambling and diffusion process, two-dimensional discrete small transform is used to separate the high-frequency coefficients and low-frequency coefficients of the image, respectively. The encryption algorithm has passed various security tests and has strong reliability and security, which can provide technical preparation for the application of the encryption algorithm in confidential communication and data hiding.

References

1. Silva-Garcí a, V.M, Flores-Carapia R, Rentería-Márquez, C, et al. Substitution box generation using Chaos: An image encryption application. *Applied Mathematicsand Computation*, 2018, 332:123-135.
2. Tang H, Sun Q, Yang X, et al. A network coding and DES based dynamic encryption scheme for moving target defense. *IEEE Access*, 2018, 6: 26059-26068.
3. Tang Y, Abdul Jalil M Khalaf, Karthikeyan Rajagopal, et al. A new nonlinear oscillator with infinite number of coexisting hidden and self-excited attractors. *Chin. Phys. B*, 2018, 27: 040502.
4. Liu S, Guo C, John T. Sheridan. A review of optical image encryption techniques. *Opt. Laser Technol*, 2014, 57: 327-342.
5. D. Coppersmith. The Data Encryption Standard (DES) and its strength against attacks. *IBM J. Res. Dev*, 1994, 38: 243-250.
6. Liu H, Wang X. Color image encryption using spatial bit-level permutation and high-dimension chaotic system. *Opt. Commun*, 2011, 284: 3895-3903.
7. Wang L, Song H and Liu P. Sheridan. A novel hybrid color image encryption algorithm using two complex chaotic systems. *Opt. Lasers Eng*, 2016, 77: 118-125.
8. Lin H, Wang C. Influences of electromagnetic radiation distribution on chaotic dynamics of a neural network. *Applied Mathematics and Computation*, 2020, 369: 124840.
9. Zhao Q, Wang C, Zhang X. A universal emulator for memristor, memcapacitor, and meminductor and its chaotic circuit. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 2019, 29(1): 013141.
10. Zhang X, Wang C. Multiscroll hyperchaotic system with hidden attractors and its circuit implementation. *International Journal of Bifurcation and Chaos*, 2019, 29(09): 1950117.
11. Zhang X, Wang C, Yao W, et al. Chaotic system with bondorbital attractors. *Nonlinear Dynamics*, 2019, 97(4): 2159-2174.
12. Deng Q, Wang C. Multi-scroll hidden attractors with two stable equilibrium points. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 2019, 29(9): 093112.
13. Yu F, Liu L, He B, et al. Analysis and FPGA realization of a novel 5D hyperchaotic four-wing memristive system, active control synchronization, and secure communication application. *Complexity*, 2019, 2019.
14. Robert A, J Matthews. On the derivation of a chaotic encryption algorithm. *Cryptologia*, 1989, 13: 29-42.
15. Jiri Fridrich. Symmetric Ciphers Based on Two-Dimensional Chaotic Maps. *Int. J. bifurc. chaos*, 1998, 8: 1259-1284.
16. Tan J, Luo Y, Zhou Z, et al. Combined Effect of Classical Chaos and Quantum Resonance on Entan glement Dynamics. *Chin. Phys. Lett*, 2016, 33: 070302.
17. Mohamed Zakariya Talhaou, Wang X. A new fractional one dimensional chaotic map and its application in high-speed Image Encryption. *Signal Processing*, 2017, 141: 109-124.
18. Liu L, Miao S. An image encryption algorithm based on Baker map with varying parameter. *Multimed Tools Appl*, 2017, 76:16511.
19. Chai X, Gan Z, Yuan K, et al. An image encryption scheme based on three-dimensional Brownian motion and chaotic system. *Chin. Phys. B*, 2017, 26: 020504.
20. Wu J, Liao X, Yang B. Color Image Encryption Based on Chaotic Systems and Elliptic Curve ElGamal Scheme. *Signal Processing*, 2017, 141: 109-124.
21. Wang S, Wang C, Xu C. An image encryption algorithm based on a hidden attractor chaos system and the Knuth–Durstenfeld algorithm. *Optics and Lasers in Engineering*, 2019, 128: 105995.

22. Wang X, Zhang J, Zhang F, et al. New chaotical image encryption algorithm based on Fisher Yatess scrambling and DNA coding. *Chin. Phys. B*, 2019, 28: 040504.
23. Mohamed Amine Midoun, Wang X, Mohamed Zakariya Talhaoui. A sensitive dynamic mutual encryption system based on a new 1D chaotic map. *Optics and Lasers in Engineering*, 2021, 139: 106485.
24. Chen H, Liu Z, Camel Tanougast, et al. A novel chaos based optical cryptosystem for multiple images using DNA-blend and gyrator transform. *Optics and Lasers in Engineering*, 2021, 138: 106448.
25. Li Y, Wang C, Chen H. A hyper-chaos-based image encryption algorithm using pixel-level permutation and bit-level permutation. *Optics and Lasers in Engineering*, 2017, 90: 238-246.
26. Li H, Wang Y, Zuo Z. Chaos-based image encryption algorithm with orbit perturbation and dynamic state variable selection mechanisms. *Optics and Lasers in Engineering*, 2019, 115: 197-207.
27. Yavuz E. A novel chaotic image encryption algorithm based on content-sensitive dynamic function switching scheme. *Optics & Laser Technology*, 2019, 114: 224-239.
28. Wang M, Wang X, Zhang Y, et al. A novel chaotic encryption scheme based on image segmentation and multiple diffusion models. *Optics & Laser Technology*, 2018, 108: 558-573.
29. Cheng G, Wang C, Chen H. A Novel Color Image Encryption Algorithm Based on Hyperchaotic System and Permutation-Diffusion Architecture. *International Journal of Bifurcation and Chaos*, 2019, 29(09): 1950115.
30. Zhou M, Wang C. A novel chaos based optical cryptosystem for multiple images using DNA-blend and gyrator transform. *Optics and Lasers in Engineering*, 2021, 138: 106448.

Authors Introduction

Mr. Wei Li

He received the B.S. degree from Tianjin University of Science and Technology, Tianjin, China. And now he is studying for a master's degree in electronic information at Tianjin University of Science and Technology.

Ms. Jingwen Liu

She is studying for a master's degree in electronic information at Tianjin University of Science and Technology.

Ms. Hongyan Jia

She received Ph. D. degree in control theory and control engineering from Nankai University in 2010. She is currently an associate professor of the department of automation in Tianjin University of Science and technology.

A Research on Image Defogging Algorithm Based on Image Enhancement

FangYan Li*, Haokang Wen, Chang Sheng, Min Zhao

College of Electronic Information and Automation, Tianjin University of Science and Technology,
300222，China

E-mail: *3162129540@qq.com

www.tust.edu.cn

Abstract

In order to solve the problem of low contrast image and loss of image details in the foggy weather, the image defogging technique is used to remove the noise in the image and improve the image contrast, so as to recover a clear and fog-free image. In this paper, we mainly introduce three image defogging algorithms: global histogram equalization, local histogram equalisation and the Retinex algorithm. The advantages and shortcomings of each algorithm are summarised through the study of the principles of each algorithm and the comparative analysis of the experimental result.

Keywords: digital image processing, image enhancement, foggy degraded image, dehazing algorithm

1. Introduction

In recent years, with the development of computer vision and image processing technology, the research and application of outdoor vision systems is growing rapidly, thus making image processing widely used in more scientific and engineering fields. These outdoor vision systems need to take outdoor scenery images as input and accurately detect the characteristics of the input images through computer vision, image processing and other processing techniques. In order to ensure that the vision system works properly around the clock, it must be able to adapt to all types of weather conditions in order to improve the reliability of the system. Foggy weather conditions are one of the most serious effects on vision of all weather conditions. The visibility of outdoor images is greatly degraded due to the presence of fog, haze, smog etc [1]. Therefore, it is of great practical importance to study how to effectively deal with image degradation obtained under weather conditions such as dust and fog, in order to recover atmospheric degraded images and enhance the information of scenic details.

The enhancement processing of foggy degraded images is a frontier subject of interdisciplinary, involving research in many fields such as computer vision and digital image processing. In recent years, due to the influence of severe weather such as heavy fog on the video surveillance system, the research of foggy image processing has attracted more and more scholars' attention and has become a research hotspot in the field of image processing. How to choose an algorithm that has better dehazing effect and is easy to apply in practice is the primary task of the current image dehazing work.

2. Algorithm Introduction

Three algorithms are mainly applied in this article: global histogram equalization algorithm, local histogram equalization and Retinex algorithm. Next, the relevant content of the three algorithms will be introduced.

Image enhancement is one of the most important issues in low-level image processing. Mainly, enhancement methods can be classified into two classes: global and local methods [2].

2.1. Global histogram equalization algorithm

The histogram equalisation algorithm is one of the more widely used global enhancement methods, mainly for enhancing the contrast of images with reduced dynamic range of grey values and for improving the dark or light areas of the whole image. The global histogram equalisation algorithm is simple and efficient to implement, and works well for single depth-of-field images, but for images with variable scene depth, it is difficult to reflect the changes in local depth of field in the image, and therefore, the contrast enhancement is less effective.

The global histogram equalization, the central idea is to change the grayscale histogram of the original image from a relatively concentrated grayscale interval to a uniform distribution in the entire grayscale range. Histogram equalization is to non-linearly stretch the image, redistribute image pixel values, so that the number of pixels in a certain gray scale range is approximately the same, and finally change the histogram distribution of a given image to a uniformly distributed histogram distribution.

2.2. Local histogram equalization algorithm

Since global histogram equalisation is for the whole image, local image regions are ignored and the details of the local image become blurred due to the increase in overall brightness. Local histogram equalisation is proposed, i.e. an image is divided into multiple regions and then histogram equalisation is performed on multiple regions. This method can improve the contrast and detail of local regions, but it has the problem of local block effect and large computational effort.

3. Retinex Algorithm

Retinex theory is a theory of image enhancement based on scientific experiments and scientific analysis of the human visual system. The basic principle model of the algorithm was first proposed by Land in 1971 as a theory of colour and a method of image enhancement based on the constancy of colour [3]. Retinex theory is based on the idea that the colour of an object is determined by its ability to reflect long (R), medium (G) and short (B) wavelengths of light, rather than by the absolute value of the intensity of the reflected light; The colour of the object is not affected by the non-homogeneity of the light and is consistent, i.e. Retinex theory is based on colour sense consistency. The algorithm balances the three metrics of image dynamic range compression, image enhancement and image colour constancy, allowing for adaptive enhancement of foggy images. The Retinex enhancement process is therefore carried out by applying the Retinex algorithm to each of the R, G and B layers of the RGB image and then integrating them into the new image.

The main steps are as followed:

According to the image formation model, the image can be seen as consisting of an incident image and a reflected image, where the incident light shines on the reflecting object and is reflected by the reflecting object to form the reflected light into the human eye. The final image formed can be expressed by the following equation.

$$\mathrm{I}(x,y) = \mathrm{S}(x,y) \cdot \mathrm{R}(x,y) \tag{1}$$

(x, y) is the coordinates of the pixels in the image, “S” represents the incident light, “R” represents the reflection characteristics of the object, “I” is the reflected light, which is captured by the camera as an image.

Take the logarithm to separate the incident light component S and the reflected light component R. $\mathrm{I}'(x,y)$is the logarithm of $\mathrm{I}(x,y)$.

$$\begin{aligned}\mathrm{I}'(x,y) &= \log[\mathrm{S}(x,y) \cdot \mathrm{R}(x,y)] \\ &= \log(\mathrm{S}(x,y)) + \log(\mathrm{R}(x,y))\end{aligned} \tag{2}$$

Convolve the original image with a Gaussian template, which is equivalent to low-pass filtering the original image to obtain a low-pass filtered image D(x,y), where F(x,y) represents Gaussian filtering function:

$$\mathrm{D}(x,y) = \mathrm{I}'(x,y) \cdot \mathrm{F}(x,y) \tag{3}$$

Subtract the low-pass filtered image from the original image to obtain the high-frequency enhanced image G(x,y).

$$\mathrm{G}(x,y) = \mathrm{I}'(x,y) - \mathrm{D}(x,y) \tag{4}$$

In the previous steps, the incident light component S and the reflected light component R are separated, so the antilog of the resultant high-frequency enhanced image G(x,y) must be taken to obtain the enhanced image R(x,y).

$$\mathrm{R}(x,y) = \exp(\mathrm{G}(x,y)) \tag{5}$$

The contrast enhancement of R(x,y) is performed to obtain the final result image. System circuit module design In the circuit design, this design adopts a voltage stabilizing module, temperature and humidity data acquisition module and voice module. These modules greatly improve the function of the device.

Researchers such as Jobson proposed the single-scale Retinex (SSR) algorithm [4],[5]. The specific formula is as follows:

$$R_i(x,y) = \log I_i(x,y) - \log[I_i(x,y) * F(x,y)] \qquad (6)$$

In the formula, Ri(x,y) is the output of Retinex in the “i” color spectrum, Ii(x,y) is the image distribution, that is, the brightness value at the position (x,y). * represents the convolution operation, F(x,y) is a wraparound function that is defined by equation (7).

$$F(x,y) = K \cdot e^{\frac{-(x^2+y^2)}{\sigma^2}} \qquad (7)$$

Among them, σ is the wrapping scale, K is the normalization constant. The wrapping function satisfies:

$$\iint F(x,y)dxdy = 1 \qquad (8)$$

The stronger the dynamic compression capability of SSR is, the better the details of the dark part of the image can be enhanced, but the color distortion of the output image is more serious.

MSR is developed on the basis of SSR algorithm [6]. And it has the advantage of maintaining both high image fidelity and compression of the dynamic range of the image. MSR also enables colour enhancement, colour constancy, local dynamic range compression, global dynamic range compression and can also be used for X-ray image enhancement. The specific formula of MSR is as followed:

$$R_i(x,y) = \sum_{k=1}^{K} w_k \{\log I_i(x,y) - \log[I_i(x,y) * F(x,y)]\}, \quad i = 1, \ldots, N \qquad (9)$$

Among them, i represents the i-th color channel, and (x,y) represents the coordinates of the pixel in the image. N is the number of color channels in the image. N=1 represents a grayscale image, N=3 represents a color image, i ∈ (R,G,B). $I_i(x,y)$ represents the i-th color channel in the input image, $R_i(x,y)$ represents the output result of the MSR of the i-th channel. $F(x,y)$ is a Gaussian function, k represents the number of Gaussian surround functions or the number of surround scales, w_k represents the weight related to the Gaussian function, $\sum_{k=1}^{K} w_k = 1$, in general, MSR takes three scales of high, medium and low, K=3. As shown in Fig.1, it is the flow chart of MSR algorithm.

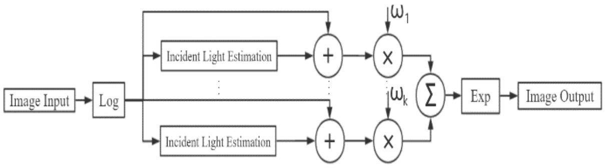


Fig.1 MSR algorithm flow chart

4. Result

The software used for image processing in this project is MATLAB (R2015a), based on Windows 10. As shown in Fig.2, it shows the results after three dehazing algorithms.

As shown in the figure, the first image in the first line is the original image, the second picture is the result of processing using the global histogram method, the first image in the second line is the local histogram method, and the last image is the MSR algorithm.

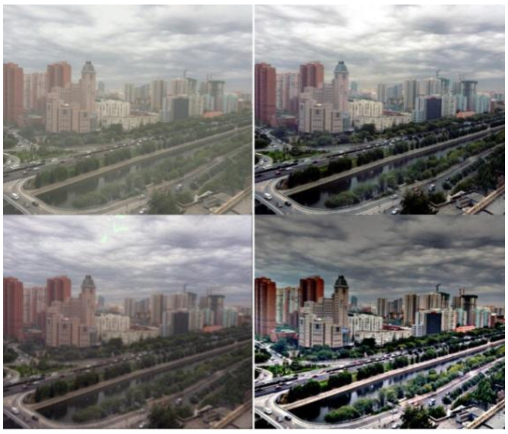


Fig.2 Three algorithm processing results

5. Conclusion

This article introduces the traditional image enhancement algorithm of global and local histogram equalization and the Retinex algorithm. Global enhancement is the dynamic adjustment of pixel values based on the statistical information of the whole image, with the aim of improving the overall contrast of the image. It should be noted that this method processes pixels independently of the location of the region and only reshapes the pixel distribution based on the image's grey scale information; in fact, the degree of scene degradation is closely related to the depth of field, so globalised enhancement is more limited when dealing with fogged images. Local enhancement highlights texture details in certain local areas of the image, eliminates block effects, suppresses over-enhancement and has some advantages over other algorithms in terms of image detail retention.

The enhanced image processed by the MSR algorithm can obtain a satisfactory enhancement effect by using the MSR algorithm. The contrast of the image is significantly improved, the detailed information of the image is fully enhanced, and the detailed features of the scene can be fully displayed and suppressed. noise. Although the MSR algorithm is a better image defogging algorithm, it still has some shortcomings.

This method regards the intensity of the image as the spatial coordinate vector of the RGB three colors, and performs arithmetic processing on it. In the synthesis of the three colors, errors are likely to occur and cause the distortion of the image.

Although it can be enhanced to a certain extent for dense fog weather, the overall effect cannot achieve the enhancement effect of general foggy weather. Since heavy fog has a greater impact on image clarity, once a foggy degraded image with multiple depths of field is formed, the algorithm does not have adaptability and does not take into account the local characteristics of the image, so local areas with different depths of field cannot be effectively enhanced.

References

1. Singh D, Kumar V. A comprehensive review of computational dehazing techniques. *Archives of Computational Methods in Engineering*, 2019, 26(5): 1395-1413.
2. Cheng H D, Shi X J. A simple and effective histogram equalization approach to image enhancement. *Digital signal processing*, 2004, 14(2): 158-170.
3. Land E H. The retinex theory of color vision. *Scientific american*, 1977, 237(6): 108-129.
4. Hines G, Rahman Z, Jobson D, et al. Single-scale retinex using digital signal processors. *Global signal processing conference*. 2005 (Paper 1324).
5. Wen H, Dai F, Wang D. A Survey of Image Dehazing Algorithm Based on Retinex Theory. 2020 5th International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS). IEEE, 2020: 38-41.
6. Rahman Z, Jobson D J, Woodell G A. Multi-scale retinex for color image enhancement. *Proceedings of 3rd IEEE International Conference on Image Processing*. IEEE, 1996, 3: 1003-1006.

Acknowledgments

This paper is partly supported by the Education Reform Project (2021-JG-03) from the Teaching Guidance Committee of Electronic Information in Higher Education of the Ministry of Education, China in 2021, and the Graduate Education Reform and Innovation Project (2021YJCB02) from Tianjin University of Science and Technology, China in 2021.

Authors Introduction

Ms. FangYan Li

She is the first-year graduate student of Tianjin University of Science and Technology. Her major is Control Science and Engineering.

Mr. Haokang Wen

He received master degree from Tianjin University of Science and Technology. His major is information processing and Internet of Things technology. His main research field is digital image processing. During his studies in school, he published several research papers.

Mr. Chang Sheng

He received master degree from Tianjin University of Science and Technology. His research interests are machine learning and deep learning, and he has published several papers during the school.

Ms. Min Zhao

She reveived master degree from Tianjin University of Science and Technology. Her research is about automatic and adaptive control.

Autonomous Microcontroller-Based Aerial Water Sampling Device

Weifang Wang[1], Mingxia Kang[2], Ruming Kang*[3]
([1]North Minzu University, Ningxia, China, [2]Tianjin University of Commerce, Tianjin, China,
[3]Chongqing University, Chongqing, China)

E-mail: **krm15053429@163.com*
www.cqu.edu.cn

Abstract

Water quality testing starts with sampling. Traditional methods may cause water quality properties to change due to the limitation of sampling methods. In addition, it is difficult and risky to take samples in parts of special terrain, such as ravines and swamps, where manpower is difficult to reach. To this end, we combine a single-chip machine STM32-based sampling device with a UAV that can span complex terrain to solve the above problems. The characteristics of water quality without human intervention and equipment automation are of great significance.

Keywords: WIFI, intelligent home, ad hoc network, remote control, monitoring

1. Introduction

At present, more than 420 billion m3 of sewage is discharged into rivers, lakes and seas worldwide each year, polluting 5.5 trillion m3 of freshwater, which is equivalent to more than 14% of the total global runoff; As people drink contaminated water, this is one of the main causes of disease and even transmission.

This design mainly solves the problem of long distance water quality sampling. Compared with other designs, it has the advantages of low cost, flexibility, simple design structure and dexterity, adaptability and high operational accuracy. Using the most common microcomputer STM32 series combined with cost-effective miniature drones, the automated water quality aerial sampling device is designed cleverly. The most advanced approach is to collect small water samples for laboratory analysis, as many of the characteristics of interest to them cannot be measured in situ easily or efficiently at cost.

Through these water samples in the laboratory, they can measure chemical properties, including phosphate, total phosphate, nitrate/nitrite, nitrogen and ammonia, as well as biological properties such as the presence of toxic microcystidins. Certain properties can be measured on site, but they require a large number of devices. These include temperature, conductivity, pH, dissolved oxygen, light, turbidity, and Secchi transparency. So we can bring the collected water samples back to the laboratory for water quality testing. In this way, we know if the water quality is healthy.

2. Hardware Design

In our approach, we use miniature drones to easily ship them in cars or backpacks to research locations. These flight robots are computer-controlled, light, commercially available, and can carry a small payload of 750g for up to 20 minutes. Fortunately, the UAV's limited payload is not a critical disadvantage, as water samples do not have to be very large (20 ml = 20g) to be scientifically useful. The UAV's limited battery flight allows it to travel nearly a kilometre back and forth, which is close enough for many water sampling applications.

The method proposed in this work is sampling from the air. The sampling device is controlled remotely by the staff using the mobile terminal to reach the designated sampling point and descend to the sampling height [1]. The sampling device is driven to move along the moving guide rail to directly above the detection hole by setting the motor and control line inside, and the air inside the rubber-tipped drip tube is discharged by squeezing the rubber-tipped drip tube through the squeezing mechanism. After the air is discharged from the tip drip tube, the telescopic mechanism retracts downwards so that the lower end of the tip drip tube is submerged under the water surface, then the squeezing mechanism is reset so that the air is pushed by air pressure into the inside of the tip drip tube, and at the same time the sample liquid is pushed into the inside of the tip drip tube. This design mainly solves the problem of long-distance water quality sampling. There are other designs compared with the advantages of low cost, flexibility, simple and dexterous design structure, adaptability, high operational accuracy UAV range is stronger4.

This paper designs flight control system with STM32 F103 series of chips as the core, including sampling device module, quad rotor UAV module, power supply module, sensor module, wireless communication module and PWM electronic speed control module and other functional modules. Fig.1 is the module contact diagram of this design.

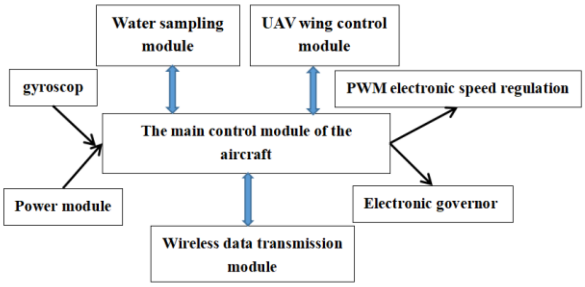

Water sampling module
UAV wing control module
gyroscop
PWM electronic speed regulation
The main control module of the aircraft
Power module
Electronic governor
Wireless data transmission module

Fig 1 Design of the module contact diagram

For the sampling device workflow: the upper end of the housing is fixed with a signal receiving antenna, the upper end of the housing is embedded with a chip and a control circuit, the upper end of the housing is fixed with a field-shaped moving rail, and the lower end of the moving rail is sliding with a connecting column; The lower end of the connecting column is fixed with a vertical downward-set telescopic mechanism. The lower end of the telescopic mechanism is fixed with a squeezing mechanism that shrinks laterally. The compression plates are fixed on both sides of the lower end of the extrusion mechanism. Fig.2 shows the side view of the structure of the design.

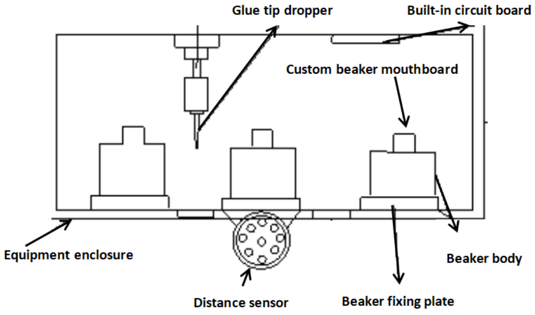

Glue tip dropper
Built-in circuit board
Custom beaker mouthboard
Equipment enclosure
Beaker body
Distance sensor
Beaker fixing plate

Fig 2 Side view of the structure

The internal fixing of the moving rail is equipped with motor and control line driving the movement of the connecting column. Together, the connecting column, motor and control line form the driving structure of the sampling device. The signal receiving antenna is electrically connected to the chip as well as the control circuit by wire. The sampling device design structure is shown in Fig.3.

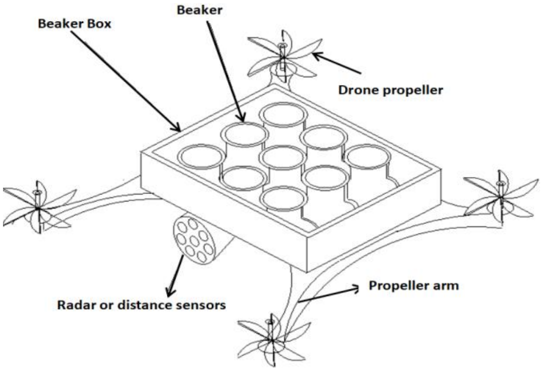

Beaker Box
Beaker
Drone propeller
Radar or distance sensors
Propeller arm

Fig 3 The sampling device design structure

3. Minimum height that can be maintained while sampling

The outdoor height control assessment was carried out on Tianta Lake in Tianjin. The water depth at this location is 2-3m. For these outdoor tests, we chose a calm day with a handheld anemometer measuring wind speeds of less than 0.38m/s. We recorded the ultrasonic, pressure sensor and Kalman filter height estimates. In this study, drones always fly at low altitudes. The figure shows that at a higher level, the software system implements a finite state automaton (FSA). Fig.4 is the structure diagram of the sampling process.

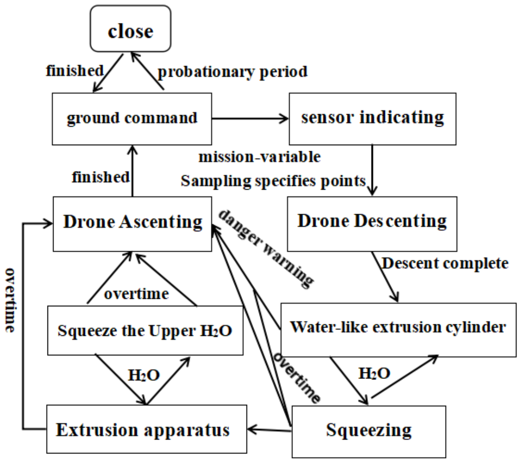

close
finished
probationary period
ground command
sensor indicating
mission-variable
Sampling specifies points
finished
Drone Ascenting
Drone Descenting
danger warning
Descent complete
overtime
overtime
Squeeze the Upper H_2O
Water-like extrusion cylinder
H_2O
H_2O
overtime
Extrusion apparatus
Squeezing

Fig 4 Diagram of the sampling process

Wind is a key environmental factor in any aerial field deployment. In the steady state of the flight system, the sampling system is specified to operate in winds of up to 10 m/s and, in our experience, it can remain roughly in position in open areas in winds of up to 15 m/s. However, when close to water, even wind speeds of less than 10 m/s can disrupt the system's ability to successfully sample water in several ways (1) The wind changes the air pressure around the body, resulting in large and rapid changes in the height recorded by the pressure altimeter. (2) The wind deflects the rubber-tipped dropper at an angle (5°), reducing the effective aspiration length and requiring the UAV to descend in altitude again, resulting in the flight system remaining in a critically stable state. (3) The wind disrupts the flow of water at the needle, resulting in less water entering the capture bottle. We know from experience that wind can interfere with altitude control and can bring the vehicle surprisingly close to swimming. However, The mission of this work required us to fly very close in order to achieve a high sampling success rate. During the sampling process this article set a 'target height' which the UAV attempted to maintain as it headed the dropper into the vial [2]. Overall, the flight system was able to operate the system normally without significant impact at water surface wind speeds ≤10m/s.

4. Reliability of Water Sampling Device

We tested the reliability and effectiveness of water sampling systems both indoors and outdoors. Indoors, we perform autonomous tasks, launching drones to 1m, flying over the tank, descending to the fully submerged sampling height of the glue drip, sampling, and then ascending to 1.5 m.

Overall, 139 of the 150 samples collected continuously from Chamber Lake (3 out of 50 tests) met the requirements (about 95% were successful). To better understand the relationship between success rates and the use of our ultrasound and pressure height controllers, half of the samples were collected using the height reported by the Vicon motion capture system. After adjusting for the controller, the success rate is 97.5%, which is ideal data. Then we tested the water samples collected (as shown in Figure 8), showing that the water samples collected using our water quality sampling device are closer to reality, and the error is about 66.23% smaller than the water samples collected manually. Fig.5 is a physical diagram of the sampling process.

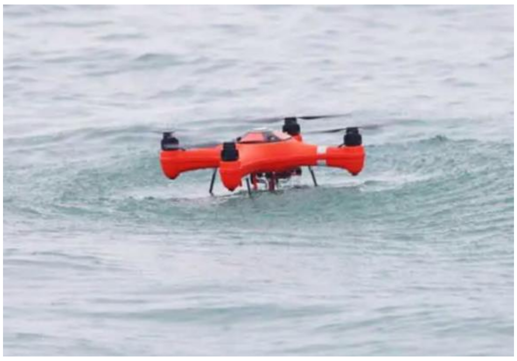


Fig 5 Physical diagram of the sampling process

The research is based on the definition of the classical quadrotor UAV system dynamics model and defines the angle ψ of rotation along the z-axis of the carrier coordinate system as the yaw angle, which is represented by the rotation matrix R2 (ψ). The angle of rotation Φ along the x-axis of the carrier coordinate system is defined as the roll angle and is represented by the rotation matrix Rx (Φ). The angle of rotation 0 along the y-axis of the carrier coordinate system is defined as the pitch angle and is represented by the rotation matrix Ry(0). The rotation matrix R(Φ) is also introduced to define the rotation relationship for the conversion of the quadrotor UAV carrier coordinate system B to the inertial coordinate system E [1].

$$\boldsymbol{R} = \begin{bmatrix} C_\psi C_\theta & C_\psi S_\theta S_\phi - S_\psi C_\phi & C_\psi S_\theta C_\phi + S_\psi S_\phi \\ S_\psi C_\theta & S_\psi S_\theta S_\phi + C_\psi C_\phi & S_\psi S_\theta C_\phi - C_\psi S_\phi \\ -S_\theta & C_\theta S_\phi & C_\theta C_\phi \end{bmatrix}$$

Each test consists of six samples, then examine the water sample bottle. Any amount less than the top of the sample vial "neck" is recorded as less than full. We completed 50 trials in total. Each test will take 2-3 minutes to fly and about 8 minutes to set up the flight system, take out water samples and regularly replace the UAV batteries.

5. Conclusion

In this work, we show a new mechanism for automatic water intake from drones, which requires less effort than existing technologies and is almost an order of magnitude faster. The system can safely fly at close range, collecting 9 30ml samples per flight. Finally, we conducted 150 outdoor tests on 6 samples in the range of 5-10m/s wind speed and found that the device can sample effective water samples with wind speeds less than 10m/s. The equipment solves the sampling of hard-to-reach terrain, such as ravines, swamps, etc., which not only improves the sampling efficiency, but also improves safety.

Future efforts include further operation and development of the system outdoors, in particular how the platform can be used in conjunction with adaptive sampling, sampling at greater depths and with other sensing and sampling mechanisms deployed in the water column. We plan to describe our risk management framework as part of a long-term analysis of failure modes and system reliability. We are seeking ways to measure the amount of water in a vial. We intend to explore how the system might operate on a wide variety of water bodies, including those with continuous flow and waves.

References

1. Kim Soram, Lee Seungyun, Kang Hyunsuk, Kim Sion, Ahn Minkyu. P300 Brain–Computer Interface-Based Drone Control in Virtual and Augmented Reality. Sensors, 2021, 21(17): 5765.
2. Liu Aike, Zhu Zhishan. Manipulator Massage System Based on STM32 Control. Journal of Physics: Conference Series, 2021, 1828(1): 012175.

Authors Introduction

Ms. Weifang Wang

She is currently studying at the School of Bioengineering, Northern Minzu University, Ningxia, China and she is currently an undergraduate student majoring in bioengineering.

Ms. Mingxia Kang

She is currently studying at the School of Pharmaceutical Engineering, Tianjin University of Commerce, where she is now an undergraduate student in pharmaceutical engineering.

Mr. Ruming Kang

He graduated from the School of Electronic Information and Automation of Tianjin University of Science and Technology in 2018 and is currently pursuing a master's degree at Chongqing University.

Intelligent Electronic Guide Dog

Mengyu Liu*, Yande Xiang, Zhi Qiao, Tao Zhu

College of Electronic Information and Automation, Tianjin University of Science and Technology,
300222, China

*E-mail: *2223282162@qq.com*

www.tust.edu.cn

Abstract

There are a large number of blind people in the world, and the scale is still increasing, so the scarcity of blind guide products has always been a difficult problem to be solved urgently. Although some intelligent blind guide products have flooded into the market in recent years, such as guide boxes, guide battles, etc., but most of their functions only stay in the obstacle avoidance stage, lack of path planning, intelligent recognition functions and are not easy to carry. Based on the above shortcomings, an intelligent mobile blind guide robot is designed in this paper. The blind guide robot is equipped with a variety of sensors and combined with the powerful computing power of the main control board to form a complete obstacle avoidance system to achieve autonomous navigation and obstacle avoidance in a complex environment. It controls GPS positioning through Jetson-Nano, plans the route, detects obstacles with the help of cameras and ultrasonic sensors, and gives users feedback through the alarm sound of the buzzer.

Keywords: *stm32, Jetson Nano, gps navigation, ultrasonic sensor, camera recognition*

1. Project Research Background

According to the data of the World Health Organization, there are currently about 285 million visually impaired people in the world, of whom 39 million are blind, and a survey in China has found that there are 21 million visually impaired people, of whom 23.8% are blind [1].

The continuous expansion of visual impairment groups has brought about an increase in the demand for blind guide facilities in the market. Not only that, with the continuous improvement of urban construction, people's living environment is also changing, the traditional blind guide equipment has been unable to meet this change. Therefore, the blind guide robot should be born and gradually be highly concerned by the industry and enterprises.

Guide dogs in China are free. As long as blind people can apply for guide dogs, guide dogs are donated free of charge to the visually impaired. However, it costs at least 130000 yuan to train a guide dog, and after training it, it is donated free of charge to the blind, and there is no capital recovery. China's first guide dog training public welfare organization, the Dalian training Base of China Guide Dog, is in trouble. Every year, only 20 guide dogs are provided to thousands of visually impaired people across the country. And the service life of the guide dog is 8 to 10 years, if the guide dog is too old, there may be errors in the work process. Therefore, for safety reasons, seeing eye dogs will retire when they work until they are 10 to 12 years old. Therefore, it is very important to design a practical guide robot to help the visually impaired.

At present, research work on guide assistance has been conducted in countries around the world, focusing on the research of walking AIDS, wearable marching AIDS, and mobile marching AIDS. As early as 2010, Japan developed an electronic guide stick, which uses ultrasonic sensors to make the visually impaired people feel the obstacles, and to effectively remind users through the vibration handles. The belt-like action accessory developed in the United States uses ultrasound to create a panoramic map of the region, allowing the visually impaired to "regain their eyes". In addition, the foreign research and development of smart trolleys and smart wheelchairs also provide the foundation and help for the research and development of mobile guide AIDS. China's research in this industry is relatively late, the current research focus is still on the guide rod, the research work on the blind guide robot is still in preparation, in the final analysis, the technology is not perfect. Whether it is sensor technology, tracking strategy and independent path planning capabilities, China needs to be further improved. On the whole, the blind guide robot belongs to the advanced products of the auxiliary blind guide tools. In order to achieve the progress of this industry, we still need to improve the core technology. Of course, the blind guide robot is a major trend in the future, and it must be the focus of our development. once the research and development of the blind guide robot in our country is on the right track, it will bring subversive changes to the existing blind guide facilities and the current situation.

2. Scheme Design

In this paper, we designed an intelligent mobile guide robot. The robot is equipped with a variety of sensors, combined with the powerful computing power of the main control board, to form a complete obstacle avoidance system, to realize autonomous navigation and obstacle avoidance in complex environments. It controls GPS positioning through the Jetson Nano, plans routes, detects obstacles with the help of cameras and ultrasonic sensors, and gives users feedback through the alarm sound of a buzzer.

2.1. Functional description

(1) Automatic obstacle avoidance of ultrasonic sensor.
(2) Bluetooth Module Motion Control.
(3) GPS positioning.
(4) Cameras identify vehicles, pedestrians, shoulders, etc, and the buzzer alarm when encountering pedestrians.

2.2. System composition

The system framework of the intelligent mobile guide robot system can be divided into the following parts: main control module, motion control module, obstacle avoidance module, positioning module, visual recognition module and power supply module.

2.3. Work principle introduction

The microcontroller stm32f103rct6 drive motor drive plate controls the speed of the rear two-wheel motor through PWM wave voltage regulation, and the stm32 microcontroller directly drives the steering machine through PWM pulse width modulation to control the steering of the first two wheels. SCM communicates with GPS module through serial port to obtain the position information. The ultrasonic module is directly controlled by the microcontroller to avoid the obstacles in front of it. The other main control board Jetson Nano controls the camera to obtain the wide-angle line-of-sight information in front of the camera, compare and analyze with other information collected from vehicles, road signs, pedestrians, and assist in obstacle avoidance.

The principle framework is shown in Fig. 1.

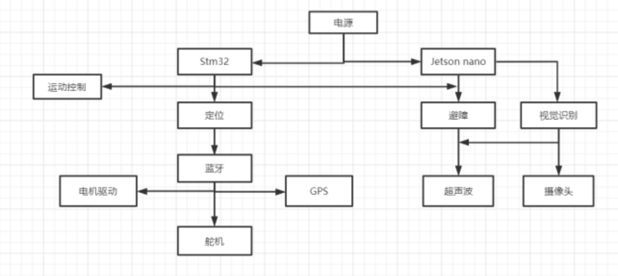


Fig 1 STM32F103ZET6 chip

2.4. Main control part design

With stm32 and Jetson Nano as the main control, stm32 is responsible for the underlying motion control, ultrasonic sensor and GPS module, Jetson Nano is responsible for visual information processing and analysis, and the two main control modules communicate through serial ports.

(1) Stm32f103rct6 Introduction

Using ArmCortex-M3 kernel, Harvard structure, with independent instruction bus and data bus.48KBSRAM, 256KBFLASH, 2 basic timers, 4 universal timers, 2 advanced timers, 2 DMA controllers (a total of 12 channels), 5 serial ports, 312-bit ADC, 112-bit DAC, 1 SDIO interface, and 51 universal IO ports. It has the advantages of powerful function, fast response, low power consumption, wide working temperature range, and is widely used in microcontroller motion control.

(2) Jetson Nano Introduction

Jetson Nano A02 is an open source hardware with good compatibility with the current commonly used Raspberry Pi motherboard. Its GPU is 128-coreMaxwell, 4GB of large capacity memory, strong image processing ability, low power consumption, high cost performance, so it is used as the main part of the visual recognition control.

2.5. Sport control design

The servo drive system is adopted, the first two wheels use the TBSK20 steering machine to directly receive the stm32 signal by PWM wave voltage regulation to control the steering, the second two wheels use the encoder motor, with TB6612 as the drive chip, and the PWM pulse width modulation technology is used to control the motor speed by changing the duty cycle, and then control the speed of the car. TB6612FNG is a DC motor drive device produced by Toshiba Semiconductor. It has a high-current MOSFET-H bridge structure, dual-channel circuit output, and can drive two motors at the same time. For PWM signal, the frequency of up to 100 kHz has a great advantage.

The motor drive schematic diagram is shown in Fig.2.

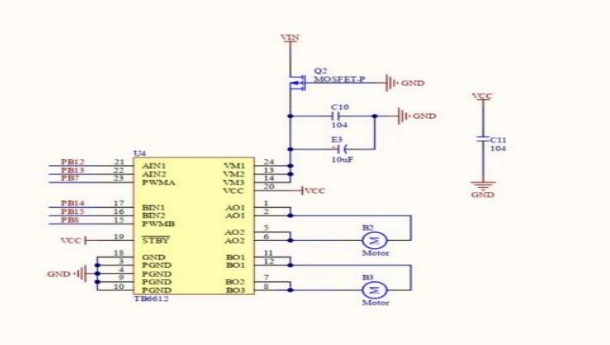

VIN
Q2
MOSFET-P
GND
C10
104
E3
10uF
VCC
C11
104
U4
AIN1
AIN2
PWMA
VM1
VM2
VM3
VCC
BIN1
BIN2
PWMB
AO1
AO2
STBY
GND
PGND
BO1
BO2
TB6612
PB12
PB13
PB7
PB14
PB15
PB8
B2
Motor
B3
Motor

Fig 2 Motor drive module

2.6. Design of obstacle avoidance module

The principle of ultrasonic sensor ranging is to obtain the distance information of the obstacle by calculating the time interval between the time of sending the source sound wave and the time of receiving the echo signal. The calculation formula is as follows [2].

$$\text{Test distance} = (\text{High level time} \times \text{speed of sound}) / 2$$

(The speed of sound is 340m/s.)

The HC-SR04 ultrasonic module has 4 pins, namely Vcc, Trig (control end), Echo (receiving end) and GND; VCC and GND are connected with 5V power supply, Trig (control end) controls the ultrasonic signal, and Echo (receiving end) receives the reflected ultrasonic signal. With a high level above 10US issued through the Trig pin, the high level output can be waited at the Echo receiver; the timer can be opened. When the output becomes low level, the value of the timer can be read and the distance can be calculated. Such a constant periodic measurement, you can reach the value of the moving measurement.

The schematic diagram of ultrasonic module is shown in Fig.3.

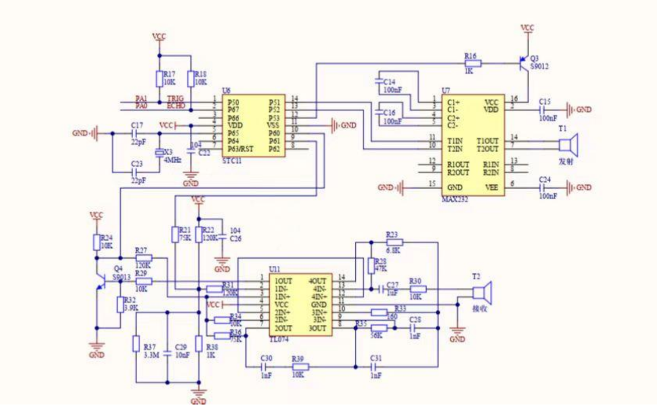

VCC
U8
STC11
U7
MAX232
Q3
S9012
T1
U11
TL074
Q4
S9013
T2
GND

Fig 3 Ultrasonic module

2.7. Gps module design

The GPS positioning module uses the ATGM336H of Zhongke Microelectronics, which has high sensitivity, supports the single system positioning of the BDS / GPS / GLONASS satellite navigation system, and the receiver module with any combination of multi-system joint positioning. ATGM336H can directly replace U-blox's MAX series of multiple GPS modules, the main interface signal is Pin-Pin compatible, consistent installation hole, low power consumption, built-in antenna detection circuit, with antenna short circuit protection function.

The schematic diagram of the GPS module is shown in Fig.4.

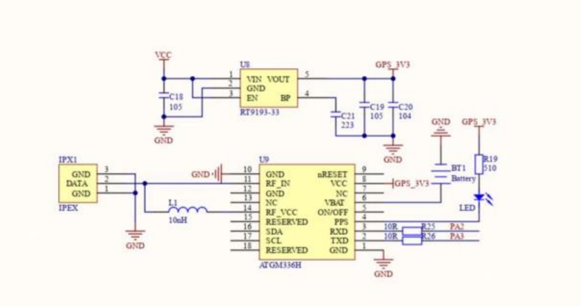

VCC
U8
VIN
VOUT
GND
EN
BP
RT9193-33
C18
105
C21
223
C19
105
C20
104
GPS_3V3
GND
IPX1
DATA
IPEX
L1
10nH
U9
RF_IN
NC
RF_VCC
RESERVED
SDA
SCL
nRESET
VCC
VBAT
ON/OFF
PPS
RXD
TXD
ATGM336H
BT1
Battery
R19
510
LED
R25
R26
PA2
PA3

Fig 4 GPS module

2.8. Visual recognition module design

With a Jetson Nano NX AI HD camera and 160 focal field wide Angle, the optical sensor chip is SONY IMX219, which is suitable for image data acquisition and information processing. The program used in this project is the yolov3 algorithm under the AlexeyAB modified darknet framework, which has a strong performance in the field of real-time target detection and can fit the computing power of Jetson Nano. Real-time running time in Jetson Nano can reach a detection rate of 10 to 20 frames, and can cope with 90% of the scenes in daily life [3].

The Jetson Nano interface diagram is shown in Fig.5.

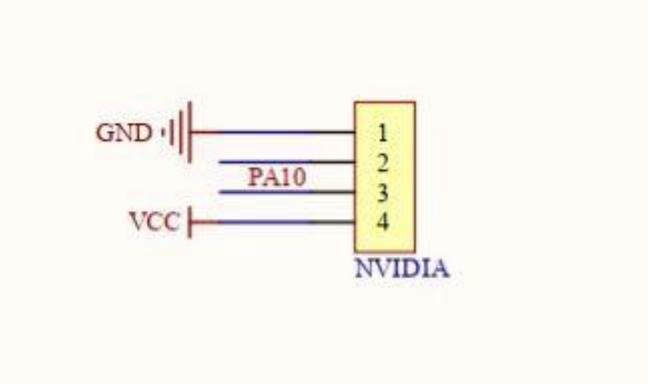

GND
PA10
VCC
1
2
3
4
NVIDIA

Fig 5 The Jetson Nano interface

2.9. Power supply module

Because there are many drive modules, the car adopts two parts to power supply stm32 and Jetson Nano respectively. First, 3 sections of 12V 1500mA power supply stm32, and another 3 sections of 12V 1500mA power supply, and AMS1117 5.0 is used to supply power to Jetson Nano with low voltage drop regulator.

The principle of antihypertensive is shown in Fig.6.

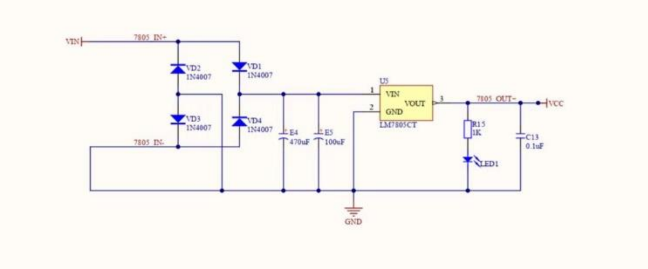
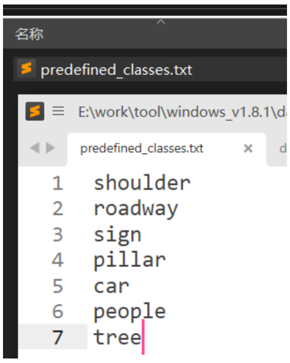

VIN
7805 IN+
VD2
1N4007
VD1
1N4007
VD3
1N4007
VD4
1N4007
7805 IN-
E4
470uF
E5
100uF
U5
VIN
VOUT
GND
LM7805CT
1
2
3
7805 OUT-
VCC
R15
1K
LED1
C13
0.1uF
GND

Fig 6.Anti-pressure module

3. Introduction of YOLO training algorithm

This project starts with devices with higher computing power than Jetson Nano, including the personal computer (windows platform) using the KIT T I data set published by the Karlsruhe Institute of Technology (KIT) and Toyota Technical University at Chicago (TTIC). It mainly includes data processing, training, use and deployment. Here is the first part. Download and use open source data sets or take a certain number of pictures in the field, and name the pictures in a certain order. The following example uses the KITTI-road dataset as an example. The dataset was manually annotated by using the labelIMG tool (The labelIMG tools: heartexlabs/label Img):

(1) Configure the preset value tag file in labelIMG: Open the predefined_classes.txt file in the data path and write the label name of the object to be detected to the file. The operation is shown in Fig.7.


名称
predefined_classes.txt
E:\work\tool\windows_v1.8.1\d
predefined_classes.txt

```
shoulder
roadway
sign
pillar
car
people
tree
```

Fig 7 Enter the object label name

(2) Using the labelIMG software.

The operations are shown in Fig.8 and Fig.9.

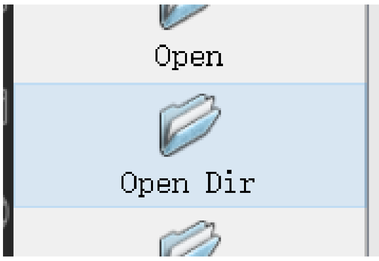

Open
Open Dir

Fig 8 Click on the Open Dir

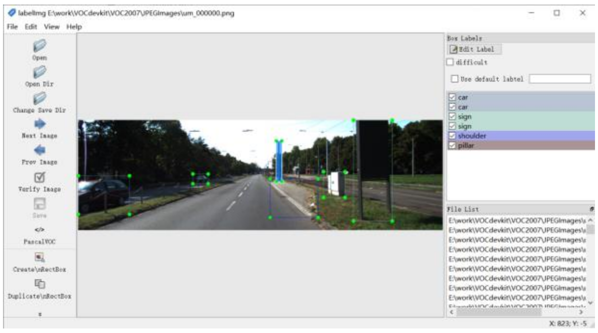

File
Edit
View
Help
Open
Open Dir
Change Save Dir
Next Image
Prev Image
Verify Image
Save
PascalVOC
Create\nRectBox
Duplicate\nRectBox
Box Labels
Edit Label
difficult
Use default label
car
car
sign
sign
shoulder
pillar
File List

Fig 9 Select the path to the downloaded image dataset or the actual image file.

(3) To mark The operations are shown in Fig.10, Fig.11 and Fig.12.

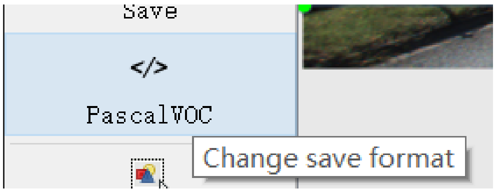

Save
PascalVOC
Change save format

Fig 10.Select the PascalVOC format dimension on the left side of the window.

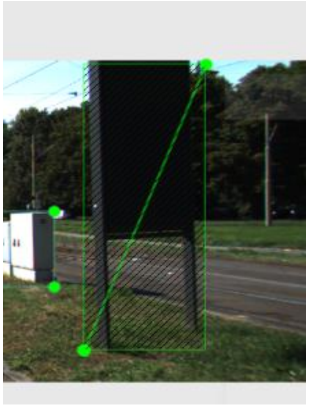


Fig 11 Press the shortcut key "w" to drag out the corresponding box on the image.

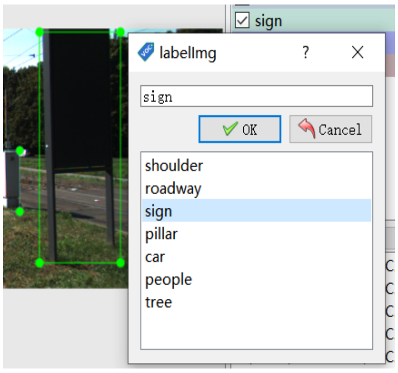
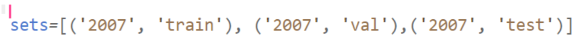


Fig 12 Then select the corresponding label in the pop-up window and click OK.

Right-key the box to modify the label or delete the box The annotation processing of the data set is completed after the annotation of all the picture files Then all the generated XML tag files are selected and cut to the scripts \ VOCdevkit \ VOC2007 \ Annotations folder, run the cut.py under scripts \ VOCdevkit \ VOC2007, and split the dataset, Then run the script under the project to further process the data, Open the voc_label.py script in the scripts directory, Edit the classes item in it, Keeping it consistent with the labels previously selected to make the dataset, The proportion of segmentation can be adjusted in the script, trainval_percent is the proportion of the datasets used for validation, The train_percent is the proportion of the data sets used for training: then run the voc_label.py in the scripts directory for format conversion, The modified classes item is consistent with the target category tag that you want to identify. To match the dataset format, the set item was modified as Fig.13.



```
sets=[('2007', 'train'), ('2007', 'val'),('2007', 'test')]
```

Fig 13 Modify the set code

This script generates both the 2007_train.txt and the 2007_val.txt files. Need to be used during the training sessions. These two files contain the path of the pictures contained in the training set randomly and the validation set.

4. Conclusion

The electronic guide dog introduced in this paper is a mobile robot, which is used to assist the blind people to travel safely. Compared with the current guide stick and guide box, it has obvious advantages and high cost performance. The main control module, motion control module, obstacle avoidance module, positioning module, visual recognition module and power supply module are introduced in detail, and the schematic diagram is given to complete the design of the whole guide dog.

References

1. Jun Liu, Yi Zhu, Jun Chen, et al. Research on the current situation of the digital viability of the blind people--Take the blind in Guiyang Blind Deaf-mute School as an example,*Education modernization*,2019,6(86):pp.351-353.
2. Siripun Thongchai and Kazuhiko Kawamura. Application of fuzzy control to a sonar-based obstacle avoidance mobile robot[C]. *Proceedings of the IEEE International Conference on Control Application*,USA,2000:pp.425-430.
3. KeSheng Xu, XiaoKui Cui, Smart football field pedestrian detection based on improved YOLOv3,*Computer system application*,2022,11:pp.1-8.

Authors Introduction

Ms. MengYu Liu

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. Zhi Qiao

He is currently pursuing his undergraduate degree at the School of Electronic Information and Automation, Tianjin University of Science and Technology.

Mr. Tao Zhu

He is currently pursuing his undergraduate degree at the School of Institute of Artificial Intelligence, Tianjin University of Science and Technology.

Research on Chaos Synchronization of Qi System and Lü System with Different Structures

Hanyuan Wang[*], **Yiting Gao**

Tianjin University of Science and Technology, 300222， China

E-mail: [*]*1463213510@qq.com*

www.tust.edu.cn

Abstract

This paper introduces a three-dimensional chaotic synchronization method is introduced, and the advantages and disadvantages of the synchronization controller designed by this method are analyzed. Firstly, the characteristics of two chaotic systems with different structures are studied. Secondly, the mathematical model is established, and the synchronization controller is designed by direct method, so that the two chaotic systems with different structures can be synchronized at different initial values. With the help of MATLAB, the error curve of the synchronous system is drawn when the synchronous controller acts on the response system.

Keywords: Chaos synchronization, Chaotic systems with different structures, Direct method, MATLAB

1. Introduction

Chaos theory is a key subject in the study of nonlinear theory in today's academic field. In 1963, the famous American scientist Lorenz first discovered the chaotic system [1]. This equation of chaotic system plays a very important role in the history of chaos, especially in the analysis of the emergence of chaotic solutions in nonlinear equations, which has great research significance. By analyzing the formation of chaotic system, scientists design a mathematical model suitable for the system, and then analyze whether the chaotic system can carry out chaotic motion under the mathematical model, and finally verify the basic characteristics of chaotic system.

The synchronization phenomenon was first discovered by the famous physicist Huygens. Once by chance, he saw two pendulums placed side by side, and they were exactly the same when they swung. This discovery opened up a branch in the field of mathematics and physical science the theory of the disaster oscillator, which revealed the synchronization phenomenon and its mechanism in nature. In 1990, Pecora and Carroll proposed a synchronization method to drive response, which made the chaotic trajectory of the system coincide under different initial conditions, and discovered the chaotic synchronization phenomenon in the circuit for the first time [2]. Scientists have found that there are many ways to synchronize chaotic systems, such as adaptive synchronization, delay synchronization, pulse synchronization and so on. In the study of two chaotic systems, firstly, the equilibrium point of the control system is analyzed. Then, according to the stability conclusion of Lyapunov's law, the correctness of the theoretical simulation results is verified by using the function method and numerical simulation [3],[4],[5]. The adaptive synchronization controller is designed and the adaptive synchronization method of chaotic system is discussed.

The discovery of chaos synchronization greatly promotes the study of chaos. Scientists have put forward the concept of chaotic system synchronization according to different directions, and the research of chaotic synchronization has made rapid progress.

2. Mathematical models of chaotic systems

The mathematical model of the Qi chaotic system is shown below:

$$\begin{cases} \dot{x} = a\left(y - x\right) + yz \\ \dot{y} = cx - xz - y \\ \dot{z} = xy - bz \end{cases} \tag{1}$$

Unknown quantities such as x, y and z are state variables of the system. The unknown quantity is the state variable of the system is $x, y, z \in \mathrm{R}$. $a = 35$, $b = 8/3$, $c = 80$ are typical parameters of the system.

The state variables of the mathematical model are transformed on the x, y and z axes respectively $\left(x, y, z\right) \to \left(x, -y, -z\right)$, $\left(x, y, z\right) \to \left(-x, y, -z\right)$, $\left(x, y, z\right) \to \left(-x, -y, z\right)$.

It is found that only when z-axis transformation is carried out, the mathematical model of the system does not change, so it can be explained that the mathematical model of the system is symmetric about the z-axis.

The partial derivative of the mathematical model of Qi chaotic system is obtained:

$$\nabla V = \frac{\partial \dot{x}}{\partial x} + \frac{\partial \dot{y}}{\partial y} + \frac{\partial \dot{z}}{\partial z} = -\left(a + b + 1\right) = -\frac{116}{3} < 0 \tag{2}$$

According to the partial derivative of the above equation, the derivative value is less than zero. According to the results, it can be concluded that the mathematical model of the system is always dissipative. The change of all systems is limited and bounded. As time increases, the orbit of the system will not spread out to infinity.

2.1 Chaotic dynamic properties

Set the initial value of the Qi chaotic system to $\left(x_0, y_0, z_0\right) = \left(1, 1, 1\right)$. The Lyapunov index of Qi chaotic system is obtained by MATLAB operation toolbox as $\lambda_1 = 4.0517 > 0$, $\lambda_2 = -0.0027 \approx 0$. Because of the Lyapunov index is $\lambda_1 = 4.0517 > 0$, it means that Qi chaotic system can have chaotic motion at this time. By putting the three Lyapunov exponents obtained into the formula for solving, the Lyapunov dimension d_L of the system is calculated as:

$$d_L = j + \frac{\sum_{i=1}^{j} \lambda_i}{\left|\lambda_{j+1}\right|} = 2 + \frac{4.0517 - 0.0027}{42.7151} = 2.0948 \tag{3}$$

j is the largest integer satisfying the condition that $\sum_{i=1}^{j} \lambda_i > 0$. According to the above formula, d_L =2.0948 , so the dimension of the chaotic system is fractional dimension.

The curves of the state variables of the chaotic system can be obtained by Matlab simulation as follows. Fig. 1 is the phase trajectory diagram of the chaotic system obtained by Simulink simulation, and Fig. 2 is the phase trajectory curve of Qi chaotic system in different coordinate systems.

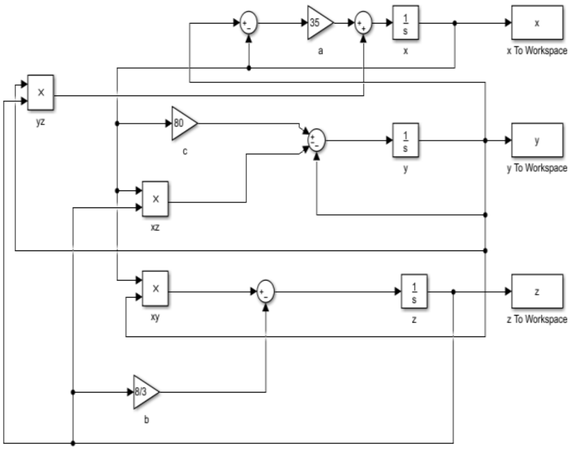
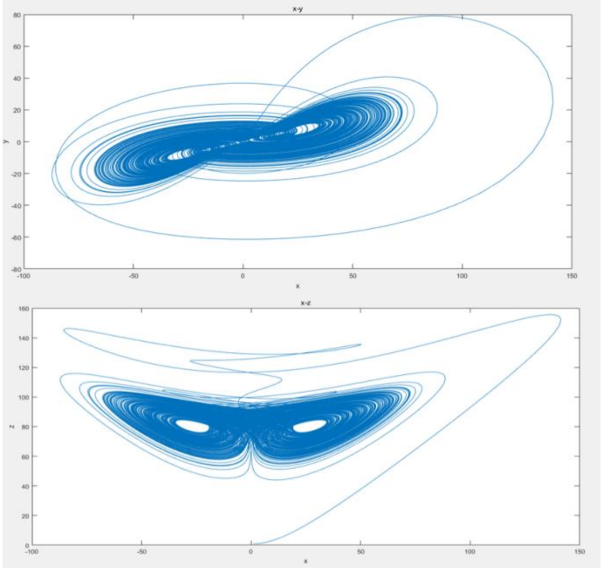


Fig. 1 Simulink simulation and construction of Qi chaotic system



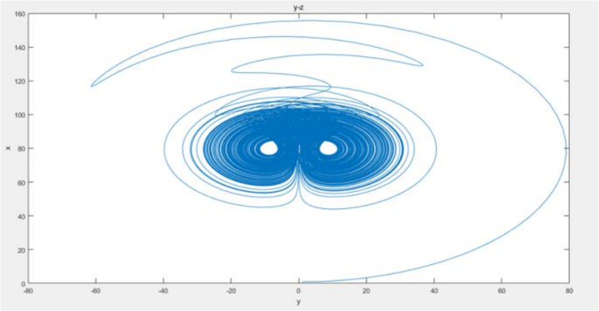


Fig. 2 Phase trajectory curve of Qi chaotic system system

Set the initial values of the system as $\left(x_{01}, y_0, z_0\right) = \left(1, 1, 1\right)$ and $\left(x_{02}, y_0, z_0\right) = \left(1.0001, 1, 1\right)$ respectively, and build a mathematical model through simulink to draw the solution curves of the corresponding three variables of the system, as shown in Fig. 3.

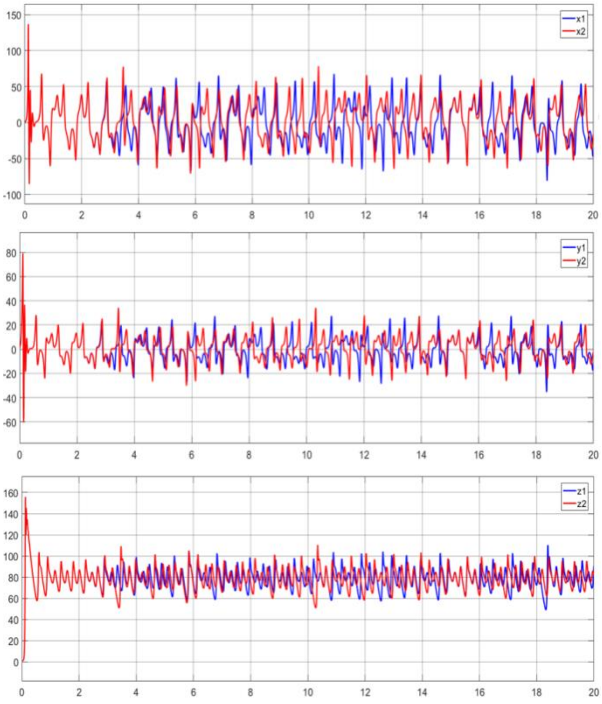


Fig. 3 Qi Solution curves of chaotic system in different states with different initial values

Keep other conditions unchanged, only change the initial value of state variable x of Qi chaotic system by 0.01%, in a short time, the two solution curves of chaotic system will change significantly. This shows that Qi chaotic system is susceptible to small changes in initial value and has a large influence. Therefore, Qi chaotic system is very sensitive to the change of initial value, which is an important characterization of the study of chaotic system.

2.2 Characteristic of equilibrium point

Set the right side of the mathematical model equation of the system as 0, and find the equilibrium state equation is:

$$\begin{cases} a\left(y - x\right) + yz = 0 \\ cx - xz - y = 0 \\ xy - bz = 0 \end{cases} \tag{4}$$

Solve the system of state, let:

$$\begin{cases} x_0 = \sqrt{\dfrac{b\left[ac + c^2 - 2a + c\sqrt{(a+c)^2 - 4a}\right]}{2a}} \\ y_0 = \dfrac{\sqrt{2ab\left[ac + c^2 - 2a + c\sqrt{(a+c)^2 - 4a}\right]}}{a + c + \sqrt{(a+c)^2 - 4a}} \\ z_0 = \dfrac{ac + c^2 - 2a + c\sqrt{(a+c)^2 - 4a}}{a + c + \sqrt{(a+c)^2 - 4a}} \end{cases} \tag{5}$$

The equilibrium point of the system under the mathematical model can be obtained by substituting parameters into the formula, respectively:

$$\begin{cases} S_1 = \left(0, 0, 0\right) \\ S_2 = \left(x_0, y_0, z_0\right) \\ S_3 = \left(-x_0, -y_0, z_0\right) \end{cases} \rightarrow \begin{cases} S_1 = \left(0, 0, 0\right) \\ S_2 = \left(26.3899, 8.0531, 79.6948\right) \\ S_3 = \left(-26.3899, -8.0531, 79.6948\right) \end{cases} \tag{6}$$

When the equilibrium point is the Qi chaotic system linearization, the coefficient of each variable is written into the matrix to obtain its Jacobian matrix, and then the equilibrium point into:

$$J_1 = J\Big|_{S_1} = \left.\begin{bmatrix} -a & a+z & y \\ c-z & -1 & -x \\ y & x & -b \end{bmatrix}\right|_{S_1} = \begin{bmatrix} -a & a & 0 \\ c & -1 & 0 \\ 0 & 0 & -b \end{bmatrix} \tag{7}$$

Transform the above matrix into a determinant, make the determinant equal to zero, and then expand the determinant to obtain the characteristic equation as follows:

$$f\left(s\right) = \left(s + b\right)\left[s^2 + \left(a+1\right)s - a\left(c - a\right)\right] = 0 \tag{8}$$

Through calculation, the characteristic root of the matrix can be obtained as:

$$s_1 = -b = -2.6667 \tag{9}$$

$$s_2 = \frac{-\left(a+1\right)+\sqrt{\left(a+1\right)^2+4a\left(c-1\right)}}{2} \approx 37.5788 \tag{10}$$

$$s_3 = \frac{-\left(a+1\right)-\sqrt{\left(a+1\right)^2+4a\left(c-1\right)}}{2} \approx -73.5788 \tag{11}$$

The equilibrium $s_1 = \left(0,0,0\right)$ is a saddle node.

The analysis method of the remaining two equilibrium points is the same as that of S_1, and S_2 and S_3 are symmetric with respect to the z-axis. So we can just analyze one of the two equilibria, and then we will analyze the equilibrium S_2.

By linearizing Qi chaotic system at the equilibrium point $s_2 = \left(26.3899, 8.0531, 79.6948\right)$ (same as S_1), its J matrix is

$$J_2 = \begin{bmatrix} -a & a+79.6948 & 8.0531 \\ c-79.6948 & -1 & -26.3899 \\ 8.0531 & 26.3899 & -b \end{bmatrix} \tag{12}$$

According to the above method, the matrix is transformed into the determinant, and the determinant is set to zero. Then, the determinant is expanded to get the characteristic equation. Through the operation, the characteristic root of J_2 are:

$$s_1 = -45.8958 \tag{13}$$

$$s_{2,3} = 3.6146 \pm 32.3465\mathrm{j} \tag{14}$$

$s_1<0$ and the values of the other two equilibrium points are a pair of conjugate complex roots, and the real parts of the two values are greater than zero. According to Rouse stability criterion, it can be concluded that the properties of the equilibrium points S_2 and S_3 are the same, both of which are unstable focal points of the system.

By observing this table, it can be seen that two of the three equilibrium points obtained by the Qi chaotic system show instability, so the orbits near the two equilibrium points will disperse out. When the time increases, the tracks at these two points will disperse out, indicating that the system has the butterfly effect.

This chaotic system has the property of dissipation, chaotic motion can occur, so the whole system is stable. Because the dissipative property has the effect of stabilizing the system, it makes the outer orbitals of the system attractor gather into the attractor. But the adjacent orbitals repel, and they have to be separated exponentially. Therefore, the system is stable overall but unstable locally, and the system will exhibit a very cumbersome structure.

Similarly, the numerical simulation shows that Lü chaotic system has the same chaotic behavior as Qi chaotic system.

3. Design of synchronous controller

3.1 Driving system

In this paper, Qi chaotic system is taken as the driving system, and the mathematical model of Qi chaotic system is given as:

$$\begin{cases} \dot{x}_1 = a_1\left(y_1 - x_1\right) + y_1 z_1 \\ \dot{y}_1 = c_1 x_1 - x_1 z_1 - y_1 \\ \dot{z}_1 = x_1 y_1 - b_1 z_1 \end{cases} \tag{15}$$

Among them, the typical parameters of Qi chaotic system are: $a = 35$, $b = 8/3$, $c = 80$.

3.2 Response system

In this paper, the Lü chaotic system is taken as the response system. The equation with the synchronization controller is as follows:

$$\begin{cases} \dot{x}_2 = a_2\left(y_2 - x_2\right) + u_{c1} \\ \dot{y}_2 = -x_2 z_2 + c_2 y_2 + u_{c2} \\ \dot{z}_2 = x_2 y_2 - b_2 z_2 + u_{c3} \end{cases} \tag{16}$$

Typical parameters of Lü chaotic system are: $a = 36$, $b = 3$, $c = 20$. The synchronization controller is $u_c = \left[u_{c1} \quad u_{c2} \quad u_{c3}\right]^T$.

By using MATLAB software to simulate the error system and state synchronization curve of the controller, judge whether these two chaotic systems with different structures can show complete synchronization.

3.3 Direct method

When the equilibrium state of the system is very stable and the output of the system has reached the equilibrium state. With the increase of time, the energy stored in the system will decrease until it reaches the minimum stable value of the equilibrium state. Lyapunov direct method is from the point of view of energy. The movement of the system will consume energy, but the capacity of the system will not be consumed to zero.

So Lyapunov by introducing a fictitious function, this function is called Lyapunov function, denoted as $v(x,t)$ or $v(x)$. Let $v(x)$ be any scalar function, where x is the unknown variable of the system, if $v(x)$ satisfies the following properties:

(1) $\dot{v}(x)=\frac{d_{v(x)}}{dt}$ is continuous and can reflect the trend of energy change;

(2) $v(x)$ is positive definite and can reflect the magnitude of energy;

(3) When $\|x\|\to\infty$, $v(x)\to\infty$ reflects the distribution of energy, function $v(x)$ is called Lyapunov function.

Given two systems: $A=(x_1,y_1,z_1)$, $Y=(x_2,y_2,z_2)$, u$_c$ indicates the control quantity.

$$\begin{cases} A=f\left(A\right) \\ B=g\left(B\right)+u_c \end{cases} \tag{17}$$

The way to synchronize two chaotic systems with different structures is to find a suitable u$_c$ that make $\lim\limits_{t\to\infty}\|y(t)-x(t)\|=0$.

For any initial values $x(0)$ and $y(0)$, the control synchronization problem of the system can be transformed into the problem of system error.

Set the error $e=\left[e_1\ e_2\ e_3\right]^T=\left[x_2-x_1\ y_2-y_1\ z_2-z_1\right]^T$, choose the Lyapunov function $v(x)=\frac{1}{2}\sum\limits_{i=1}^{n}e_i^2$, obviously $v(x)$ is positive definite, if $v(x)$ is negative definite, then the error equation of state is asymptotically stable at the origin, that is, $\dot{v}(x)=\dot{e}_1e_1+\dot{e}_2e_2+\dot{e}_3e_3=0$ at $t\to\infty$ so you have to find a suitable u$_c$ to make $v(x)$ negative definite.

Set the controller to:

$$u_C=\left[u_{c1}\quad u_{c2}\quad u_{c3}\right]^T \tag{18}$$

By comparing the mathematical model of the drive system with that of the response system, the error system can be written as follows:

$$\begin{aligned} \dot{e}_1&=a_2(e_2-e_1)+(a_2-a_1)(y_1-x_1)-y_1z_1+u_{c1} \\ \dot{e}_2&=-x_2z_2+c_2e_2+(c_2+1)y_1+x_1z_1-c_1x_1+u_{c2} \\ \dot{e}_3&=x_2y_2-b_2e_3-x_1y_1+(b_1-b_2)z_1+u_{c3} \end{aligned} \tag{19}$$

Choose the Lyapunov function $v(x)=\frac{1}{2}\sum\limits_{i=1}^{n}e_i^2$ and calculate $\dot{v}(x)=\dot{e}_1e_1+\dot{e}_2e_2+\dot{e}_3e_3$.

$$u_c=\begin{bmatrix}(a_2-1)e_1-a_2e_2+(a_2-a_1)(x_1-y_1)+y_1z_1 \\ x_2z_2-(1+c_2)e_2-(c_2+1)y_1-x_1z_1+c_1x_1 \\ x_1y_1+(b_2-1)e_3-x_2y_2+(b_2-b_1)z_1\end{bmatrix} \tag{20}$$

The error formula and the typical parameters of the two systems are obtained one generation after another.

$$\begin{cases} \dot{x}=35(y_1-x_1)+(x_1-x_2)+y_1z_1 \\ \dot{y}=-y_2-x_1z_1+80x_1 \\ \dot{z}=x_1y_1-z_2-\frac{5}{3}z_1 \end{cases} \tag{21}$$

Finally, according to the equation of the system and the mathematical model of the Qi chaotic system, the model of the two systems is built, and the error curve and state synchronization curve of the Qi system and the Lü system are drawn to verify whether this method can synchronize the two systems with different structures.

Here, the initial values of the two systems are: $\left(x_1,y_1,z_1\right)=\left(1,1,1\right)$, $\left(x_2,y_2,z_2\right)=(30,30,30)$ .Simulation software was used to draw the error curves and state curves of the two systems, as shown in the Fig. 4 and Fig. 5.

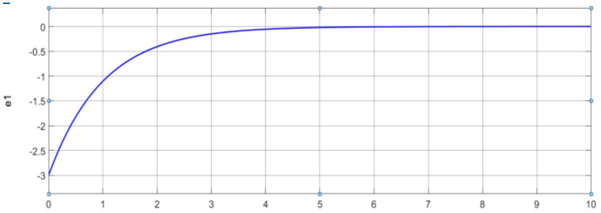
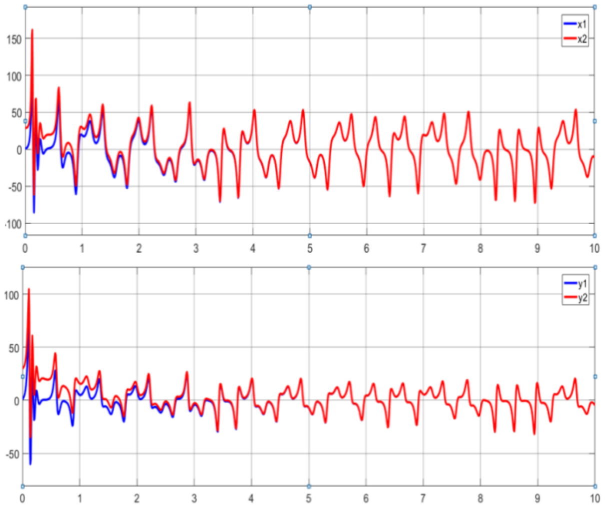


Fig. 4 Simulation error curve of synchronous controller designed by center translation method



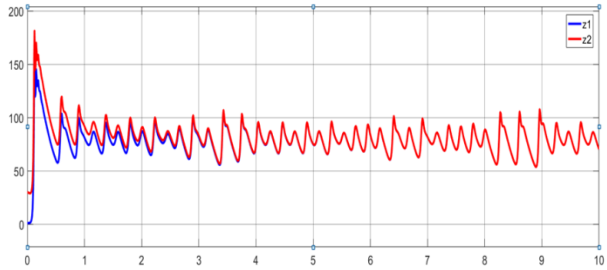


Fig. 5 Simulated state synchronization curve of synchronization controller designed with direct method

From the simulated image, it can be seen that the system can reach equilibrium in about 4.5 seconds.

4. Conclusion

By using direct method to design synchronous controller is simple in theory and easy to implement. In practical application, it effectively saves manpower, time and energy, and can effectively realize the synchronization problem of two different structures.

References

1. Lorenz E N, Deterministic nonperiodic flow, *Journal of atmospheric sciences*, 1963, 20(2): pp. 130-141.
2. Pecora L M, Carroll T L, Synchronization in chaotic systems, *Physical review letters*, 1990, 64(8): pp. 821.
3. Shahi M, Fallah Kazemi M, Adaptive sliding mode control approach, *Transactions of the Institute of Measurement and Control*, 2017, 39(1): pp. 86-95.
4. Zhang H, Huang W, Wang Z, et al, Adaptive synchronization between two different chaotic systems with unknown parameters, *Physics Letters A*, 2006, 350(5-6): pp. 363-366.
5. Huang J, Chaos synchronization between two novel different hyperchaotic systems with unknown parameters, *Nonlinear Analysis: Theory, Methods & Applications*, 2008, 69(11): pp. 4174-4181.

Authors Introduction

Ms. Hanyuan Wang

She received her B.S. degree from College of Tianjin University of Science and Technology, China in 2021. She is currently a Master course student in Tianjin University of Science and Technology. Her research area is about Inertial Technology.

Ms. Yiting Gao

She is currently a master course student in Tianjin University of Science and Technology, majoring in neural network, deep learning.

Generation and Analysis of a Multi-scroll Conservative Chaotic System

Jiaxin Li*, Yong Liu, Min Zhao

College of Electronic Information and Automation, Tianjin University of Science and Technology,
300222, China

*E-mail: *790861600@qq.com*

www.tust.edu.cn

Abstract

Multi scroll conservative chaotic system is a new kind of chaotic system, which has attracted extensive attention due to its complex dynamic characteristics. In this paper, we study a conservative chaotic system and introduce a sine function without multiple angles to make the conservative chaotic system generate multiple scrolls, so as to construct a multiple scroll conservative chaotic system. The system generates one-dimensional linear multi roll and two-dimensional grid like multi roll distributions by adjusting nonlinear functions.

Keywords: Multi-Scroll, Conservative Chaotic System, Balance Point Analysis, Lyapunov Index

1. Introduction

In recent years, the research on multi-scroll chaotic attractors [1] has become an important branch of chaos research. The dynamic behavior of multi-scroll chaotic system is more complex than that of single-scroll and double-scroll chaotic system. At present, due to the shortcomings of the analysis method, the characteristics of the multi-scroll chaotic system is only partially analyzed. But in practice, the multi-scroll chaotic system has its special value, for example, it plays a great role in the fields of secure communication and fuzzy recognition. At the same time, because the chaos attractors generated by the multi-scroll chaotic system are usually obvious, the study of the multi-scroll chaotic system can deepen the understanding of the generation mechanism of chaos attractors in the multi-scroll chaotic system, and is also conducive to the development of new chaotic systems.

2. Generation and Analysis of a Multi-scroll Conservative Chaotic System

The multi-scroll chaotic system with more complex dynamic characteristics is more suitable for chaos security and chaos compression and other applications. Based on the conservative chaotic system, the non-double Angle sine function is used to generate multiple scrolls in the conservative chaotic system, and the dynamic characteristics of the multi-scroll attractor are analyzed by means of Lyapunov exponential diagram, bifurcation diagram and attractor phase diagram.

2.1. The emergence of multi-scroll chaotic systems

The multi-scroll conservative chaotic system is constructed, and the non-double Angle sine function is introduced into the equilibrium equation (1).

$$\begin{cases} 0 = ay + byw \\ 0 = -ax \\ 0 = cw \\ 0 = -bxy - cz \end{cases} \tag{1}$$

Set $\nabla H(x) = (\sin(x), y, z, w)$ to get the new system equation:

$$\begin{cases} \dot{x} = ay + byw \\ \dot{y} = -a\sin(x) \\ \dot{z} = cw \\ \dot{w} = -b\sin(x)y - cz \end{cases} \tag{2}$$

The divergence of system equation (2) is $\nabla \cdot f = \frac{\partial \dot{x}}{\partial x} + \frac{\partial \dot{y}}{\partial y} + \frac{\partial \dot{z}}{\partial z} + \frac{\partial \dot{w}}{\partial w} = 0$, The divergence tells us that the system is still conservative.

Let $\dot{x} = 0, \dot{y} = 0, \dot{z} = 0, \dot{w} = 0$ to get equilibrium equation (3).

$$\begin{cases} 0 = ay + byw \\ 0 = -\mathrm{a}\sin(x) \\ 0 = cw \\ 0 = -b\sin(x)y - cz \end{cases} \tag{3}$$

Its equilibrium points should be classified into two categories. The first category is $(2k\pi,0,0,0)$ $(k\epsilon 1,2,3,\cdots)$, the second type is $((2k+1)\pi,0,0,0)$ $(k\epsilon 0,1,2,3,\cdots)$, such equilibrium points are common in multi roll chaotic systems. Equation (4) is its Jacobian matrix [2].

$$J(x)=\begin{bmatrix} 0 & a+bw & 0 & by \\ -a\,cos(x) & 0 & 0 & 0 \\ 0 & 0 & 0 & c \\ -bcos(x)y & -bsin(x) & -c & 0 \end{bmatrix} \tag{4}$$

Order $J-\lambda E=0$, when the balance point is the first type of balance point, $\lambda_1=-ai,\lambda_2=ai,\lambda_3=-ci,\lambda_4=+ci,a>0,b>0,c>0$, such balance point is the center point; When the balance point is the second type of balance point, $\lambda_1=a,\lambda_2=-a,\lambda_3=-ci,\lambda_4=+ci,a>0,b>0,c>0$, such equilibrium points are unstable saddle points [3].

2.2. Fundamental dynamic analysis of a multi-scroll chaotic system

Take an initial value of change as$[\pi/2+2k\pi,1,1,1],k=(0,1,2)$. When k=0, the initial value is [π/2,1,1,1], and the parameters a=9, c=14, the Lyapunov exponent of system equation (2) is shown in Fig.1.

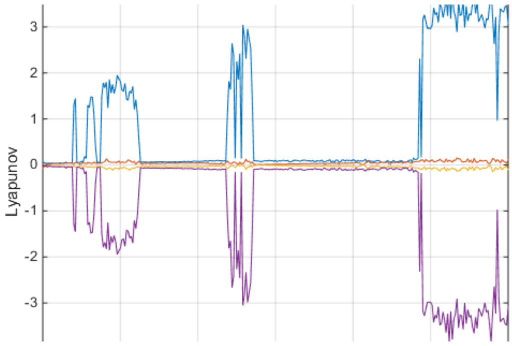


Fig.1. Laplace exponent diagram when a=9, c=14

Bifurcation Diagram Corresponding to Lyapunov Exponential Graph when Parameters a=9, c=14 is shown in Fig.2.

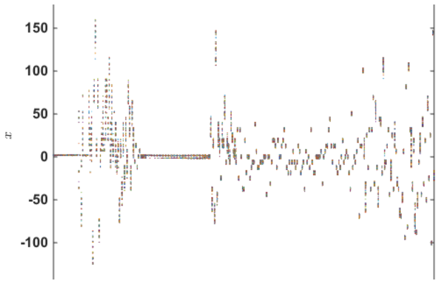


Fig.2. Bifurcation Diagram

When k=0, the initial value is [π/2,1,1,1], and the parameters a=4, c=4, the Lyapunov exponent of system equation (2) is shown in Fig.3.

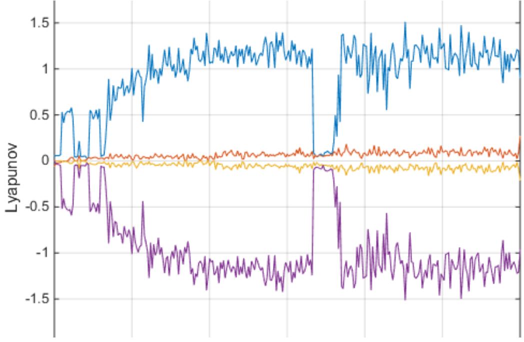


Fig.3. Laplace exponent diagram when a=4, c=4

Bifurcation Diagram Corresponding to Lyapunov Exponential Graph when Parameters a=4, c=4 is shown in Fig.4.

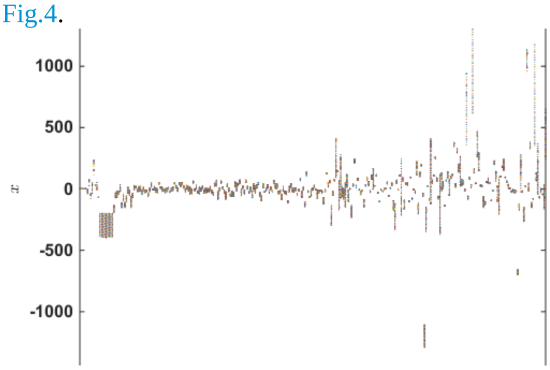


Fig.4. Bifurcation Diagram

By observing Lyapunov exponential graph and bifurcation diagram, it can be seen from system (2) that when parameters change, the distribution of Lyapunov exponential graph changes significantly. System (2) can generate chaos in significantly different parameter ranges.

2.3. A multi-scroll attractor and its stability analysis

The change of parameters obviously changes the interval in which the system can generate chaos. Here, the phase diagram of the system is observed by changing the initial value of the system, and the attractor characteristics of the system are analyzed by changing the phase diagram.

Take the system parameter $a=b=c=4$, then only the initial value of the system is changed.

Set initial value $y_0=z_0=w_0=1$, and the Lyapunov exponent with initial value x is shown in Fig.5.

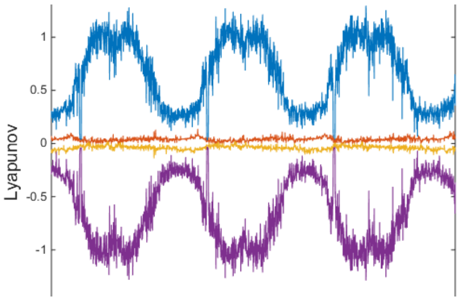


Fig.5. Laplace exponent diagram when a=4, c=4

Bifurcation Diagram Corresponding to Lyapunov Exponential Graph when Parameters a=4, c=4 is shown in Fig.6.

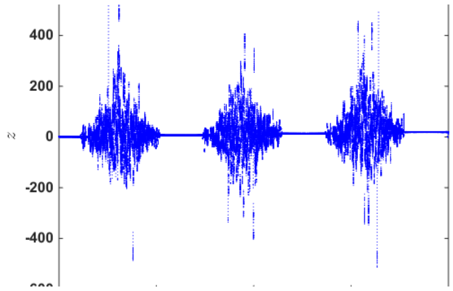


Fig.6. Bifurcation Diagram

By comparing the Lyapunov exponent diagram in Fig.5 and the bifurcation diagram in Fig.6, it can be seen that the range of chaos generated by the system is periodic. Within the range of $[0, 2\pi]$, the change of Lyapunov exponent within this range is similar to the image of a sine function. Then, taking 2π as the period, it can be seen that its image is the same as that within the range of $[0, 2\pi]$. That is, when the system changes with the initial value x, if $x = p, p \in (0,2\pi)$ is taken, then x can be taken as $p + 6k, k \in R^N$. The results obtained by the system can be the same.

When k=0, the initial value is [π/2,1,1,1], and the phase diagram generated by system equation (2) is shown in Fig.7.

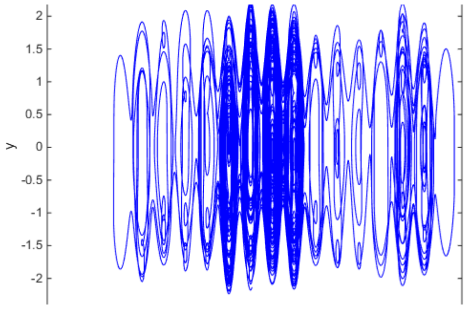


Fig.7. 16-scroll chaotic attractors

When k=1, the initial value is [5π/2,1,1,1], and the phase diagram generated by system equation (2) is shown in Fig.8.

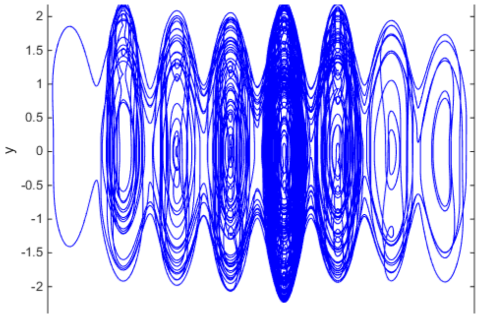


Fig.8. 8-scroll chaotic attractors

When k=2, the initial value is [9π/2,1,1,1], and the phase diagram generated by system equation (2) is shown in Fig.9.

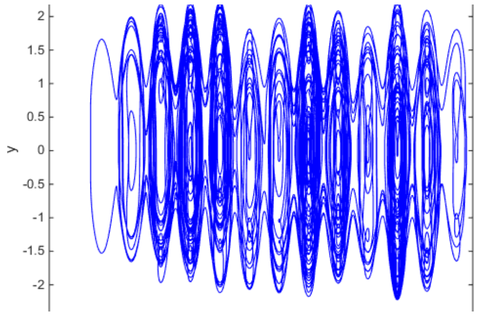


Fig.9. 13-scroll chaotic attractors

From Fig.7, Fig.8 and Fig.9, it can be seen that the number of scrolls in the x-y phase diagram varies significantly with the initial value, that is, the system has multiple stability, and the system is still a conservative system from the appearance. It can be seen from the multi roll phase diagram that the introduction of $\sin(x)$ leads to a one-dimensional multi roll chaotic attractor, which verifies the multi stability. Similarly, the introduction of angle free sine functions to y, z, w makes the system generate multi roll.

In Jacobian matrix (4), a two-dimensional nonlinear function $\nabla H(x) = (\sin(x), \sin(y), z, w)$ is introduced to generate multi scroll chaotic attractors

Take $a = b = c = 4$ to calculate the eigenvalues of different equilibrium points and analyze their characteristics. Characteristic value and type of balance point is shown in Table1.

Take $N = 2n\pi$ to limit the width of nonlinear function [4]. Equation (5) is the formula of nonlinear function.

Table 1. Characteristic value and type of balance point

balance point (x_0, y_0, z_0, w_0)	characteristic value $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$	type
$(2k\pi, 2k\pi, 0,0)$	$(-4i, 4i, -4i, +4i)$	Center point
$(2k\pi, (2k+1)\pi, 0,0)$	$(4, -4, -4i, +4i)$	Unstable saddle point
$((2k+1)\pi, 2k\pi, 0,0)$	$(4, -4, -4i, +4i)$	Unstable saddle point
$((2k+1)\pi, (2k+1)\pi, 0,0)$	$(-4i, 4i, -4i, +4i)$	Center point

$$f(u) = \begin{cases} u+N, u<-N \\ \sin(u), -N \le u \le N \\ u-N, u>N \end{cases} \tag{5}$$

Lyapunov Exponential Graphs Varying with Initial Values is shown in Fig.10.

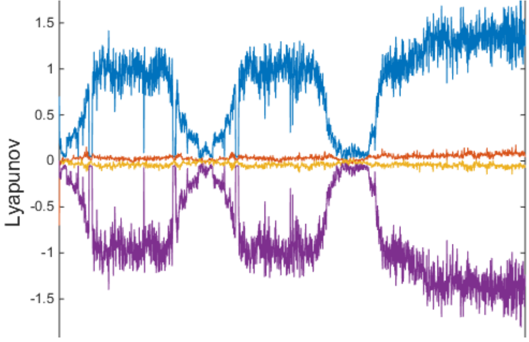


Fig.10. Laplace exponent diagram

In Fig.10, before $x = 4\pi$, the change rule of the Lyapunov exponent of the system that is greater than zero is still an approximate periodic change. After $x = 4\pi$, the change rule of the system has changed due to the use of two sine functions without multiple angles, breaking the previous change rule. It can be seen from the image that the Lyapunov exponent has a trend of gradual expansion after 4π.

The initial value is $[\pi/2, \pi/2, 1, 1]$, and the parameter is $a = b = c = 4$. We can get the two-dimensional multi roll attractor phase diagram, and the number of attractors is related to the value of n. If the number of scroll attractors generated by the system is M, then there is a relationship (6).

$$M = (2n)^2 + (2n+1)^2 \tag{6}$$

If the order of the generated two-dimensional multi scroll is R, the order relationship (7) can be obtained.

$$R = 4n+1 \tag{7}$$

Taking $n = 1$, the phase diagram of the scroll attractor in Fig.11 is obtained.

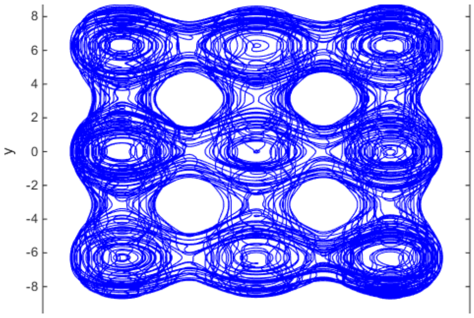


Fig.11. 5×5 multi-scroll chaotic attractor phase diagram

Fig.11 shows that the number of attractors is $M = (2)^2 + (2+1)^2 = 13$.

Taking $n = 2$, the phase diagram of the scroll attractor in Fig.12 is obtained.

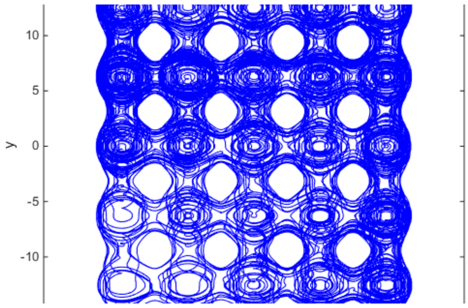


Fig.12. 9×9 multi-scroll chaotic attractor phase diagram 1

On the basis of $n = 2$, change the initial value, let $[\pi/2, 5\pi/2, 1, 1]$, and draw the phase diagram, as shown in Fig.13.

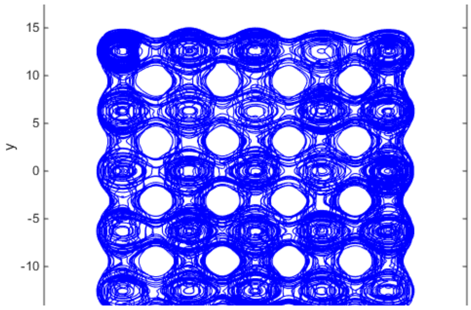


Fig.13. 9×9 multi-scroll chaotic attractor phase diagram 2

By changing the initial value and comparing Fig.12 and Fig.13, it can be seen that changing the initial value can change the internal motion track of the phase diagram and ensure that the number of scrolls does not change.

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.

Acknowledgments

This paper is partly supported by the Education Reform Project (2021-JG-03) from the Teaching Guidance Committee of Electronic Information in Higher Education of the Ministry of Education, China in 2021, and the Graduate Education Reform and Innovation Project (2021YJCB02) from Tianjin University of Science and Technology, China in 2021.

References

1. Ding, P. , Feng, X. , & Lin, F. . (2020). Generation of 3-d grid multi-scroll chaotic attractors based on sign function and sine function. Electronics, 9(12), 2145.
2. Guoyuan Qi & Jianbing Hu.(2020). Modelling of both energy and volume conservative chaotic systems and their mechanism analyses. *Communications in Nonlinear Science and Numerical Simulation* (C). doi:10.1016/j.cnsns.2020.105171.
3. Hildeberto E Cabral & Kenneth R Meyer.(1999).Stability of equilibria and fixed points of conservative systems. *Nonlinearity* (5). doi:10.1088/0951-7715/12/5/309.
4. Yazheng Wu,Chunhua Wang & Quanli Deng.(2021). A new 3D multi-scroll chaotic system generated with three types of hidden attractors. *The European Physical Journal Special Topics* (prepublish). doi:10.1140/EPJS/S11734-021-00119-8.

Authors Introduction

Mr. Jiaxin Li

In 2021, 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. Yong Liu

He received his Bachelor of Engineering degree from the school of Electronic Information and Automation of Tianjin University of Science and Technology in 2021. His main research direction is chaotic systems.

Ms. Min Zhao

She reveived her master degree from Tianjin University of Science and Technology, China in 2019. Her research is about automatic and adaptive control.

A Design of Fire Detection Device Based on YOLOv5

Zhiyang Li*, Yande Xiang, Haoyu Guo, Yu Chen, Wenxuan Pan, Fengzhi Dai

Tianjin University of Science and Technology, 300222， China

*E-mail: *a17730209318@163.com*

www.tust.edu.cn

Abstract

In recent years, fires have become more and more frequent, which has a great impact on people's production and life and even their lives. This paper designs a fire detection device based on YOLOv5, which is mainly composed of Raspberry Pi, Openmv and buzzer, which can be widely used in narrow corridors, parking lots, shopping malls, forests and other scenarios. The device has the characteristics of high recognition rate, fast recognition speed and strong sensitivity, and has excellent recognition effect in fire detection.

Keywords: fire detection, YOLOv5, Raspberry Pi , machine learning

1. Introduction

As one of the hidden dangers that threaten the safety of human life and production, fire has always been the focus of people's attention. Traditional fire monitoring devices detect fires according to temperature, not only poor sensitivity, but also long feedback time, often when firefighters receive alarm messages, the fire chamber has been uncontrollable. In this paper, a set of fire detection devices have been designed by Raspberry Pi, Openmv and Buzzer design, which can perform real-time sensitive detection of fires.

The rest of this article is organized below. The second part describes the working principle of the entire device. The third part introduces the advantages of the YOLOv5 model. The fourth part introduces the functions of each component. In the fifth part, the training set is explained and the detection results are presented. Part VI summarizes the main contents of this article.

2. Overview of Fire Detection Devices

This paper designs a fire monitoring device based on machine vision, which is suitable for most scenes. A convolution neural network is constructed by machine learning, and a large number of training neural networks are trained. The trained model is transplanted to Raspberry Pi, which is used as the main controller to control the operation of other modules.

In terms of the use of components, this paper uses Openmv as a camera to capture the pictures in the scene, and transmits each picture to Raspberry Pi. Raspberry Pi processes and recognizes images. If a flame scene is detected, it will send a signal to the buzzer, and the buzzer will give an alarm to remind the inspectors of the fire, thus achieving the effect of real-time fire monitoring [1]. The design of the The whole device is shown in Fig.1.

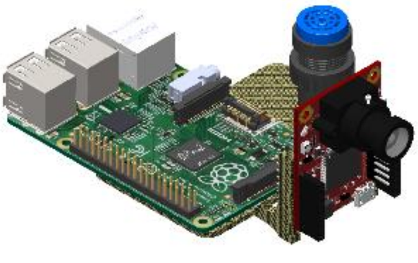


Fig.1 The appearance of the devices

3. Introduction of YOLOv5

Generally speaking, YOLO algorithm is a single-stage end-to-end detection algorithm based on anchor-free. After the images are input into the network for feature extraction and fusion, the prediction frame position and class probability of the detection target are obtained. Compared with previous generation YOLO algorithms, YOLOv5 has smaller model, flexible deployment and better detection accuracy and speed. It is suitable for real-time target detection. Its workflow diagram is shown in Fig.2 below.

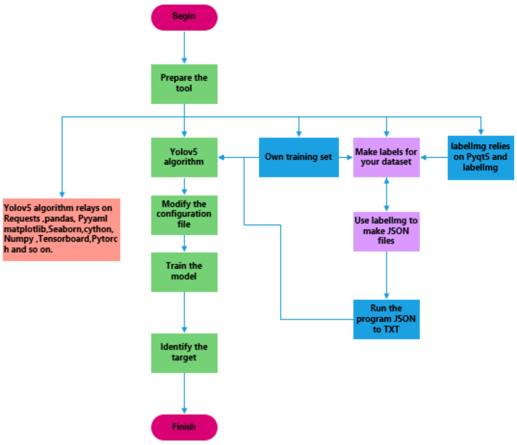

Begin
Prepare the tool
Yolov5 algorithm
Own training set
Make labels for your dataset
labelImg relies on Pyqt5 and labelImg
Yolov5 algorithm relays on Requests ,pandas, Pyyaml matplotlib,Seaborn,cython, Numpy ,Tensorboard,Pytorch and so on.
Modify the configuration file
Use labelImg to make JSON files
Train the model
Run the program JSON to TXT
Identify the target
Finish

Fig.2 The workflow diagram of the YOLOv5

YOLOv5 is divided into four models: YOLOv5s, YOLOv5m, YOLOv5l and YOLOv5x according to different depths and widths of feature maps. YOLOv5s is the smallest model, so this paper uses YOLOv5s model for fire detection.YOLOv5 is an extension of the YOLO series, and you can also see it as an improvement based on YOLOv3 and YOLOv4.

YOLOv5 does not have a corresponding paper description, but the author actively opens the source code on Github, and through the analysis of the source code, we can quickly understand the network architecture and working principle of YOLOv5 [2].

4. Component Design

4.1. Raspberry Pi

It is a microcomputer motherboard based on ARM, with SD/MicroSD card as the memory hard disk. There are 1/2/4 USB interfaces and a 10/100 Ethernet interface around the card motherboard, which can connect the keyboard, mouse and network cable. It also has TV output interface for video analog signals and HDMI high-definition video output interface. All the above components are integrated on a motherboard only slightly larger than credit cards. Its appearance is shown in Fig.3 below.

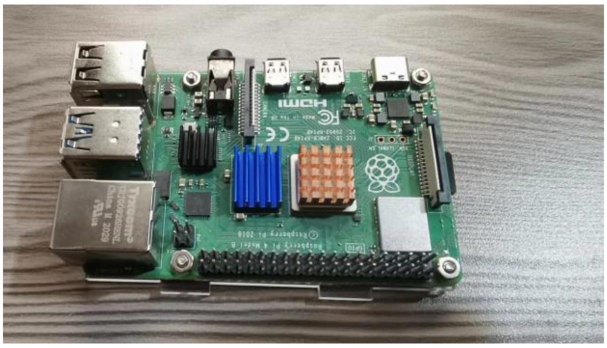


Fig.3 The appearance of the Raspberry Pi

With the basic functions of all PCs, you can perform many functions such as spreadsheets, word processing, playing games, playing high-definition videos, etc. just by turning on the TV and keyboard. Raspberry Pi B only provides computer board, without memory, power supply, keyboard, case or connection. In this project, as the master controller, it is responsible for scheduling various functional modules.

4.2. OpenMV

OpenMV is an open source, low cost and powerful machine vision module. With STM32F767CPU as the core, oV7725 camera chip is integrated. On the compact hardware module, the core machine vision algorithm is efficiently implemented with C language, and Python programming interface is provided. Fig.4 below shows the pin definition for Openmv.

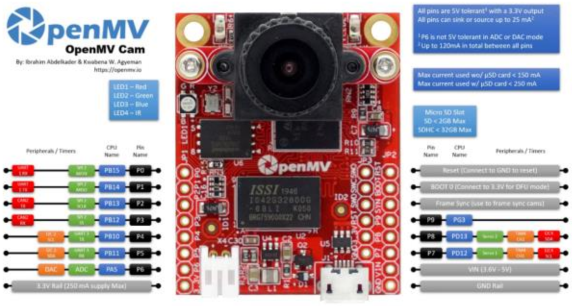


Fig.4 The pin definition for Openmv

In the construction of the fire monitoring device, the Openmv equipped with MT9V034 sensor module is used as the camera to collect pictures in real time and transmit them to Raspberry Pi for detection, which can perfectly provide convenience for accurate visual support and support color and grayscale output. In VGA/QVGA format, it can stably output 90fps frame rate. Output at 200fps frame rate in QQVGA format and 400fps frame rate in QQQVGA format.

4.3. Buzzer

Electromagnetic buzzer is composed of oscillator, electromagnetic coil, magnet, diaphragm and casing. After the power supply is turned on, the audio signal current generated by the oscillator passes through the electromagnetic coil, so that the electromagnetic coil generates a magnetic field, and the vibrating diaphragm periodically vibrates and sounds under the interaction of the electromagnetic coil and the magnet. When receiving the signal from Raspberry Pi, it can quickly send out an alarm to remind us to take corresponding measures. Its appearance is shown in Fig.5 below.

Fig.5 The appearance of the Buzzer

5. Test Results

In this paper, labelImg annotation tool is used to annotate the dataset images, and it is saved in the txt format of YOLO series. By arranging the data out of order, 1442 training sets, 617 test sets and 617 verification sets were randomly selected. Machine learning training through thousands of pictures of fires also has a good detection effect for tiny flames [3]. The detection effect is shown in Fig.6.

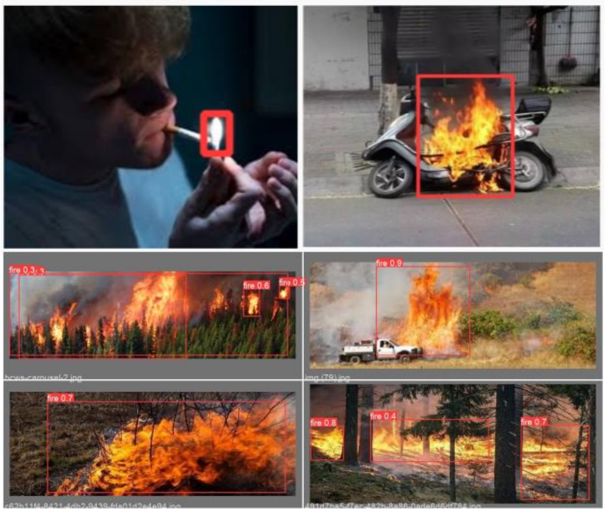


Fig.6 The detection effect of the device

6. Conclusion

Aiming at the problems of poor sensitivity and long response time of similar fire detection devices in the market, a new fire detection device based on machine vision is designed. Compared with the existing temperature and fire detection devices, the machine vision recognition technology has wide application and high practicability.

Acknowledgments

This paper is partly supported by the Education Reform Project (2021-JG-03) from the Teaching Guidance Committee of Electronic Information in Higher Education of the Ministry of Education, China in 2021, and the Graduate Education Reform and Innovation Project (2021YJCB02) from Tianjin University of Science and Technology, China in 2021.

References

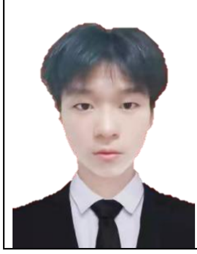
1. Wu Huafeng, Hu Yanglin, Wang Weijun, et al., Ship Fire Detection Based on an Improved YOLO Algorithm with a Lightweight Convolutional Neural Network Model, *Sensors*, 2022, 22(19): 7420.
2. Jiang T, Li C, Yang M, et al., An Improved YOLOv5s Algorithm for Object Detection with an Attention Mechanism. *Electronics*, 2022, 11(16): 2494.
3. Avazov Kuldoshbay, Mukhiddinov Mukhriddin, Makhmudov Fazliddin, et al., Fire Detection Method in Smart City Environments Using a Deep-Learning-Based Approach, *Electronics*, 2021,11(1): 73.

Authors Introduction

Mr. Zhiyang Li

He is a second-year undergraduate student at Tianjin University of Science and Technology. His research interests are computer vision and machine learning.

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. Haoyu Guo

He is currently studying in the School of Biology of Tianjin University of Science and Technology and has achieved certain results in related professional competitions.

Ms.Yu Chen

She is currently studying for an undergraduate degree at Tianjin University of Science and Technology.

Mr. Wenxuan Pan

He is currently pursuing his undergraduate degree at the School of Electronic Information and Automation, Tianjin University of Science and Technology.

Dr. Fengzhi Dai

4. 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.

Application of Convolutional Neural Network in Accurate Breast Cancer Identification

Zhiyang Li*, Haoyu Guo, Yande Xiang, Wentao Kuang, Lu Chen

Tianjin University of Science and Technology, 300222， China;

*E-mail: * a17730209318@163.com*

www.tust.edu.cn

Abstract

As one of the most common cancers in women, breast cancer has the highest incidence in the world. Nearly 600,000 people die from breast cancer each year, and early detection is essential for breast cancer treatment. In recent years, the rapid development of artificial intelligence has provided unprecedented ideas for the precise diagnosis and treatment of breast cancer. In this paper, the practical application of artificial intelligence convolutional neural network in breast cancer recognition is studied, which greatly improves the detection speed and saves a lot of time for doctors to further judge the condition.

Keywords: breast cancer, artificial intelligence, convolutional neural network, recognition

1. Introduction

In 2021, there will be 2.5 million breast cancer patients in China. The overall survival rate of breast cancer in the country in the past five years is less than 50%. The incidence of breast cancer in urban areas has reached 34.3 cases per 100,000 people, twice that of rural areas. Globally, women lose more disability-adjusted life years to breast cancer than any other type of cancer. Women in every country in the world develop breast cancer at any age after puberty, but the incidence is increasing later in life. Invasive ductal carcinoma is one of the most common types of breast cancer, and 80% of breast cancers are invasive ductal carcinoma. Currently, the diagnosis of invasive ductal carcinoma often requires a combination of tests, including a physical examination and some imaging tests. This process is not only time-consuming, but also extremely cumbersome. For some places where medical facilities are backward, the conditions for testing are not even available.

As a new technology developed by today's science and technology, artificial intelligence has made remarkable progress in the application of imaging and pathology[1]. This paper studies the practical application of artificial intelligence convolutional neural network in breast cancer recognition, which can achieve the effect of batch detection. Using AI can improve the efficiency of breast cancer diagnosis, reduce the workload of doctors, and get better feedback in clinical medicine.

The rest of this article is organized as follows. The second part introduces the source of the dataset, and the third part introduces the convolutional neural network and explains the architecture of the convolutional neural network used in this article. The fourth part is the effect of model detection, which shows that the recognition effect of the model is better through the confusion matrix and learning curve. The fifth part is a summary of the full text.

2. BreakHis Dataset

Spanhol et al. published the BreakHis (breast cancer histopathological database) dataset in 2016. The dataset contains 7909 pathological images of breast histopathology from 82 patients. The dataset currently contains four different types of benign breast tumors histologically: adenopathy (A), fibroadenoma (F), lobular tumors (PT), and tubular adenomas (TA); and four malignancies (breast cancer): carcinoma (DC), lobular carcinoma (LC), mucinous carcinoma (MC) and carcinoma (PC). The sample was derived from a breast tissue biopsy section, stained with hematoxylin and eosin (HE) and labeled by a pathologist.

This allows the BreakHis dataset to not only advance the research of benign and malignant binary classification algorithms, but also promote the research of pathological classification algorithms with more significant clinical significance. The dataset BreaKHis is divided into two broad categories: benign tumors and malignant tumors. Histological benign is a term that refers to lesions that do not meet any criteria for malignancy, e.g., obvious cellular atypia, mitosis, basement membrane disruption, metastasis, etc. As a rule, benign tumors are relatively "innocent", grow slowly and remain localized. Malignancy is synonymous with cancer: lesions can invade and destroy adjacent structures (local infiltrative) and spread to distant places (metastasis) leading to death. The samples present in the current dataset are collected by the SOB method, also known as partial mastectomy or excisional biopsy. Compared to any needle biopsy method, this type of procedure allows the removal of larger tissue samples and general anesthesia in the hospital. This dataset can provide strong support for the smooth progress of the study. A partial picture of the dataset is shown in Fig.1.

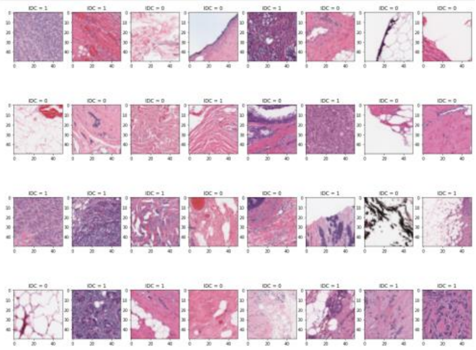


Fig.1 A partial picture of the dataset

3. Convolutional Neural Networks

Convolutional neural networks are variants of multilayer perceptrons (MLPs) developed by biologists Huebor and Wiesel in earlier research on the feline visual cortex, where cells have a complex structure that is very sensitive to subregions of visual input space called receptive fields. Convolutional neural network is a multi-layer supervised learning neural network, and the convolutional layer and pool sampling layer of the hidden layer are the core modules to realize the feature extraction function of convolutional neural network[2].

The network model uses the gradient descent method to minimize the loss function, and adjusts the weight parameters in the network layer by layer, and improves the accuracy of the network through frequent iterative training[3]. Compared with the original neural network, the convolutional network structure can better adapt to the structure of the image, and perform feature extraction and classification at the same time, so that feature extraction can help feature classification. And weight sharing can reduce the training parameters of the network, making the neural network structure simpler and more adaptable[4]. The figure below shows the architecture of a convolutional neural network in breast cancer. The neural network architecture used in the model is shown in Fig.2.

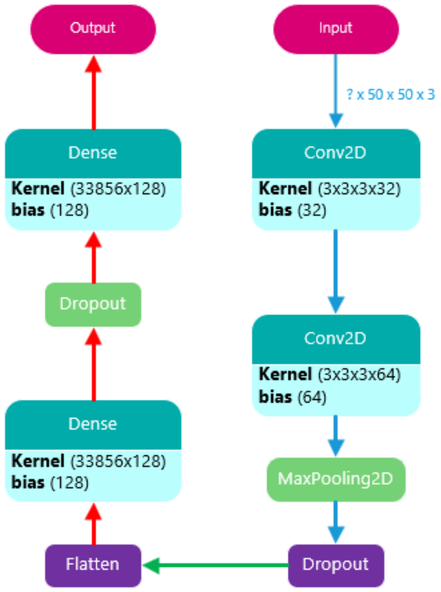


Fig.2 Neural network architecture

4. The Effect of Detection

The confusion matrix, also known as the error matrix, is a standard format for representing accuracy evaluation, represented in the form of a matrix of n rows and n columns. The specific evaluation indicators include overall accuracy, cartographic accuracy, user accuracy, etc., which reflect the accuracy of image classification from different aspects. In artificial intelligence, confusion matrices are visualization tools, especially for supervised learning, and in unsupervised learning, they are generally called matching matrices. The confusion matrix after detection is shown in Fig.3.

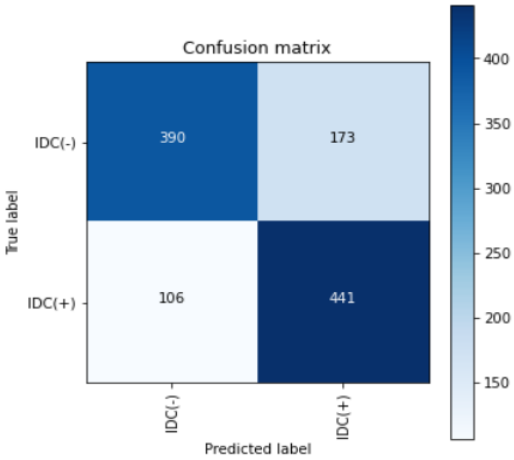


Fig.3 Detect confusion matrix

The confusion matrix shows that the overall recognition accuracy reaches 75%, and the IDC(-) recognition effect can reach 79%. From the graph, we can see that the learning curve shows that the validation set score is always higher than the training set score. The learning curve after the test is shown in Fig.4.

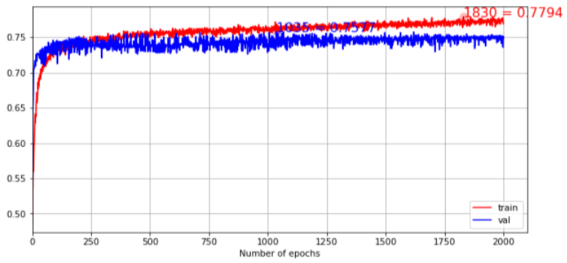


Fig.4 learning curve

The learning curve shows that the training curve and cross-validation curve of different shapes will not have too much overfitting, and both the confusion matrix and the learning curve indicate that the model does not have high bias and has high reference value. In summary, the recognition effect of the model is better.

5. Conclusion

Machine learning technology is an important branch of artificial intelligence, and it is proposed that machine learning-based technology is expected to play a greater role in cancer screening. The times are constantly developing, and human society, which is entering the intelligent era, is facing new challenges and opportunities. This paper uses convolutional neural networks for breast cancer recognition, which has a good recognition effect and can help doctors carry out preliminary screening, thereby greatly saving time.

References

1. Wang J, Liu Q, Xie H, et al. Boosted efficientnet: Detection of lymph node metastases in breast cancer using convolutional neural networks. *Cancers*, 2021, 13(4): 661.
2. Masud M, Eldin Rashed A E, Hossain M S. Convolutional neural network-based models for diagnosis of breast cancer. *Neural Computing and Applications*, 2020: 1-12.
3. Zarbega T S A, Gültepe Y. Semantic segmentation of cell nuclei in breast cancer using convolutional neural network. *International Journal of Computer Applications*, 2020, 176(22): 1-8.
4. Ahmed S R A, UÇAN O N, Duru A D, et al. Breast cancer detection and image evaluation using augmented deep convolutional neural networks. *Aurum journal of engineering systems and architecture*, 2018, 2(2): 121-129.

Authors Introduction

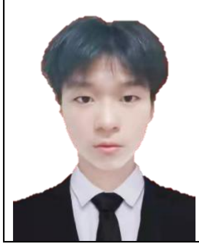
Mr. Zhiyang Li

He is a second-year undergraduate student at Tianjin University of Science and Technology. His research interests are computer vision and machine learning.

Mr. Haoyu Guo

He is currently studying in the School of Biology of Tianjin University of Science and Technology and has achieved certain results in related professional competitions.

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. Wentao Kuang

He is currently pursuing his bachelor's degree at the School of Mechanical Engineering, Tianjin University of Science and Technology.

Ms. Lu Chen

She is currently studying for an undergraduate degree at Tianjin University of Science and Technology.

Intelligent Infusion Service Based on Open MV

Haoran Gong*, Zongyi Li, Qi Chu, Siyuan Liu, Feiyang Qu, Lu Wang

Tianjin University of Science and Technology, Tianjin, China

*mail: *1465157817@qq.com*

www.tust.edu.cn

Abstract

This paper presents an intelligent infusion service system designed to reduce the workload of doctors and nurses and for the health of patients. The system is applied to the drip stand to create a "new type of drip stand" that will effectively solve a number of problems. It is dedicated to: "intelligent infusion process, improving the efficiency of health care workers and reducing their work stress and psychological burden". It embodies the concept of economy and structural rationality.

Keywords: intelligent infusion service system, new type of drip stand, economy, structural rationality

1. Introduction

Intravenous infusions are not only fast-acting, but they are also a widespread and effective medical treatment as they can be administered intravenously to reach the whole body quickly through the blood circulation. It is a widespread and effective method of medical treatment.

However, sometimes there are too many patients for the doctor or nurse to change the fluid or tube in time. When the pressure in the infusion line is lower than the pressure in the blood vessels, the blood will flow out of the infusion line, which is often referred to as "bleeding back". If the bleeding is prolonged, the blood will coagulate, causing clotting of the needle, which can lead to medical accidents and disputes between doctors and patients [1].

This can lead to medical malpractice and disputes between doctors and patients, or even death in severe cases!

Moreover, most of today's young people are only children, so they have to go to the hospital alone when they are sick and have no one to accompany them. When they get a transfusion, it can be very difficult to move around. The tall and bulky IV stand is seen by the patients as a jackal, blocking their access to toilets, water, food and so on.

2. Design Bbrief

Based on the idea of "reducing the work pressure of doctors and nurses, and for the sake of patients' health", we have designed an intelligent infusion service system, which will be applied to the drip stand to create a "new type of drip stand", which will effectively solve a series of problems, as shown in Figure 1

.

2.1. *Highly accurate feedback on fluid levels*

The intelligent infusion service system, which is applied to the IV stand, enables real-time image acquisition of the height of the medicine in the IV bottle via the camera on top of the stand and combines it with an algorithm to calculate whether the bottle has run out of medicine and needs to be changed [2]. If it needs to be changed, the system connects to the hospital's nursing station via a remote WIFI module via a local area network, allowing the nurses to change it at the first opportunity.This allows the nurses to receive the information in time to change the medication for the patient, ensuring that every patient is attended to. This ensures that every patient is attended to.

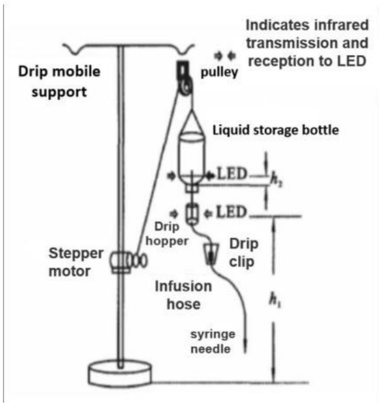


Figure 1. A conceptual model for smart drip

2.2. *drip rack mobile system*

The intelligent infusion service system also allows for patient monitoring via a pan-tilt camera mounted on the vertical bar of the drip stand. Facial recognition" is used to identify and bind the patient through facial analysis. The camera can rotate 360 degrees to see where the patient is The camera can be rotated 360 degrees to observe the patient's position and follow him/her, with the aid of a PID algorithm to set the following distance by means of feedback and negative feedback. The PID algorithm assists in the adjustment of feedback and negative feedback to set the following distance, so that the drip holder can follow the patient's walking and move to achieve "freeing the patient's hands" and realising the true "one person, one lever". This allows the patient to follow the patient's movement, freeing up the patient's hands and enabling true "one person, one lever" auto-following.

2.3. *Privacy security system*

When a patient finishes using the drip stand, the healthcare provider simply restarts the system and the previous patient's information is automatically erased. This not only allows for repeated and efficient use of the 'intelligent' drip rack, but also ensures that the privacy of the previous patient is not compromised. This not only allows for repeated and efficient use of the 'intelligent' drip rack, but also ensures that the privacy of the previous patient is not compromised.

2.4. *Design Advantageization*

1. Intelligent infusion process [3].
2 . Improve the work efficiency of medical and nursing staff.
3 . Reduce the work pressure and psychological burden of health care workers
4. Reduces the risk of blood return and air entering the infusion
5. Automatic sound and light alarm when the infusion is about to be completed, no need to "keep an eye on the bottle", reflecting the concept of humanization.
6. Suitable for all colours of medicine.
7. Simple structure, small size, light weight, reflecting the concept of simplicity;
8. Low price, can be used repeatedly;
9. Sensitive response, safe and reliable;
10. No electromagnetic radiation interference;
11. Mechanical monitoring, no contact with the liquid, safe and hygienic;
12. It is possible to infuse each bottle of medicine as much as possible without wasting the medicine, reflecting the concept of saving;
13. The infusion control device effectively ensures that the working schedule of the medical staff is reasonable.

3. Design process

3.1. *Total design solution*

The intelligent infusion service system, which is applied to the IV stand, enables real-time image capture of the height of the medicine in the IV bottle via the Open MV camera on top of the stand and combines it with an algorithm to calculate whether the bottle has run out of medicine and needs to be replaced. This is shown in Figure 2 below [4].

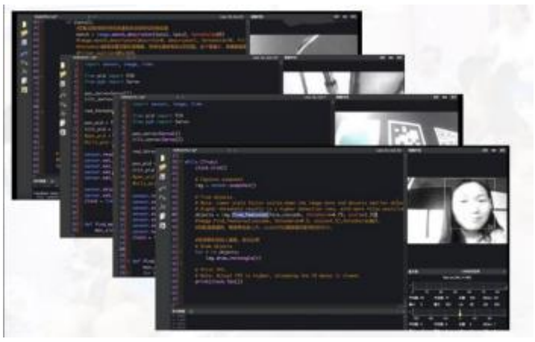


Figure 2. Open MV identifies the persona

If it needs to be changed, the system connects to the hospital's nurses' station via a remote WIFI module via a local area network, allowing nurses to receive the information in time to change the medication for the patient, ensuring that every patient is attended to.The Intelligent Infusion Service system also allows the patient to be identified and bound to the patient through facial analysis by means of a pan-tilt camera mounted on the vertical pole of the drip stand, while the distance between the patient and the infusion pole is determined through a PID algorithm, as shown in Figure 3 below, and followed in real time to fully address the problem of patient movement.

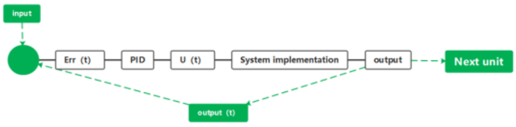


Figure 3. PID unit cycle once process

The intelligent infusion service structure includes various functions such as camera acquisition, wireless remote control alarm, vision and other sensors reading information, automatic following, Bluetooth control, IOT management system, etc. as shown in Figure 4.

In summary, we designed the original idea to free nurses and caregivers to the maximum extent possible and to facilitate the development of the medical profession. To achieve technological intelligence + medical treatment.

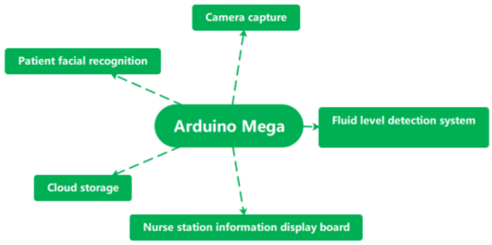

Camera capture
Patient facial recognition
Arduino Mega
Fluid level detection system
Cloud storage
Nurse station information display board

Figure 4. Operating applications of the control panel

3.2. *Introduction to PID*

Next we look at the basic PID equation, as shown in equation (1) below [5].

$$u(t)=Kp[e(t)+\frac{1}{Ti}\int_0^t e(t)dt+Td\frac{de(t)}{dt}] \quad (1)$$

There are two general expressions for PID, a positional PID and an incremental PID, which calculates the change in value of the output of the system in the current control system.

At present, the intelligent infusion system we have designed uses incremental PID, which calculates the difference between the current position and the last position and the expected position, so that the system can quickly and accurately track the patient (patient) in order not to rip off the infusion tube or bottle. Safety problems are thus avoided to a great extent.

The above formulae are collated and deformed to give the following equation, as in equation (2).

$$u(t)=Kp\times e(t)+Kp\times\frac{1}{Ti}\int_0^t e(t)dt+Kp\times Td\times Td\times\frac{de(t)}{dt} \quad (2)$$

4. Example of operation

4.1. *Method of use*

We have simplified the use into the following steps:

Step 1: After the infusion is normal, clip the upper end of the infusion tube into the detection slot.

Step 2: Press the S/R (Start/Reset) key, the alarm will sound "Bi-Bi" twice and the alarm will turn on.

Step 3: Press the S/R key again, the alarm will sound "Bi" and the green light will flash, entering the initialisation state.

Step 4: After about 15s, the alarm will give a long "Bi" sound and the green light will flash slowly to enter the working state.

Step 5: When the infusion is completed, a continuous "Bi-Bi" beep will sound and the green light will flash rapidly, and the infusion tube will be blocked.

Step 6: Press the S/R key, the alarm will sound "Bi-Bi" twice and the pressure wheel will release the infusion tube.

Step 7: Switch off: Press and hold the S/R key, the green light will turn off after about 3s and the alarm will switch off.

Step 8: Charging: Plug in the charging transformer in the off state and press and hold the S/R key for 5 seconds, then release the key when the green light comes on for the second time. After about 4 to 8 hours the red light goes out, indicating that charging is complete. The alarm can be used normally in the charged state.

Step 9: When the alarm is working, a short "Bi" tone is emitted every 30 seconds, indicating that the battery is low and needs to be charged. Plug in the charging transformer to recharge immediately. If the alarm is not recharged after approximately 10 minutes, the alarm will automatically switch off after the "Bi" tone and the infusion will be blocked.

4.2. *Illustrated overview*

To make it easier to understand, we give a picture to explain, which shows the components of the system, the interconnection between the various positions [6], in layman's terms and easy to understand as follows in Figure 5.

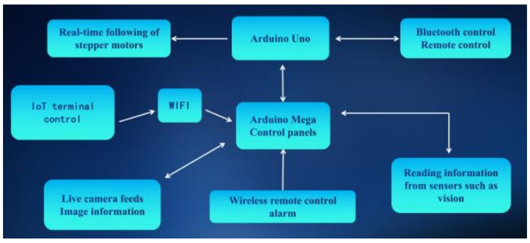

Real-time following of stepper motors
Arduino Uno
Bluetooth control Remote control
IoT terminal control
WIFI
Arduino Mega Control panels
Reading information from sensors such as vision
Live camera feeds Image information
Wireless remote control alarm

Figure 5. Single system linkage map for each service

5. Concluding remarks

The rapid development of modern technology today has made life faster and medical care + smart technology has become the mainstream of society. In the future, most of the traditional medical treatments will be replaced by modern medical treatments and the information age will usher in a "revolutionary upgrade" of modern medical treatments. This not only reflects a comprehensive upgrade in the medical field, but also gives rise to many more applications. Technology for good, high technology can create even better conditions for the general public.

We have every reason to believe that in the near future we will be able to fit in medical technology that will help patients to overcome their difficulties and be widely appreciated by them [7]. We have greater confidence in this, because it is the country's hope for the future of human healthcare and a satisfactory answer to a major problem such as human health.

References

1. Liu Yiping. Why does blood flow back during infusion? [J]. Discussion on Physics Teaching, 2013, 31(10): 63.
2. Xu W. Design of Smart Home monitoring System based on WiFi and Android [D]. Southwest Jiaotong University, 2017.
3. Ming Chen. Based on the application of Internet of things of intelligent medical system [J]. Computer programming skills and maintenance, 2022(11): 125-127. doi:10.16184/j.cnki.com.prg. 2022.11.023.
4. WEI S J. Design of a Face recognition device with Complementary light Performance [J]. Electronic Technique, 2022, 51(02): 172-173. (in Chinese)
5. SHENG Liuqing. Application of Adaptive PID Control Algorithm in Intelligent Vehicle Control System [J]. Journal of Beijing institute of printing, 2021, 29 (5) : 147-150. The DOI: 10.19461/j.carol carroll nki. 1004-8626.2021.05.046.
6. Ye Yun. Research on the application of Internet of Things technology under the background of Big data era [J]. Industrial Innovation Research, 2022 (22): 76-78.
7. Liao dragon. Wisdom medical information platform construction and application [J]. Computer knowledge and technology, 2022, 17 (19) : 55-57. DOI: 10.14004 / j.carol carroll nki CKT. 2022.1353.

Authors Introduction

Mr. Haoran Gong

He participated in the American College Mathematical Contest in Modeling as a modeler and data analyst in early 2022 and won the s Award. During his freshman year, he published one international CPCI paper. In 2022, he attended the "Hi cool" Global Entrepreneur Summit with his mentor.

Mr. Zongyi Li

He is a 2021 undergraduate student of robotics engineering in the School of Telecommunications of Tianjin University of Science and Technology. He has participated in competitions such as the North China Five Provinces, the Provincial Three of the Electronic Design Competition and won two bronze medals in the Challenge Cup School.

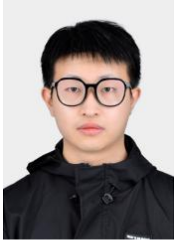
Mr. Qi Chu

He is a current undergraduate student of Tianjin University of Science and Technology, majoring in Robotics Engineering, Class of 2021. He won a scholarship at the university level. In 2022, he won the second prize in the North China Robotics Competition and the third prize at the municipal level.

Mr. Siyuan Liu

He will receive his bachelor's degree from Tianjin University of Science and Technology in China in 2024. His research interests include robot dynamics and attitude calculation of unmanned aerial vehicles.

Mr. Feiyang Qu



He is an undergraduate student of Robotics Engineering in the School of Electronic Information and Automation of Tianjin University of Science and Technology. He has a good command of C++, Python and other languages, CAD, UG, PrAe and other software, and has participated in mathematical modeling competitions.

Ms. Lu Wang

She is an undergraduate student of Light Industry Science and Engineering at Tianjin University of Science and Technology. She was awarded a university-level scholarship and has participated in the Student Innovation and Entrepreneurship Competition and the Internet Plus Competition, winning places respectively.

Adaptive STDP Learning with Lateral Inhibition for Neuromorphic Systems

Ashish Gautam
Institute of Industrial Science, The University of Tokyo, 4-6-1 Komaba,
Meguro, Tokyo 153-8505, Japan

Takashi Kohno
Institute of Industrial Science, The University of Tokyo, 4-6-1 Komaba,
Meguro, Tokyo 153-8505, Japan

E-mail: asgautam@iis.u-tokyo.ac.jp , kohno@g.ecc.u-tokyo.ac.jp

Abstract

Implementing biologically plausible learning rules on neuromorphic chips is essential to explore the learning mechanisms in the brain. Spike-timing dependent plasticity (STDP) is one such rule but its multi-bit circuit implementation requires too much area. In our previous study, we proposed a hardware-friendly learning rule named adaptive STDP and experimentally showed that its performance was similar to STDP learning in a very basic biologically plausible spike pattern detection task using a single neuron. In this study, we extend the adaptive STDP learning rule with lateral inhibition, a common motif observed in the brain, and apply to a spike pattern detection model with multiple neurons that compete to detect multiple patterns. Our results show that the performance is similar to that with STDP learning.

Keywords: Adaptive STDP, Pattern detection, Synaptic efficacy, Neuromorphic chips, Lateral inhibition

1. Introduction

Significant amounts of resources are being spent worldwide to understand the information processing and learning mechanisms in the brain. Design of neuromorphic chips with biologically plausible neuron and synapse circuits is an integral part of this effort [1]. In the brain, neuronal cells interact at the microcircuit level via various (yet unknown) network motifs. Exploration of these motifs via a bottom-up construction at the level of neuronal cells and synapses is crucial to improve our understanding of microcircuits in the brain. Neuromorphic chips contribute to this exploration via this “analysis by synthesis” methodology. Their neuron and synapse circuits are used to create scalable versions of known network motifs and can help develop insights into their network dynamics via real-time emulation.

In this study, we focus on the neuromorphic implementation of one such widely observed network motif, lateral inhibition with spike-timing dependent plasticity (STDP), and present post-silicon-validated circuit models for the same. Lateral inhibition is easily implemented in neuromorphic chips but ideal STDP learning models generally require synapse circuits with very high efficacy resolution (>10 bits). Resolution of synapses in the brain is not clearly known. However, recent findings with high resolution imaging techniques show that individual synapses have multi-bit resolution [2]. A high resolution is achieved in pure analog circuits that store efficacy on a capacitor [3]. However, due to capacitor leakage, the learned efficacy is lost over time. Bi-stable synapse circuits [4] compensate for this leakage and have two long-term stable states but have low resolution of about 1.5 bits. In mixed-signal circuits, efficacy is implemented using digital-to-analog converters (DACs) and digital memories. The area and power consumption of DAC doubles for every 1-bit increase in its resolution. Hence, even though lower-resolution synapses deteriorate the performance of STDP learning, the synapse resolution is restricted to about 4 to 6 bits in most chips [1], [5].

As a potential solution to this problem, an adaptive STDP learning rule that uses just 4-bit synaptic efficacy was proposed in a previous study [6]. It was successfully applied to a very basic biologically plausible spike pattern detection model comprising a single neuron and its performance in numerical simulations as well as on a mixed-signal neuromorphic chip was demonstrated to be similar to ideal STDP learning [6], [7]. In this study, the application of adaptive STDP is validated on a more realistic biologically plausible network comprising multiple neurons that laterally inhibit each other. Models of this motif were described in a previous study that used the ideal STDP learning rule and synapses with 64-bit floating-point resolution efficacy [8]. We use the same motif but apply adaptive STDP learning rule with 4-bit synapses and compare the performance with ideal STDP learning.

The rest of the manuscript is organized as follows. The network model with details of the input spike trains, neuron, synapse, and the learning rule is presented in section 2. Results are presented in section 3 and the final section concludes with a discussion of the results and possible future works.

2. Network Model

The target network comprises nine neurons that inhibit each other in an all to all fashion. Each of them receives stochastic input spikes via 2048 4-bit excitatory synapses empowered with adaptive STDP learning. Three different spike patterns are embedded repeatedly in these stochastic spike trains and the goal of the network is to recognize the embedded spike patterns. The input spike trains are similar to the ones used in the reference study [8]. They assume only the statistical properties generally assumed in neuroscience. The embedded spike patterns are characterized purely by spike times mimicking the temporal neural code observed in various regions of the brain. The procedure to generate the spike train is described next.

2.1. *Input Spike Train Model*

Input spike trains for each afferent (total 2048 afferents) were generated via an inhomogeneous Poisson process with the instantaneous firing rate varying between 0 to 90 Hz. The smallest interval for the frequency to change from 0 to 90 Hz was 50 ms. Each afferent spiked at least once in 50 ms establishing 54 Hz as the average frequency of each spike train. Upon generation of 225 s long stochastic spike train, three randomly selected 50 ms long segments (spike patterns to be embedded) were copied from it. This ensures that the spike patterns have similar characteristics to the stochastic Poisson spikes. Next, the spike train was divided into 50 ms long sections and based on the chosen pattern appearance frequency (11.1% for each pattern) a certain number of these sections were replaced by the spike patterns to be embedded. In other words, the three spike patterns occupy one-third of the total simulation time. Additionally, a Gaussian jitter with zero mean and 1ms standard deviation is added to the spike patterns during the copy-paste procedure. This copy-and-replace procedure is applied only to half of the randomly chosen afferents. The remaining afferents encode only stochastic spikes. In this copy-and-paste process, consecutive 50 ms sections were avoided. The population average spiking rate inside and outside the 50 ms long spike patterns is the same and the patterns are characterized only by the precise spike timing of the afferents. Subsequently, 10 Hz spontaneous spikes were added to all the spike trains increasing the population average firing rate of the afferents to approximately 64 Hz.

2.2. *Neuron, Synapse, and Learning Models*

Structure of the network is shown in Fig.1. The neurons are modeled using the reduced compartment technique. It comprises two compartments, a passive dendritic compartment (shaded grey in Fig.1) and an active somatic compartment (shaded green) that can generate neuronal spikes. A unidirectional resistor (R_c) connects them and only allows the current to flow into or out of the somatic compartment based on their potential difference. Detailed arguments for choosing these models, their ideal model equations, and their circuit details are described in our previous studies [6], [7]. Here we focus on the implementation of lateral inhibition.

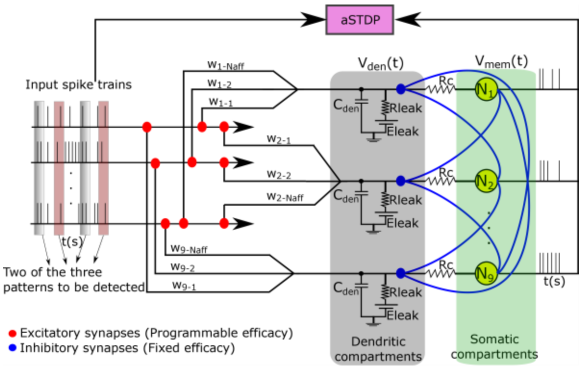


Fig. 1: Network for lateral inhibition with adaptive STDP learning. $\boldsymbol{N_{aff}}$ is the number of afferents.

To implement the lateral inhibition, the somatic compartment of every neuron is connected to the dendritic compartments of other neurons via inhibitory synapses. With successive coincidences of spike patterns, the adaptive STDP learning rule directs the neurons to spike in the presence of these patterns [6]. For every neuronal spike, each neuron hyperpolarizes the dendritic membrane potential of all other neurons via inhibitory connections (blue lines in Fig.1). The dynamics of the dendritic compartment are given by:

$$C_{den}\frac{dV_{den}}{dt} = I_{syn\_exc} + I_{syn\_inh} - \frac{V_{den}-E_{leak}}{R_{leak}}, \qquad (1)$$

where V_{den} is the dendritic membrane potential, C_{den} is the dendritic capacitance fixed at 48 pF, R_{leak} and E_{leak} are the parameters of the leak resistor set at 50 MΩ and 305 mV, respectively. I_{syn_exc} is the excitatory synaptic current generated by input synapses and I_{syn_inh} is the inhibitory synaptic current that implement lateral inhibition. They are described by

$$I_{syn}(t) \ = \ I_{sw}.\left(-exp^{-t/\tau_r} + exp^{-t/\tau_d}\right)\Big/a_{scale}, \quad (2)$$

where τ_r and τ_d together control the rising and falling time constant of the synaptic current, their values are set at 1ms and 3ms (10ms) for I_{syn_exc} (I_{syn_inh}), respectively. The scaling factor a_{scale} sets the amplitude of the synaptic current equal to the value of the synaptic efficacy denoted by I_{sw}. The efficacy of inhibitory synapses is constant. The initial synaptic efficacies of excitatory synapses are set randomly and are modified by the adaptive STDP learning rule. It is given by

$$\Delta w_j = \begin{cases} +1 \text{ bit, if } t_j \le t_i, t_i - t_j < t_{\text{pre}}, \text{and w} < w_{\max} \\ -1 \text{ bit, if } t_j > t_i, t_j - t_i < t_{\text{post}}, \text{ and w} > w_{\min} \end{cases}, \quad (3)$$

where t_j (t_i) represents the timing of the presynaptic (postsynaptic) spike and t_{pre} (t_{post}) is the maximum delay of the postsynaptic (presynaptic) spike after the presynaptic (postsynaptic) spike that leads to potentiation, LTP (depression, LTD). The 4-bit efficacy saturates at its maximum and minimum values of $w_{\max}$ and $w_{\min}$ representing synaptic current of values 15 pA and 0 pA, respectively. The learning parameter t_{pre} is 13 ms and t_{post} is adapted from 13 ms to increasingly higher values during the learning process saturating at 40 ms.

3. Results

Here we present the simulation results quantifying the performance of the network in Fig. 1. Simulations were run for 100 different inputs generated by the procedure described in section 2 and a run was considered to be successful if all three spike patterns were detected with a hit rate greater than 98 % and false alarms under 1 Hz in the last 75 s of the run (one-third of the total run time), similar to the criterion in [8].

The observed success rate was 70%, that is in 70 out of 100 runs, all three spike patterns were detected by at least 1 neuron. In 24 runs, only two spike patterns were detected and in 6 runs only 1 spike pattern was detected. These results are similar to the reference study that used STDP learning on the same task where it is reported that in more than two-thirds of the cases, all three spike patterns were detected by at least one neuron [8]. The superimposed membrane potentials of the neurons during a successful run are shown in Fig.2(A) and the same in the last second of the run is shown in Fig.2(B). Three different color-shaded regions represent the 50-ms long spike patterns within which the neurons learn to spike.

4. Discussion

The results above demonstrate that adaptive STDP learning with 4-bit synapses when applied to a lateral inhibitory network has a performance similar to STDP learning with high-resolution synapses. The circuits (of neuron, synapse, and learning rule) whose ideal models were used for simulation have been validated in silicon in previous studies [6], [7] and can be used for large-scale implementation and exploration of lateral inhibitory networks on neuromorphic chips.

The network motif, input spike train model, and biomimetic somatic compartment were chosen because of their biological plausibility. The reference study [8] used the single-compartment leaky integrate-and-fire

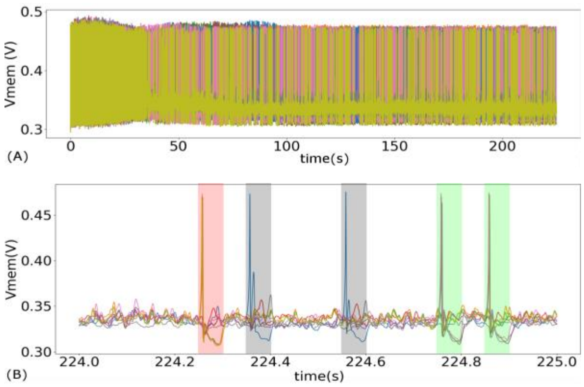


Fig. 2: Superimposed membrane potentials of nine neurons (A) for the entire duration of a run; (B) in the last second of the run. The neurons learn to spike within the patterns.

(LIF) neuron model. The unidirectional two-compartment neuron model used in this study is not biomimetic but is more biologically plausible than the widely used single-compartment neuron model as it incorporates the passive properties of the dendrites (ignoring their spatial morphology). The two compartments also provide flexibility to apply inhibition to either one of them or use biomimetic shunting inhibitory synapses for inhibition. The variations in performance and network dynamics due to these modifications will be explored in future studies.

All network parameters were tuned manually but the network is not sensitive to precise parameter values and similar results were obtained using a wide range of parameters. Their values can be chosen either based on biophysical values or to minimize silicon area (e.g., small C_{den} and high R_{leak}).

The network was limited to a single layer and the neurons learned a small segment in the 50-ms long patterns. As seen in Fig.2(B), the neurons learn to spike near the beginning of the patterns and thus only detect the initial segment of the pattern. Multiple layered network may detect patterns of longer duration. Also, during learning, only the parameter controlling depression, (t_{post}) was adapted externally. The potentiation parameter t_{pre} can also be adapted while learning and these adaptations can be made dependent on the network state rather than via external programming. These potential modifications will be explored in future studies.

Acknowledgment

This study was partially supported by JSPS KAKENHI Grant Number 21H04887, DLab, The University of Tokyo in collaboration with Cadence Design Systems, Inc.

References

[1] C. Pehle *et al.*, “The BrainScaleS-2 Accelerated Neuromorphic System With Hybrid Plasticity,” *Front Neurosci*, vol. 16, p. 158, Feb. 2022.
[2] K. K. L. Liu, M. F. Hagan, and J. E. Lisman, “Gradation (approx. 10 size states) of synaptic strength by quantal addition of structural modules,” Philos Trans R Soc Lond B Biol Sci, vol. 372, no. 1715, Mar. 2017.
[3] M. R. Azghadi, N. Iannella, S. F. Al-Sarawi, G. Indiveri, and D. Abbott, “Spike-based synaptic plasticity in silicon: Design, implementation, application, and challenges,” Proceedings of the IEEE, vol. 102, no. 5, pp. 717–737, Apr. 2014.
[4] N. Qiao et al., “A reconfigurable on-line learning spiking neuromorphic processor comprising 256 neurons and 128K synapses,” Front Neurosci, vol. 9, no. 141, pp. 1–17, Apr. 2015.
[5] J. Schemmel, A. Grübl, K. Meier, and E. Mueller, “Implementing synaptic plasticity in a VLSI spiking neural network model,” in Proc. IEEE Int. Conf. on Neural Netw., Jul. 2006, pp. 1–6.
[6] A. Gautam and T. Kohno, “An Adaptive STDP Learning Rule for Neuromorphic Systems,” Front Neurosci, vol. 15, pp. 1–12, 2021.
[7] A. Gautam and T. Kohno, “Adaptive STDP-based On-chip Spike Pattern Detection,” TechRxiv Jun. 2022; preprint.
[8] T. Masquelier, R. Guyonneau, and S. J. Thorpe, “Competitive STDP-based spike pattern learning,” Neural Comput, vol. 21, no. 5, pp. 1259–1276, May 2009.

Authors Introduction

Ashish Gautam

He received his B. Tech degree in electronics and instrumentation engineering in 2012 from Pondicherry University, India. He received his PhD degree in electrical engineering in 2021 from the University of Tokyo, Tokyo, Japan where he is currently a Project Researcher.

Takashi Kohno

He has been with the Institute of Industrial Science at the University of Tokyo, Japan since 2006 where he is currently a Professor. He received the BE degree in medicine in 1996 and the PhD degree in mathematical engineering in 2002 from the University of Tokyo, Tokyo, Japan.

Spike pattern detection with close-to-biology spiking neuronal network

Takuya Nanami
Institute of Industrial Science, The University of Tokyo, Tokyo, Japan.

Takashi Kohno
Institute of Industrial Science, The University of Tokyo, Tokyo, Japan.

E-mail: nanami@iis.u-tokyo.ac.jp, kohno@g.ecc.u-tokyo.ac.jp

Abstract

The nervous system contains a variety of neuron types, each with different electrophysiological properties. However, their roles in information processing are poorly understood. Using the piecewise quadratic neuron (PQN) model that can reproduce a variety of electrophysiological properties, we demonstrate that performance on a biologically plausible task of spike pattern detection varies depending on the electrophysiological properties.

Keywords: Spiking neural network, Pattern detection, PQN model

1. Introduction

In the nervous system, neurons exhibit a variety of electrophysiological properties. For example, when given a gradually increasing stimulus input, Class I neurons in the Hodgkin's classification [1] start firing at a frequency close to zero, whereas Class II neurons start firing at a certain higher frequency. In addition, when given a step current input, neurons in the regular spiking class exhibit spike-frequency adaptation, in which the firing frequency gradually decreases, whereas neurons in the fast spiking class fire at an almost constant frequency. Furthermore, bursting neuron classes such as elliptic bursting and parabolic bursting are known to alternate between intense firing and resting state in response to constant stimuli. The mathematical structures behind these diverse electrophysiological properties of neurons have been intensively investigated, and spiking neuron models that reproduce them have also been developed [2][3][4][5][6][7][8][9][10][11][12][13]. However, the role of these electrophysiological properties of neurons in information processing in the nervous system remains unclear. This is because the information processing mechanisms in the brain microcircuits are too complex and poorly understood to assess the contribution of each electrophysiological properties. In addition, there are various technical hurdles in measuring neuronal activities in vivo, which also make it difficult to analyze the role of electrophysiological properties.

In this study, we investigate the relationship between the electrophysiological properties of neurons and the learning properties of a network using in silico simulations. We employ simple and biologically plausible network structure and task proposed in the previous studies [14][15]. For the spiking neuron model, we adopt the piecewise quadratic neuron (PQN) model [13], which is a lightweight model that can reproduce a wide variety of electrophysiological properties. Although there are a variety of dynamics behind the neuronal activities, here we focus on the most fundamental dynamics, the spike generation. We investigate how the success rate of pattern detection changes while varying the mathematical structure of the fast subsystem responsible for the spike generation. We show how the success rate depends on the dynamical structure in the fast subsystem.

The remainder of this paper is organized as follows: Section 2 describes the details of the methods, Section 3 shows the simulation results, and Section 4 summarizes the work and suggests ideas for future.

2. Methods

2.1. *Network model*

The network structure, input data, and learning rule, which we briefly describe in this section, are largely the same as those proposed in the previous study [15]. The differences are in the neuron and synapse models, the number of input nodes (reduced from 2000 to 512), and the number of patterns to be detected. Figure 1 shows the network architecture comprising input nodes, output neurons, and excitatory and inhibitory synapses. 512 input nodes transmit spike trains to 9 output neurons through excitatory synapses, whose synaptic efficacies are updated based on the spike timing-dependent plasticity (STDP) rule. The output neurons are the PQN model. They are connected by the inhibitory synapses each other, whose synaptic efficacies are fixed. A stimulus input to the j-th neuron is written as follows:

$$I_{jk} = p_0 \sum_{i=1}^{512} w_{ji}^{e} x_i - p_1 \sum_{i=1(i\neq j)}^{9} w_{ji}^{i} s_i \qquad (1)$$

where x_i represents the activation of the i-th input node, which is 0 or 1. s_i is an inhibitory synaptic current from the i-th neuron. w_{ji}^{e} is the efficacy of the excitatory synapse from the i-th node to j-th neuron. w_{ji}^{i} is the efficacy of the inhibitory synapse from the i-th neuron to j-th neuron. Parameters p_0 and p_1 control the amount of excitatory and inhibitory stimulus.

Figure 2 shows an example of the input data. Three different spike patterns (colored dots) are embedded in random spikes (black dots). These spike trains are generated based on the Poisson process. First, an 80-second training period is given to learn excitatory synapses based on the STDP. Then, a 20-second test period is given to assess whether the output neurons are firing in response to the spike patterns. The task is considered successful when all spike patterns are detected by at least one output neuron with a probability of 90% or greater during the test period.

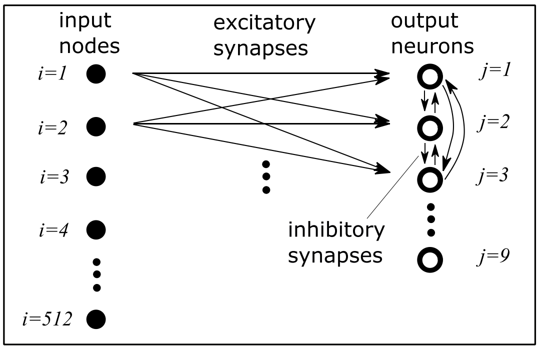


Figure 1. Network architecture.

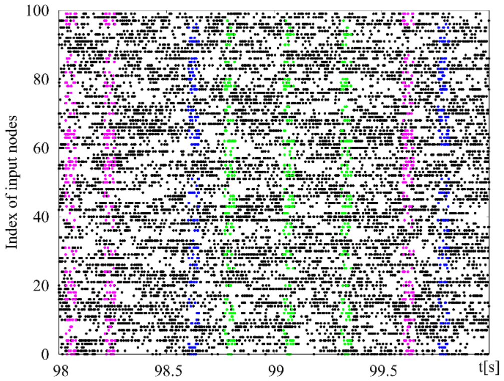


Figure 2. Input data.

The rule for updating the synaptic efficacy based on the STDP is given by

$$\Delta w_{ji}^{e} = \begin{cases} a^{+} \exp\left(\dfrac{t_i - t_j}{\tau^{+}}\right) & \text{if } t_i \le t_j, \\ a^{-} \exp\left(-\dfrac{t_i - t_j}{\tau^{-}}\right) & \text{if } t_i > t_j. \end{cases} \qquad (1)$$

where t_j is the spike timing of the presynaptic node and t_j is the spike timing of the postsynaptic neuron. τ^{+} and τ^{-} control the time constant of the long-term potentiation and depression, respectively. a^{+} and a^{-} determine the learning rate.

2.2. *Neuron model*

The PQN model is a qualitative spiking neuron model that supports a wide variety of spiking properties with a limited computational cost. The equations for the PQN model in two-variable mode are as follows:

$$\frac{dv}{dt} = \frac{\emptyset}{\tau}\left(f(v) - n + I_0 + kI_{stim}\right), \qquad (1)$$

$$\frac{dn}{dt} = \frac{1}{\tau}\left(g(v) - n\right), \qquad (2)$$

$$f(v) = \begin{cases} a_{fn}\left(v - b_{fn}\right)^2 + c_{fn} & (v < 0) \\ a_{fp}\left(v - b_{fp}\right)^2 + c_{fp} & (v \ge 0), \end{cases} \qquad (3)$$

$$g(v) = \begin{cases} a_{gn}\left(v - b_{gn}\right)^2 + \mathrm{c}_{\mathrm{gn}} & (v < r_g) \\ a_{gp}\left(v - b_{gp}\right)^2 + c_{gp} & (v \ge r_g), \end{cases} \qquad (4)$$

where v and n correspond to the membrane potential and recovery variable, respectively. Parameter I_0 is a bias constant. I_{stim} is the stimulus input and k is its scaling parameter. The parameters τ and $\emptyset$ determine the time constants of the variables. The parameters r_g, a_x, b_x, and c_x, where x is fn, fp, gn, or gp, are constants that determine the nullclines of the variables. Constants b_{fp}, c_{fp}, b_{gp}, and c_{gp}, are determined by other parameters such that the nullclines are continuous and smooth (see [13]). All of the variables and parameters are purely abstract with no physical units.

We prepared parameter sets, each with different dynamics of the spike generation, by varying the two key parameters, a_{gn} and $\emptyset$. The parameter a_{gn} determines the slope of the left portion of the n-nullcline. For example, if the a_{gn} is reduced from 0.75 to 0, the slope on the left portion of the n-nullcline decreases and the position of the bifurcation point shifts to the right (Fig. 3 (A-B)). The parameter $\emptyset$ controls the time constant of v and affects the trajectory of the stable limit cycle. If the $\emptyset$ increases from 0.35 to 1, the trajectory is extended especially in the negative direction of the v-axis (Fig. 3 (B-C)). We prepared twelve parameter sets consisting of the combination of four different values of a_{gn} (1.75, 0.75, 0.25, and 0) and three different values of $\emptyset$ (0.35, 0.5, and 1).

3. Results

Figure 4 shows the success rate for each parameter set. We varied the number of patterns from 1 to 5 and measured the average success rate. We performed 128 trials for each number of patterns, resulting in a total of 640 trials. For fair comparison, parameters, I_0, τ, p_0, and p_1 were chosen to achieve the best success rate. First, the values of these parameters were determined randomly, and then these parameters were iteratively updated by the differential evolution (DE) algorithm [16] while evaluating the success rate. We used 64 individuals and 200 generations for the DE algorithm. The results show that as the a_{gn} becomes smaller and $\emptyset$ becomes larger, a higher success rate tends to be obtained.

4. Conclusion

In this study, we prepared twelve types of dynamical structures in the neurons and investigated how the differences in their dynamics affect performance in a biologically plausible task. The results showed that better performance was obtained when the a_{gn} is small and $\emptyset$ is large. As the a_{gn} becomes smaller, the bifurcation

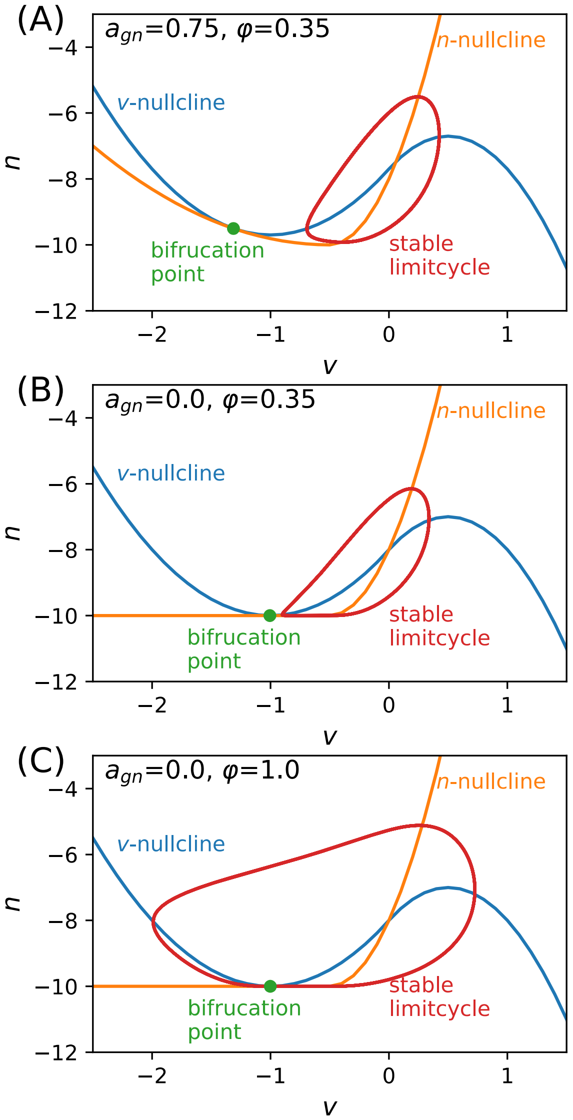


Figure 3. Mathematical structure.

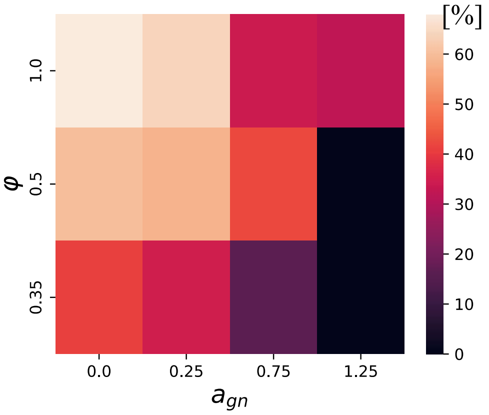


Figure 4. Success rate.

point approaches v = 0. Note that v = 0 is supposed to be the point of transmitter release in the PQN network. When the bifurcation point is close to this point, neurons require less time to fire. Considering the exponential decay of the STDP curve, such neurons can obtain a larger Δw than those fire more slowly. It may cause the higher success rate.

As the $\emptyset$ increases, the stable limit cycle is extended in the negative direction of the v-axis. Consequently, the trajectory goes through a narrow channel and takes a long time to generate the second spike. Considering that inputs of sufficient magnitude are usually instantaneous, such neurons will rarely generate a second spike. Since the second spike may cause pairing with spike signals unrelated to the target pattern, the neurons with a larger $\emptyset$ will stably detect patterns by avoiding the second spike.

In our future work, we will verify these hypotheses and investigate the impact of other neuronal dynamics, such as spike frequency adaptation and bursting, on the success rate of the task.

Acknowledgment

This study was partially supported by JSPS KAKENHI Grant Number 21K17849.

References

1. A. L. Hodgkin, “The local electric changes associated with repetitive action in a non-medullated axon.” The Journal of physiology, vol. 107, no. 2, pp. 165–181, Mar. 1948.
2. X.-J. Wang, “Ionic basis for intrinsic 40 hz neuronal oscillations”, NeuroReport, vol. 5, pp. 221–224, 1993.
3. R. E. Plant, “Bifurcation and resonance in a model for bursting nerve cells”, Journal of Mathematical Biology, vol. 67, pp. 15–32, 1981.
4. J. Rinzel and B. Ermentrout, “Analysis of neural excitability and oscillations”, in Methods in Neuronal Mod- eling, Second Edition, C. Koch and I. Segev, Eds. MIT Press, 1998, ch. 7, pp. 251–292.
5. E. M. Izhikevich, “Simple model of spiking neurons”, IEEE Trans. Neural Networks, pp. 1569–1572, 2003.
6. R. Brette and W. Gerstner, “Adaptive Exponential Integrate-and-Fire Model as an Effective Description of Neuronal Activity”, J. Neurophysiol., vol. 94, pp. 3637–3642, 2005.
7. E. Izhikevich, “Dynamical systems in neuroscience”, MIT Press, p. 111, Jul. 2007.
8. M. Pospischil, M. Toledo-Rodriguez, C. Monier, Z. Piwkowska, T. Bal, Y. Fregnac, H. Markram, and A. Destexhe, “Minimal hodgkin-huxley type models for different classes of cortical and thalamic neurons”, Biological Cybernetics, vol. 99, no. 4-5, pp. 427–441, 2008.
9. T. Nanami and T. Kohno, “Simple cortical and thalamic neuron models for digital arithmetic circuit implementation”, Frontiers in Neuroscience, section Neuromorphic Engineering, vol. 10, no. 181, 2016.
10. T. Nanami, K. Aihara, and T. Kohno, “Elliptic and parabolic bursting in a digital silicon neuron model”, in 2016 International Symposium on Nonlinear Theory and Its Applications, November 2016, pp. 198–201.
11. T. Nanami, F. Grassia, and T. Kohno, “A parameter optimization method for digital spiking silicon neuron model”, Journal of Robotics Networking and Artificial Life, vol. 4, no. 1, pp. 97–101, 2017.
12. T. Nanami, F. Grassia, and T. Kohno, “A metaheuristic approach for parameter fitting in digital spiking silicon neuron model”, Journal of Robotics Networking and Artificial Life, vol. 5, no. 1, pp. 32–36, 2018.
13. T. Nanami and T. Kohno, “Piecewise Quadratic Neuron model : a tool for close-to-biology spiking neuronal network simulation on dedicated hardware”, Frontiers in Neuroscience, section Neuromorphic Engineering, in press.
14. T. Masquelier, R. Guyonneau, and S. J. Thorpe, “Spike timing dependent plasticity finds the start of repeating patterns in continuous spike trains”, PLOS ONE, vol. 3, pp. 1–9, 01 2008.
15. T. Masquelier, R. Guyonneau, and S. J. Thorpe, “Competitive STDP-Based Spike Pattern Learning”, Neural Computation, vol. 21, no. 5, pp. 1259–1276, 05 2009.
16. R. Storn and K. Price, “A simple and efficient heuristic for global optimization over continuous spaces”, Journal of Global Optimization, vol. 11, no. 4, pp. 341–359, 1997.

Authors Introduction

Dr. Takuya Nanami

He received the Ph.D from University of Tokyo in 2019. He is a research associate in Institute of Industrial Science, University of Tokyo since 2020.

Prof. Takashi Kohno

He received the B.E. degree in medicine (1996) and the Ph.D. degree in mathematical engineering (2002) from University of Tokyo, Japan. He is a professor in Institute of Industrial Science, University of Tokyo since 2018.

Object Status Detection in Cluttered Environment for Robot Grasping Using Mask-RCNN

Kasman
Graduated School of Creative Informatics, Kyushu Institute of Technology
Castail Iizuka 318, 268-1, Kawazu, Iizuka-City, Fukuoka, 820-0067, Japan

Eiji Hayashi
Department of Mechanical Information Science and Technology, Kyushu Institute of Technology
680-4, Kawazu, Iizuka-City, Fukuoka, 820-8502, Japan
E-mail: kasman@mmcs.mse.kyutech.ac.jp, haya@mse.kyutech.ac.jp, kasman@umi.ac.id
www.kyutech.ac.jp

Abstract

Detecting object status in cluttered manipulator's robot environment before grasping is quite challenging to recognize the target because of unstructured and uncertainty scenes. Using Mask R-CNN for detecting the status of the object i.e. free for picking, close, overlapping and piling to the other objects is very useful as computer vision before the manipulator doing next procedures to complete its task. This paper provides a systematic summary and analysis target detecting and recognizing object status using Mask R-CNN. Unlike related solution methods that use machine vision and deep learning directly and combine together for doing robot controlling, pushing and grasping, we are doing image processing separately and simply for detecting the object's status before performing like pushing the object for making free and easy grasping. Experiment with this method shows that it has good accuracy, easy to implement for detecting and classification using the algorithm.

Keywords: object detection, Mask RCNN, cluttered environment

1. Introduction

Object detection and classification for robotic applications has emerged as an important goal because object images in cluttered environment provide basic information for several natural image application. object detection should determine the features in image that is contain multi-objectives complex problem considering to classification and localization single or multi-object in image. One of method to approve object detection is using Mask-RCNN. Mask R-CNN is an extension of the Faster R-CNN model for object detection, localization and instance segmentation in natural images.

There are some paper [1], [2], [3] using object classification in cluttered environment combined using grasping strategies to compensate perception error distance from the obstacle or free space around the object, using single perspective [3] and dual perspectives [4] camera for classifying the object in the cluttered environment embedded with reinforcement learning for doing pushing and grasping together with. There are also many research applications in machine learning successful to use Mask RCNN for classifying and identifying the target aims like in smart farming [5], [6] in medical purposes [10], [11], , navigations and remote sensing [8], [9].

This proposed system aims to implement object detection and classification using deep learning Mask RCNN to classify the object status like free or not in the cluttered or uncertainty environment applying in robot manipulation application and then masking the target objects with the boundary mask for determining the borderline mask of the target status. From the propose system we can mark the status of the object ie. complimentary from the other objects or obstacle or not that means the object is close to the other objects or the object over stack to the other so that the object will appear overlapping boundary in images. Output from this propose can be used in manipulator robot to do pick, push or sliding the object based on the object status in advance.

2. Proposed Method

In this paper we introduce an approach called Mask R-CNN which is an extension of Faster RCNN, Faster R-CNN yields two branches class name and bounding box. Mask R-CNN adds extra branch of mask along with the class name and the bounding box

Mask-RCNN [12] was popularized by He et al. in 2018 as a development of Faster RCNN to confess an accurate pixel-based segmentation. This method includes two main steps namely: Feature Pyramid Network (FPN) and Region Proposal Network (RPN). In the feature pyramid network, a various number of proposals was produced about the regions where there might be the background like an object different from the background based on the input image.

In this paper we obtain instance segmentation using Mask R-CNN for masking the object and object status. It consists of (CNN) convolutional backbone, which is a pretrained model like VGGNET,AlexNet, GoogleNet, ResNet etc. Any model we use for masking using RCNN will involve some block like RPN, ROI and some network layers. First, RPN is a Regional Proposal Network which is needed to generate a region of proposal for the given input image. It will produce the feature map of the image. A ROI Align Layer is used for generating fixed size of feature map. A Fully Connected Layer exists for the classification of objects in the image with the bounding box. A Mask Branch is used to generate the mask to the identified object by the fully connected Layer.

In this paper, we used a pre-trained architecture on Coco (80 class) dataset but we customized to 10 classes residue. Generally, the size of the recent model is substantially smaller due to the usage of global average pooling rather than fully-connected layers. We choose ResNet50 as a feature extractor network which encodes input image into 32x32x2048 feature map. The FPN (feature pyramid network) extracts regions of interest from features of different levels according to the size of the feature which feeds as input to Next stage (RPN). In Region Proposal Network (RPN), the regions scanned individually and predicted whether or not an object is present. The actual input image is never scanned by RPN instead RPN network scans the feature map, making it much faster. Next, each of regions of interest proposed by the RPN as inputs and outputs a classification (SoftMax) and a bounding box (regressor). Finally, Mask- RCNN adds a new branch to output a binary mask that indicates whether the given pixel is or not part of an object. This added branch is a Fully Convolutional Network on top of the backbone architecture. The proposed method consists of two main steps: Training and testing steps as illustrated in Figure 1.

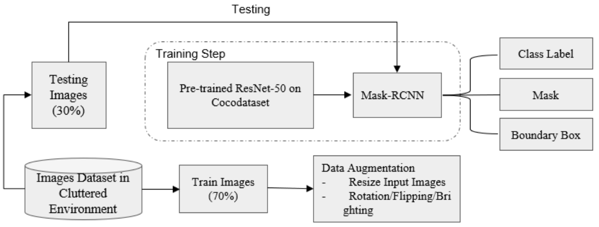


Figure 1. The proposed Status Object detection Architecture

.

The below steps will show the working of the Mask R-CNN

- The input image is given to the Mask R- CNN then the pertained model will give feature map as the output to the RPN.
- The RPN (Regional Proposal Network) will take output of the CNN and gives the multiple regions of interest using light weight binary classifiers.
- The feature Map along with the ROI of the image is given to the ROI Align layer which gives the fixed size feature map as the output, by wrapping the multiple bounded boxes into fixed dimensions.
- Then the fixed size feature map from the ROI Align is given to the fully connected layers and Mask Branch.
- The fully connected Layer will classify the objects in the image with bounded boxes.
- Then the mask branch outputs a binary mask for each ROI of the image.

2.1. Loss Function

Mask R-CNN utilized a multi-task loss function that combined the loss of classification, localization and segmentation mask as illustrated in Eq. (1).

$$L = L_{cls} + L_{bbox} + L_{mask} \quad (1)$$

Where L_{cls} , L_{bbox} are same as in Faster R-CNN [13]. The added mask L_{mask} illustrated in Eq. (2). as the average binary cross-entropy that only includes *k th* mask if the region is associated with the ground truth class k .

$$L_{mask} = -\frac{1}{m^2}\sum_{lsi,j\,sm} y_{ij} \log \hat{y}_{ij}^{k} + (1 - y_{ij})\log(1 - \hat{y}_{ij}^{k}) \quad (2)$$

Where the mask branch generates a mask of dimension m x m for each *RoI* and each class y_{ij} and k, $\hat{y}$, ij are cell *(i, j)* label of the true mask and the predicted value respectively.

2.2. Training Step

Mask-RCNN requires a large number of annotated data for training to cover up the overfitting. To overcome the problem of limited annotated dataset in cluttered environment in robot application, we adopted transfer learning by selected the pre-trained network weights of the resnet50 model, which was successfully trained with the coco dataset. We utilized the pre-trained resnet50 and fine-tuned the network weights to the Coco dataset. we use 30 epochs with learning rate 0.001. We used different augmentation methods such as horizontal flip, vertical flip, image rotation, and image translation to enlarge the training data. One can observe that this domain-specific fine-tuning allows learning good network weights for a high-capacity CNN for coco dataset.

2.3 Testing step

The learned model used directly to predict class label, boundary box, and masked segment for each image in testing data. To evaluate the learned model performance, the predicted labels and boundary box is matched with those in the dataset.

3. Result and Discussion

This section illustrates the results of the target status detection and classification on the cluttered environment. Figure 2 depicts the output of Mask RCNN algorithm in fig. (a) output loss and validation loss recorded in 30 epochs while doing training step, (b) target status detection is free, and not free (c) because close and the distance below the threshold and (d) the object detection overlapping to the other objects. The target status from the picture shows the segmentation of the mask for each object detection and the boundary line for detecting how close the object to the other objects

4. Conclusion

The proposed model will overcome the limitations of the object detection for the status of target in the cluttered environment which is implemented in robot manipulator. In the cluttered environment there are many objects that are close each other, pilled to the other up. the proposed method can determine and classify the status of the target weather free or not. The boundary line along with to each object indicate the limitation and the border threshold for each object. when the target is too close each other, the boundary line of the objects in the mask will cross. If the target is classified free, the manipulator robot can do pick the object from target place and places to the destination place. nonetheless if the status of object is not free the boundary line will show cross and the cross will indicate how close the object to the other and the point which is the manipulator can do next step like pushing or separating the object from the other so that the robot can do grasping in advanced.

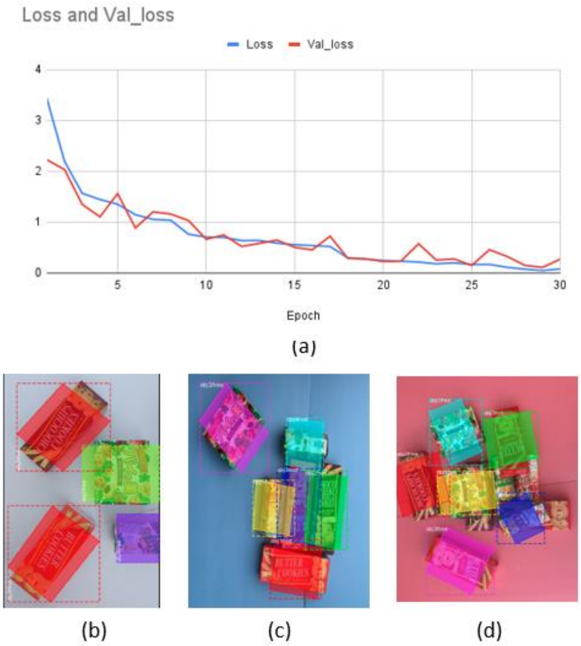


Figure 2. The output of Mask RCNN.
(a). Loss and Validation Loss
(b), (c), (d). Object status detection and Classification.

References

1. I. Sarantopoulos, K. Marios, “Total Singulation with Modular Reinforcement Learning”, IEEE Robotics and Automation Letters, Vol. 6, No. 2, April 2021.
2. I. Sarantopoulos, K. Marios, “Split Deep Q-Learning for Robust Object Singulation|”, IEEE International Conference on Robotics and Automation (ICRA), 31 May – 31 August 2020 Paris France.
3. K. Marios, M Sotiris “Robust Object Grasping in clutter via singulation|”, IEEE International Conference on Robotics and Automation (ICRA), 31 May – 31 August 2020 Paris France.
4. P. Gan, L. Jinhu, G. Shangbin, “A Pushing-Grasping Collaborative Method Based On Deep Q-Network Algorithm in Dual Perspectives”,
5. T. Pallpothu, M. Singh, Riya S, “Cotton leaf disease detection using mask RCNN”, AIP Conference Proceedings, May 2022.
6. T. Pallpothu, M. Singh, Riya S, “Cotton leaf disease detection using mask RCNN”, AIP Conference Proceedings, May 2022.

7. C. Hsien Hsia, T. William Chang, C. Chiang.” Mask R-CNN with New Data Augmentation Features for Smart Detection of Retail Products”, MDPI Journals, March 2022, Volume 12 Issue 6.
8. Y. Gan, S. You, Z. Luo, K. Liu, T. Zhang, “Object Detection in Remote Sensing Images with Mask R-CNN”, Journal of Physics: Conference series, Volume 1673, International Confrecence on Computer Science and Applications 25-27 September 2020, China.
9. K. Zhao, J. Kang, J. Jung, and G.Soh, “Object Detection and Instance Segmentation in Remote Sensing Imagery Based on Precise Mask R-CNN”, In: Proc. of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp.247-251, 2018.
10. T. Padma, C. Usha Kumari, D. Yamini, K. Pravalika, “Image Segmentation Using Mask R-CNN for Tumor Detection from Medical Images”, International Conference on Electronics and Renewable Systems (ICEARS), March 2022.
11. K. Zhao, J. Kang, J. Jung, and G. Soh, “ An Improved Mask R-CNN Model for Multiorgan Segmentation”, In: Proc. Of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 247-251, 2018
12. Kaiming He, Gergia G, Piots D, Ross G, “Mask R-CNN”, Computer Vision and Pattern Recognition, 2017

Authors Introduction

Mr. Kasman

He received his Bachelor and Master Degree from the Department of Electrical Engineering, Moslem University of Indonesia 1998 and Sepuluh Nopember Surabaya. Indonesia 2009, respectively. Currently, he is a student of Doctoral Degree Program in Kyushu Institute of Technology, Japan

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).

Deep-Learning-Based Designed Weight Picking Noodle-like Object

Nattapat Koomklang, Sakmongkon Chumkamon, Prem Gamolped, Tomofumi Tsuji, Eiji Hayashi
Department of Mechanical Information Science and Technology, Kyushu Institute of Technology,
680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan

Abbe Mowshowitz
Department of Computer Science, The City College of New York,
160 Convent Avenue, New York, NY 10031, USA

E-mail: koomklang.nattapat417@mail.kyutech.jp
http://www.kyutech.ac.jp/

Abstract

For food packaging line, manual picking up of the noodle-like objects according to specific weight requires worker's experience for picking quickly and accurately. This article presents a robot arm with a 6-finger gripper picking up the noodle-like objects in specific weight using deep-learning-based to find the best possible action. For measuring the action, we use direct variation to probability of picking action at specific weight in this research use value for the likelihood of weight probability given an action. To find likelihood of weight probably deep-learning-based and use normal distribution for model distribution of the systems. For evaluation we passed any possible action to the network and find action that get maximum likelihood.

Keywords: Deep Convolutional Neural network, Machine Learning,

1. Introduction

Noodle-like objects being picked by robots to a specific weight is challenging for designing a control system. Nowadays, packing noodles into food containers requires experienced workers to pack them up quickly because they can estimate the importance of each pick visually and with experience. This paper proposes a Deep-Learning Neural Network to estimate the robot's grasping action for picking up noodles-like objects from RGB-D camera input with a 6-finguer gripper and cartesian robot movement.

Deep-Learning Neural Network is used in a variety of ways to control robots, especially in picking up things such as dex-net [1] used to predict confidential for grasping object from a parallel-jaw, and vacuum-based suction cup gripper, TossingBot [2] is the network finding the best grasping point for throwing to designed destination. The main concept for the robots grasping network [3], [4], [5], [6], [7] is to find the confidentiality of the state and action of the robot. Commonly confidentiality is represented by the probability of positive action, E.g., successed and failed action. This paper used the likelihood of weight estimation given state (RGB-D image) and robot action to represent the confidence score.

At present, Deep learning has succeeded in the detection, classification, and regression task, especially autoencoder [8], which is the key to representing the latent feature in this research. We use RESNET18 [9] and variational inference [10] for encoding data from RGB-D images. After that, pass encoding data and action of the gripper to fully connected networks for the coefficient of variation in a gaussian distribution.

In this research, the datasets used for training the network are collected from random actions in real environments with fake noodle-like objects with 20,000 data.

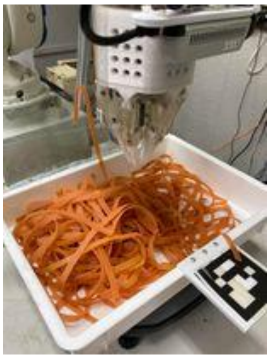


Fig. 1 6-fingers gripper and noodle-like objects

2. Methods

2.1. *Model*

Normally for a measure, the confidence score can represent the probability of success. But collecting datasets from the real environment is hard to collect covering all possible action. In this work we represent to use the probability of weight estimation instead because it can predict with regression model. From *Bayes' theorem* we can assume that the probability of weight estimation is direct variations to the probability of success.

$$P(successed|s, x, t_w) \propto P(w|s, x) \qquad (1)$$

Where s is stage in this research is RGB-D image at grasping point at the center of the image. x is picking action and t_w is target picking weight. w is estimation weight.

This research proposed a deep neural network model that predict the picking weight from RGB-D image and action $x = \{x_{gripper\ width},\ x_{gripper\ depth}\}$ where $x_{gripper\ width}$ is a distance between the end-tip of gripper. And $x_{gripper\ depth}$ is a distance from reference base to the end-tip of gripper in $z - axis$. The model predicting the weight and its covarion. The model shows in Fig. 2

For RGB-D image, we applied RESNET18 and variational inference for estimate latent parameters in gaussian distribution coefficients. After that we use reparameterization trick (In eq.2) for convert gaussian distribution coefficients to latent parameters (z)

$$z = \mu + \rho\epsilon \qquad (2)$$

Finally, we concatenation latent parameters z and action x through 4-layer fully connected neural networks with leaky relu activation function.

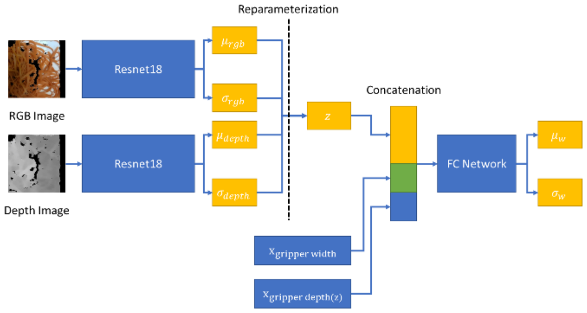


Fig. 2 Weight Estimate Network

2.2. *Loss*

The optimization of the network is the evidence lower bound (ELBO) like in other variational methods as show in eq. Fig. **3**

$$L_{\theta,\emptyset} = \log\big(p_\theta(w|x)\big) - \beta D_{KL}(q_\emptyset(z|x)||p(z|x))(3)$$

The first term, the ELBO term, is a lower bound on the log-likelihood of the data. The second term is the Kullback-Leibler (KL) divergent where β is reduce hyper-parameter.

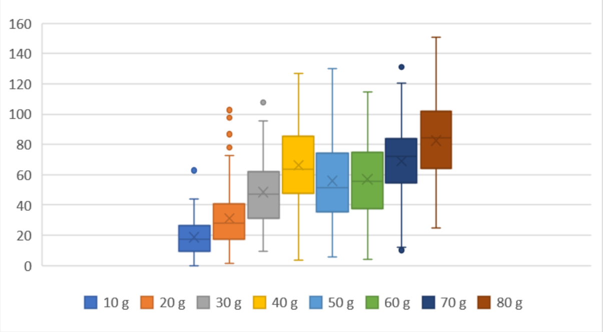


Fig. 3 Picking weight in experimental at 10 – 80 target weight

3. Results

To evaluate the model, we use the network training with 2000 epochs and sampling with possible action to the network between gripper width 30 to 80 mm with step 2.5 mm and gripper depth 0.025 to 0.020 m with step 0.025 m and x, y points every 1 cm grid. After that, we select the action from the maximum likelihood value. The data was collected by setting target weights at 10, 20, 30, 40, 50, 60, 70, 80 g. each target weight accumulated around 100 data points.

Fig. 3 shows the picking weight at each target weight. From the experimental results, it tends to be picked according to the specified weight. But its distribution is relatively high may be due to the sampling for the practical, not enough action point to get the real action.

4. Conclusion

In this paper, we proposed a deep-learning neural network to measure the confidence score for evaluating the action of picking noodle-like with a specified weight. The results of the experiment tend to be picked according to the specified weight. But the distribution is relatively high.

References

[1] J. Mahler, M. Matl, V. Satish, M. Danielczuk, B. DeRose, S. McKinley and K. Goldberg, "Learning ambidextrous robot grasping policies," *Science Robotics,* vol. 4, p. eaau4984, 2019.

[2] A. Zeng, S. Song, J. Lee, A. Rodriguez and T. Funkhouser, *TossingBot: Learning to Throw Arbitrary Objects with Residual Physics,* arXiv, 2019.

[3] H. Cao, G. Chen, Z. Li, J. Lin and A. Knoll, *Lightweight Convolutional Neural Network with Gaussian-based Grasping Representation for Robotic Grasping Detection,* arXiv, 2021.

[4] F.-J. Chu, R. Xu and P. A. Vela, *Real-world Multi-object, Multi-grasp Detection,* arXiv, 2018.

[5] I. Lenz, H. Lee and A. Saxena, *Deep Learning for Detecting Robotic Grasps,* arXiv, 2013.

[6] D. Liu, X. Tao, L. Yuan, Y. Du and M. Cong, "Robotic Objects Detection and Grasping in Clutter Based on Cascaded Deep Convolutional Neural Network," *IEEE Transactions on Instrumentation and Measurement,* vol. 71, pp. 1-10, 2022.

[7] G. Wu, W. Chen, H. Cheng, W. Zuo, D. Zhang and J. You, "Multi-Object Grasping Detection With Hierarchical Feature Fusion," *IEEE Access,* vol. 7, pp. 43884-43894, 2019.

[8] U. Michelucci, *An Introduction to Autoencoders,* arXiv, 2022.

[9] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* 2016.

[10] D. M. Blei, A. Kucukelbir and J. D. McAuliffe, "Variational inference: A review for statisticians," *Journal of the American statistical Association,* vol. 112, p. 859–877, 2017.

[11] D. P. Kingma and M. Welling, "An Introduction to Variational Autoencoders," *Foundations and Trends® in Machine Learning,* vol. 12, p. 307–392, 2019.

Authors Introduction

Nattapat Koomklang

He received master's degree in engineering in 2018, King Mongkut's University of Technology Ladkrabang in Thailand. He is currently a doctor student at Kyushu Institute of Technology and conducts research at Hayashi Laboratory.

Prem Gamolped

He received master's degree in engineering in 2021 from Kyushu Institute of Technology. He is currently a doctor student at Kyushu Institute of Technology and conducts research at Hayashi Laboratory.

Dr. Sakmongkon Chumkamon

Dr. Sakmongkon Chumkamon received Doctor of Engineering degree from Kyushu Institute of Technology in 2017. He was a postdoctoral researcher at Guangdong University of Technology in 2017-2019. Presently he is a postdoctoral researcher in Kyushu Institute of Technology since 2019. His research interests include factory automation robots and social robots.

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).

Prof. Abbe Mowshowitz

Prof. Abbe Mowshowitz received the Ph.D. degree from University of Michigan in 1967. He has been professor of computer science at the City College of New York and member of the doctoral faculty at the Graduate Center of the City University of New York since 1984. His current research interests lie in two areas are organizational and managerial issues in computing, and network science. In addition to teaching and research, He has acted as consultant on the uses and impacts of information technology (especially computer networks) to a wide range of public and private organizations in North America and Europe.

Research on grasping of string foods in the home meal replacement industry

Akihiro Ooya
Department of intelligent and control systems, Kyusyu Institute of technology,
680-4 Kawazu, Iizuka, Fukuoka, 820-8502, Japan

Sakmongkon Chumkamon, Prem Gamolped, Tomofumi Tsuji, Eiji Hayashi
Department of Mechanical Information Science and Technology, Kyushu Institute of Technology,
680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan

Abbe Mowshowitz
Department of Computer Science, The City College of New York,
160 Convent Avenue, New York, NY 10031, USA

E-mail: ooya.akihiro534@mail.kyutech.jp,
http://www.kyutech.ac.jp/

Abstract

In recent years, automation by industrial robots has been desired in Japanese food manufacturing plants. This paper describes the development of an autonomous robot for automating the preparation of home meal replacement. The serving of lunchtime meals includes not only solid foods such as rice balls, but also string foods such as spaghetti. Unlike solidified foods, string foods require quantitative grasping. However, in the grasping experiments of string foods, the spaghetti deteriorates with time, and thus, a problem arises where an accurate quantitative grasping experiment cannot be performed. Therefore, in this study, we perform a quantitative grasping experiment by deep reinforcement learning using a material like string foods to verify the grasping and serving system for string foods.

Keywords: Factory automation robots, Arm robots, Deep reinforcement learning, Fake noodle, ROS

1. Introduction

In recent years, automation by industrial robots has been desired in Japanese food manufacturing plants. In this paper, we develop an autonomous work robot for automating the serving operation of midday meal. In the preparation of home-style meals, there are not only solid foods such as fried bean curd and rice balls, but also string foods such as spaghetti. Unlike solidified foods, string foods require quantitative grasping. However, in the grasping experiment of string foods, the spaghetti deteriorates with time, and thus, a problem arises that prevents an accurate quantitative grasping experiment from being carried out. Therefore, in this study, for the purpose of verifying the grasping and serving system for string foods, quantitative grasping by deep reinforcement learning is performed using materials with similar properties to string foods.

2. System configuration

The configuration of the string object grasping system is shown in Figure 1. In this system, a robot detects and grasps an object based on information from an RGB-D camera. The system consists of a visualization processing unit and a robot control unit. First, the visualization processing unit receives information from the RGB-D camera mounted on the robot, and object detection and center of gravity detection of the object are performed. Next, the robot control unit receives the information from the visualization processing unit, and based on the grasping information learned by deep reinforcement learning, the robot grasps the food by sending motion planning signals to the control unit using ROS.

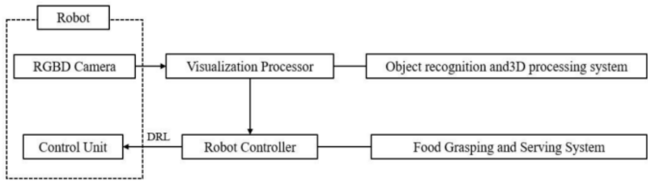

Robot
RGBD Camera
Visualization Processor
Object recognition and3D processing system
Control Unit
DRL
Robot Controller
Food Grasping and Serving System

Fig1. System configuration diagram

3. Robot configuration

Figure 2 shows the appearance of the robot used in this study. Since this robot is designed to perform the same tasks as a human in a food manufacturing factory, it has a vertically articulated 7-axis arm with a high degree of freedom, and is equipped with an RGB-D camera, a force torque sensor, and a tong gripper. The gripper part is an actuator mechanism powered by a stepping motor.

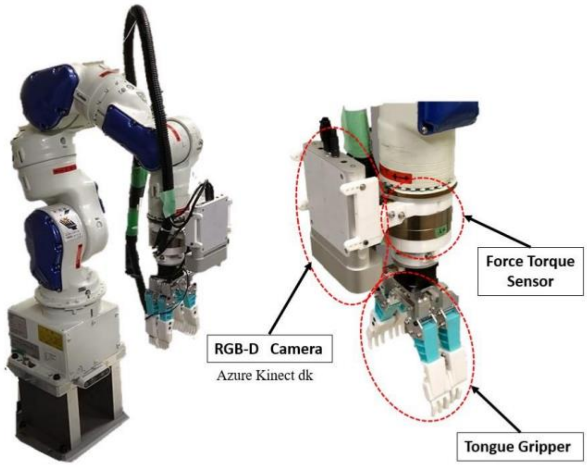

Force Torque Sensor
RGB-D Camera
Azure Kinect dk
Tongue Gripper

Fig2. Appearance of the robot

4. Deep Reinforcement Learning Models

Figure 3 shows an overview of the deep reinforcement learning used in this study.The deep reinforcement learning algorithm uses SAC [1]. The agent consists of SAC and CNN for large-scale image recognition [2]. Grasping is performed by the agent taking action against the environment based on the information from the RGB-D camera. During grasping, a weight measurement is taken, and the closer the value is to the target value, the higher the reward, and the weight and reward data are stored in a replay buffer. The data stored in the replay buffer is retrieved at random and used to train the agents. As the agents are trained, they optimize and update the network. The system learns by repeating this sequence of actions.

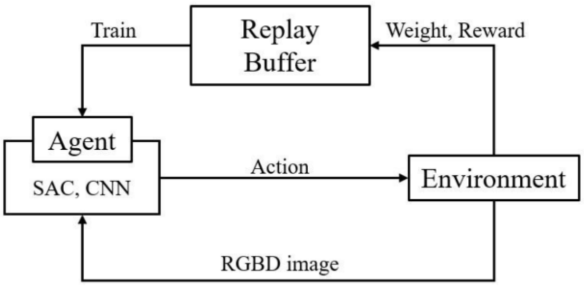

Train
Replay Buffer
Weight, Reward
Agent
SAC, CNN
Action
Environment
RGBD image

Fig3. Overview of deep reinforcement learning

5. Experiment

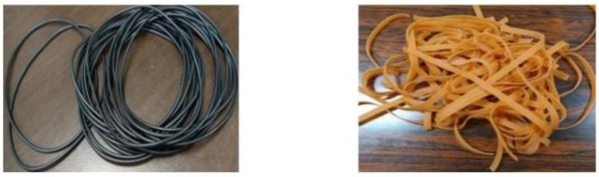


Fig4. (left) o-ring (right) rubber

The o-ring and rubber band shown in Figure 4 were used as materials with similar properties to the string foods, and were trained by a deep reinforcement learning model. In the previous study, the robot grasped spaghetti with a target weight of 50 g. Considering the robot's grasping ability, we set the number of training sessions to 2000, the target weight to 30 g, and the reward to be higher when the grasped weight was within ±10% of the target value. The mean absolute error MAE[g], mean squared error MSE[g^2], standard deviation, maximum error, and minimum error are used as evaluation indices and are shown in Table 2. As a characteristic, MAE evaluates the influence of outliers to a small extent, while MSE evaluates the influence of outliers to a large extent The calculation formulas for MAE and MSE are shown in Equations (1) and (2) as number of data N, correct answer value y_i and predicted value$\hat{y}_i$.

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|\hat{y}_i - y_i| \qquad (1)$$

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2 \qquad (2)$$

5-1. Result

The amount of grasping is shown in Figure 5, and the rewards for learning are shown in Figure 6.
Table 1 shows the experimental results, and Table 2 summarizes the data for several evaluation index.

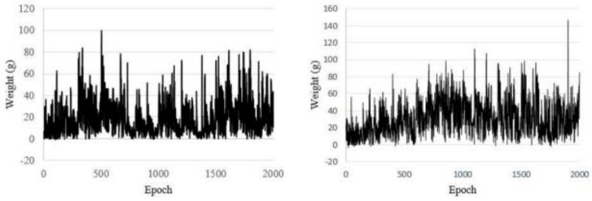


Fig5. Grasping amount （left） o-ring（right） rubber

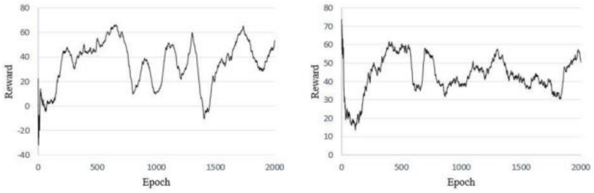


Fig6. Reward （left） o-ring（right） rubber

Table1. Experimental results

	Before learning		
	O-ring	Rubber	Spaghetti
Average grasping amount [g]	9.19	12.35	15.47
Average reward	0.14	28.35	-15.23
	After learning		
Average grasping amount [g]	25.80	24.89	43.70
Average reward	50.30	60.79	53.29

Table2. Evaluation index

	O-ring	Rubber	Spaghetti
Mean absolute error [g]	14.62	11.82	21.66
Mean squared error [g^2]	319.02	216.40	778.85
Standard deviation[g]	17.37	13.80	27.19
Maximum error [g]	58.47	51.65	76.47
Minimum error [g]	0.01	0.62	0.31

5-2. consideration

Table3. Error distribution

	0～10.0	10.1～20.0	20.1～30.0	31.1～40.0	41.1～
O-ring	37	38	20	3	2
Rubber	46	41	11	0	2

Although the rewards in Figure 6 should increase with each learning session and approach 100, they did not increase as expected after 2000 sessions in this experiment. However, Table 1 shows that the average rewards for the o-ring and rubber band are better after learning than before learning, indicating that learning has an effect. Therefore, it is conceivable that the rewards may approach 100 with further learning. Table 2 shows that both the o-ring and the rubber band show less variation than the spaghetti. However, since we were concerned about the size of the maximum error, the distribution is summarized in Table 3.

The maximum error was 5% above 30 g, suggesting that further study is needed.

6. Conclusion

In this study, we verified a system that enables grasping of food strings. As a result, it was confirmed that the grasping accuracy was increased by deep reinforcement learning, but it was found that this study alone was not sufficient to validate the system. In the future, we will search for more suitable products that can be used as food strings and improve the system. We will continue to search for more suitable products that can be used as food strings and to improve the system.

References

1. Tuomas Haarnoja, Aurick Zhou, Pieter Abbeel, Sergey Levine, Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor, 2018.
2. Karen Simonyan, Andrew Zisserman, very deep convolutional networks for large scaleimagerecognition 2014

Authors Introduction

Mr. Akihiro Ooya

He received bachelor degree in Engineering in 2022 from department of Intelligent Systems Engineering, Kyushu Institute of Technology in Japan. He is currently a Master student at Kyushu Institute of Technology and conducts research at Hayashi Laboratory.

Dr. Sakmongkon Chumkamon

Dr. Sakmongkon Chumkamon received Doctor of Engineering degree from Kyushu Institute of Technology in 2017. He was a postdoctoral researcher at Guangdong University of Technology in 2017-2019. Presently he is a postdoctoral researcher in Kyushu Institute of Technology since 2019. His research interests include factory automation robots and social robots.

Prem Gamolped

He received Master degree in Engineering 2021 from Kyushu Institute of Technology in 2022. He is currently a Doctor student at Kyushu Institute of Technology and conducts research at Hayashi Laboratory.

Mr. Tomofumi Tsuji

He received bachelor degree in Engineering in 2021 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. Abbe Mowshowitz

Prof. Abbe Mowshowitz received the Ph.D. degree from University of Michigan in 1967. He has been professor of computer science at the City College of New York and member of the doctoral faculty at the Graduate Center of the City University of New York since 1984. His current research interests lie in two areas are organizational and managerial issues in computing, and network science. In addition to teaching and research, He has acted as consultant on the uses and impacts of information technology (especially computer networks) to a wide range of public and private organizations in North America and Europe.

Prof. Eiji Hayashi i

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 Drifting Debris Detection System using Deep Learning on Coastal Cleanup

Shintaro Ogawa
Department of Intelligent and Control Systems, Kyushu Institute of Technology, 680-4, Kawazu
Iizuka-city, Fukuoka, 820-8502, Japan

Takumi Tomokawa[1], Sylvain Geiser[1], Tan Chi Jie[1], Sakmongkon Chumkamon[1], Ayumu Tominaga[2], Eiji Hayashi[1]
[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-city, Fukuoka, 802-0985, Japan

E-mail: ogawa.shintaro553@mail.kyutech.jp, m-san@mmcs.mse.kyutech.ac.jp
tominaga@kct.ac.jp, haya@mse.kyutech.ac.jp,
http://www.kyutech.ac.jp/

Abstract

In this paper, we developed a litter detector using deep learning to efficiently survey litter on beaches. The litter detector was developed using an HTC network. The HTC network and the mask R-CNN network were compared to evaluate the detector. The results showed that the HTC network was affected by small objects in the images.

Keywords: DNN, HTC, Mask R-CNN, Field Robot, Object detection, TACO

1. Introduction

Beach litter, a type of marine debris, has become a problem. The survey of litter is often conducted manually, which is a time-consuming task. In this study, a litter detection system was developed to reduce the amount of labor required for litter surveys. The creation of a litter detector is important for efficient and accurate litter surveys. The litter can range from small items such as microplastics to large items such as fishing equipment, and there are many beaches that are difficult for humans to access. In this study, we developed a detector that performs instance segmentation of beach images using deep learning. This detector detects trash pixel by pixel and classifies which type of trash it is. The accuracy of the detector was verified using a dataset containing images collected at a beach where cleanup activities were actually conducted.

2. HTC-based debris detection system

In this study, a deep learning model called HTC (Hybrid Task Cascade) was used for litter detection. HTC is an instance segmentation model that combines Cascade R-CNN and Mask R-CNN. Cascade R-CNN is improved from Faster R-CNN and expresses the threshold change of IoU (Intersection over Union, Eq. (1), which indicates the degree of overlap of two regions in a neural network. Mask R-CNN is a model for object detection, Faster R-CNN, with added segmentation functionality. Thus, HTC is created by combining two existing models.

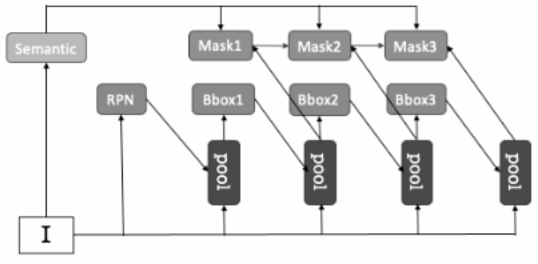


Fig. 1 Model structure of HTC

2.1. *Obtaining the dataset*

The dataset was created by combining a public dataset of general trash called TACO[1] and images taken at actual cleanup sites such as Hokuto Mizukumi Park. TACO is a dataset of overseas trash and includes many images taken at sites other than beaches, with a total of 1,500 images, and the latter includes 186 images taken at actual cleanup sites. The data set consists of 1686 images, 1636 of which are training data and 50 of which are evaluation data, for a total of 60 classes. In this study, in order to compare the detection accuracy between the different datasets, we prepared and trained two datasets, one using only TACO and the other using TACO plus images taken at the actual cleaning sites.

2.2. *Creating of Teacher Data*

In supervised learning, it is necessary to create correct answer data, called teacher data. In supervised learning, it is necessary to create correct data called teacher data. In supervised learning, the process of mapping teacher labels to target objects in images is called annotation. In this study, the annotation tool coco-annotator[2] was used to annotate each class of litter, and the teacher data was created. The image shown in Figure 2 shows the left image before annotation and the right image after annotation, in which a plastic bottle is annotated by dividing it into a bottle part and a cap part.

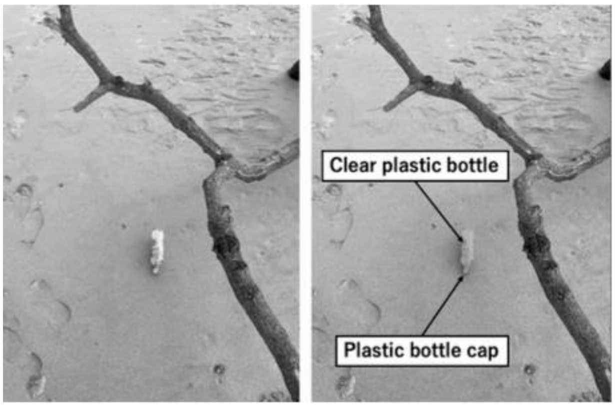


Fig. 2. Example Dataset

2.3. *Identifier creation and evaluation methods*

The discriminator was trained 2000 times with a learning rate of 0.002. Multi-class cross-entropy was used as the loss function. In addition to HTC, Mask R-CNN was used for training to compare real-time performance. The discriminators were evaluated using the unlearned discriminants. For evaluation of the discriminator, 50 untrained images were prepared and evaluated with the two training models. The accuracy was compared by calculating the goodness of fit, recall, and F value in the range of IoU values above 0.5. IoU is a measure of the accuracy of the expected region in object detection, and F value is the harmonic mean of the goodness of fit and recall. Equations (1) to (4) below show the equations for IoU, fit ratio, recall ratio, and F value.

$$IoU = \frac{Area\ of\ Intersection}{Area\ of\ Union} \tag{1}$$

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \tag{2}$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \tag{3}$$

3. Experimental results

Below are the identification results of the evaluation data for each model and dataset as shown in Figure 3.1 and Figure 3.2.

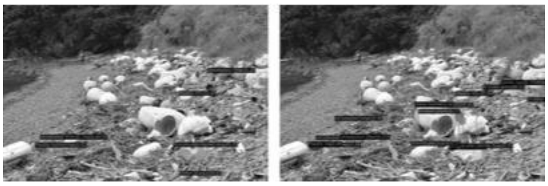


Fig. 3.1. Example of Identification Result
(Left: HTC, Right: Mask R-CNN, Dataset: TACO)

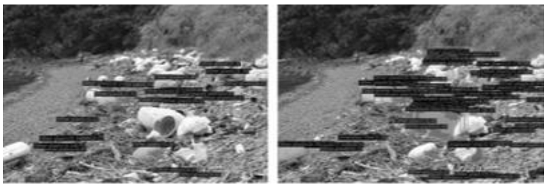


Fig. 3.2. Example of Identification Result
(Left: HTC, Right: Mask R-CNN, Dataset: TACO+Real env.)

The experimental results are evaluated in term of the fit rate, repeatability, and F value for the class as a whole, and the fit rate for PET bottles as shown in Table 1 while

Table 2 shows the individual precision performance for plastic bottle class.

Table 1. Precision, Recall, F-score

	Precision	Recall	F-score
HTC(TACO)	0.034	0.074	0.047
Mask R-CNN(TACO)	0.025	0.03	0.027
HTC(TACO+Real env.)	0.041	0.053	0.046
Mask R-CNN (TACO+Real env.)	0.038	0.053	0.044

Table 2. Precision in plastic bottle

	Precision
HTC(TACO)	0.164
Mask R-CNN(TACO)	0.127
HTC(TACO+Real env.)	0.141
Mask R-CNN (TACO+Real env.)	0.167

4. Consideration

Although Table 1 shows a relatively low precision score, the individual class of plastic bottle demonstrate a better precision result as shown in Table 2. Taking the identification of pet bottles as an example, the Mask R-CNN showed a higher conformance rate with the addition of real environment data, while the HTC showed a higher conformance rate without the addition of real environment data. The F-measure of the class as a whole does not increase much with the addition of the dataset for HTC. Why is it that HTC does not obtain good results even with the addition of a dataset? One possible reason is the compatibility between the added real-world dataset and HTC's model structure. As a characteristic of the data set, the real environment data set contained more litter than TACO, including many small litters. As shown in Figure 1, the model structure of HTC is such that the features of the Region Proposal Network (RPN) are optimized through the pooling layer. Therefore, it is thought that the detection of small debris was omitted in the feature optimization process due to the addition of the real environment data set. Classification of small objects is also extremely difficult, and misclassification is considered to be one of the reasons. For prospects, to solve these problems, the input resolution of the model should be increased, and the number of pixels occupied by small objects should be increased by raising the resolution of the input images.

In addition, it may be possible to prevent misclassification by avoiding annotation of too small garbage and reducing the number of classes to a minimum in the creation of training data.

5. Summary

In this study, we attempted to detect debris at a beach cleaning site using deep learning. We attempted to detect debris at a beach cleanup site using deep learning. As a result, we found that the number of classes for the dataset. The results suggest that it is necessary to set an appropriate number of classes for the dataset and to devise a way to detect small objects and that it is necessary to devise a method for small objects. In the future, we will aim to further improve the detection accuracy. We will further improve the detection accuracy and build a system with the aim of increasing its usefulness in actual cleaning sites. We will construct a system with the aim of further improving the detection accuracy and increasing the usefulness of the system in actual cleaning sites.

References

1. "TACO Dataset ", http://tacodataset.org/
2. Justine Brooks, https://github.com/jsbroks/coco-annotator

Authors Introduction

Mr. Shintaro Ogawa

He received bachelor degree in Engineering in 2022 from intelligent and Control Systems, Kyushu Institute of Technology in Japan. He is currently a Master student at Kyushu Institute of Technology and conducts research at Hayashi Laboratory.

Dr. Sakmongkon

Dr. Sakmongkon Chumkamon received Doctor of Engineering degree from Kyushu Institute of Technology in 2017. He was a postdoctoral Guangdong Technology Presently he is a postdoctoral researcher in Kyushu Institute of Technology since 2019. His research interests include factory automation robots and social robots.

Projected Assist. Prof. Ayumu Tominaga

Projected Assist. Prof. Ayumu Tominaga 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

Mr. Takumi Tomokawa

He received bachelor degree in Engineering in 2021 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.

Mr. Tan Chi Jie

He received his Bachelor of Engineering Electronics Majoring in Robotics and Automation from the Faculty of Engineering, Multimedia University, Malaysia in 2020. He is currently a Master student at Kyushu Institute of Technology and conducts research at Hayashi Laboratory.

Mr. Sylvain Geiser

He studied Engineering and Computer Science in Ecole des Mines de Nancy, France, between September 2018 and June 2020. He is currently a Master student at Kyushu Institute of Technology, Japan, and conducts research at Hayashi Laboratory.

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).

Soft Object Dexterous Manipulation Using Deep Reinforcement Learning

Sornsiri Promma[1], Sakmongkon Chumkamon[1], and Eiji Hayashi[1]
[1]Department of Mechanical Information Science and Technology, Kyushu Institute of Technology,
680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan

Abbe Mowshowitz[2]
[2]Department of Computer Science, The City College of New York,
160 Convent Avenue, New York, NY 10031, USA

E-mail: promma.sornsiri363@mail.kyutech.jp, m-san@mmcs.mse.kyutech.ac.jp,
haya@mse.kyutech.ac.jp, amowshowitz@ccny.cuny.edu
http://www.kyutech.ac.jp/

Abstract

Manipulation of objects is one of the basic tasks that has been studied for a long time in the robotic field. Many experiments were setting the environment and using Deep Reinforcement Learning to train robot arm to grasp various objects. Still, most of those objects are solid objects, whereas nowadays, robot arms are used for grasping soft objects as well. In this study, we develop object manipulation tasks in pybullet simulation, focusing on soft objects by using Deep Reinforcement Learning (DRL) based on Soft Actor-Critic and Proximal Policy Optimization algorithm which aims to make the robot able to grasp soft objects in the exact position with proper force that does not damage them.

Keywords: Object manipulation, Soft Object, Deep Reinforcement Learning (DRL), Soft Actor-Critic, Proximal Policy Optimization

1. Introduction

Manipulation of objects is one of the basic tasks that has been studied for a long time in the robotic field. For humans, it is a simple movement that we can do immediately without thinking, whereas, it is quite challenging for the robot to grasp the objects without making them fall. To teach the robot, one of the effective ways that have been widely used nowadays is Deep Reinforcement Learning, using a robot arm to try to grasp objects repeatedly.

Many papers have succeeded in using robot arms to grasp different shapes of objects, focusing on solid ones. On the contrary, soft-body objects are rarely studied and are more challenging. The main factor when grasping solid objects mostly depended on precise position only, whereas, grasping the solid object needs to consider proper force, elasticity, and the shape both before and after touching them.

So, in this study, we develop object manipulation tasks in pybullet simulation, focusing on soft objects by using Deep Reinforcement Learning (DRL) based on Soft Actor-Critic and Proximal Policy Optimization algorithm which aims to make the robot able to grasp soft objects in the exact position with proper force that does not damage them.

2. System Overview

2.1. *Simulation platform*

The simulation platform that we chose for this research is pybullet. Pybullet is an open-source physics simulation that is widely used in the robotic field. It is easy to use and integrated with Reinforcement Learning such as Stable Baseline which is also implemented on pybullet. Moreover, not only is it able to import various formats of mesh and articulated bodies, it can import those as soft objects with high accuracy of collision as well.

2.2. *Environment Configuration*

In pybullet, we are setting the environment that will be used for training Deep Reinforcement Learning, using Open AI gym, a python API that is used for developing Reinforcement Learning agents. In this paper, for the agent, the robot arm that we are using is Xarm7 [1], a 7-axis robotic arm with a pair of grippers, implemented in pybullet, as we can see in Fig 1. Its movement will be calculated as inverse kinematics that receive the position from the DRL model. For soft object, we used mesh in VTK format from bullet3 library which is in a tube shape, so that the agent needs to grasp it in the proper position. After importing, soft object needs to be adjusted some parameters to make it has suitable properties such as not too soft or bounce too high.

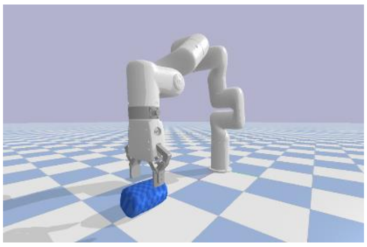


Fig. 1. Xarm7 with a pair of grippers and soft object in tube shape that is used for training in pybullet

2.3. *Deep Reinforcement Learning Model*

As described, to teach the robot to grasp soft object, we used Deep Reinforcement Learning (DRL), same as many papers that do object manipulation tasks with solid objects. DRL is not required to collect any data which is suitable for the task in the robotic field. Additionally, it can learn complex behaviors, taking action repeatedly based on observation to get the reward and evaluate them.

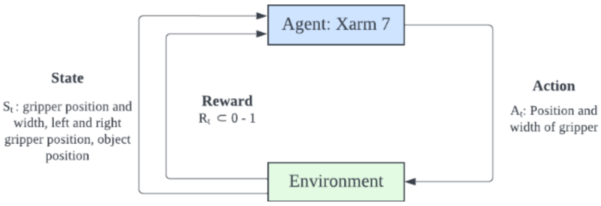

State
S_t : gripper position and width, left and right gripper position, object position
Agent: Xarm 7
Reward
$R_t \subset 0 - 1$
Environment
Action
A_t: Position and width of gripper

Fig. 2. Process of training Xarm7 to grasp soft object using Deep Reinforcement Learning

According to the Fig 2, in our experiment, the agent is Xarm7. The agent will take the action received from the model, which consists of the next position (x, y, z) and gripper width. After it performed, it will observe by get the value gripper position and width, left and right gripper position, plus gripper position. At the same time, it will calculate rewards as the result of the action, which detail will be described in the next section. Lastly, state and reward will be calculated together, so that model can determine the next step for the agent. This one-time process will be calculated as one step, and it will do repeatedly until the episode ends.

To make the robot arm knows its action result, we need to define a proper reward function. Shaped reward function is used in order to evaluate each episode, which is calculated from the equation below.

$$R_{total} = R_{distObjGoal} + R_{distObjGripper} + R_{gripper} \tag{1}$$

From the Eq (1), $R_{distObjGoal}$ is calculated from distance between object and goal in z-axis only. The higher it is, the more reward it gets. The maximum reward for $R_{distObjGoal}$ is when gripper grasps the object to 1 on z-axis. Next, $R_{distObjGripper}$ is calculated from distance between object and gripper in 3-axis (x, y, z) Agent will get more reward if the gripper is near the object. For $R_{gripper}$, this reward depends on the gripper state if it opens or closes in the right position. For example, when it is not in the grasping position, it will get higher reward if the gripper stays open. In contrast, if it is in a position of being able to grasp the object, it will get reward if it closes and touches the object. Finally, we normalize R_{total} to be in the range of 0 to 1.

The Deep Reinforcement Learning algorithms that we used for training the agent are originally from Stable Baseline3 [2], which are PPO and SAC. We chose these two algorithms based on their types, which are on-policy and off-policy respectively so that we can compare the training result. Firstly, PPO, Proximal Policy Optimization algorithm is an on-policy model-free DRL algorithm that uses a policy gradient approach to optimize a policy by avoiding policy to not change too drastically from one iteration to the next which will make training more stable. To prevent this, it uses clipped objective functions to keep the difference between the current policy and the old one [3]. On-policy DRL algorithm concept such as PPO can be seen in Fig 3.

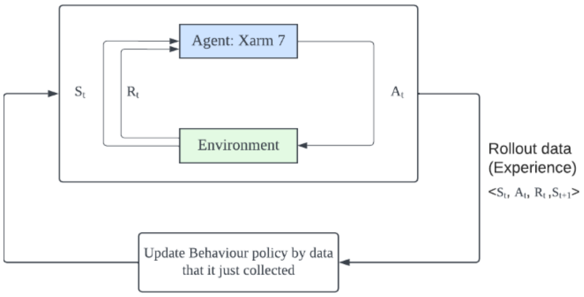

Agent: Xarm 7
S_t
R_t
A_t
Environment
Rollout data (Experience)
<S_t, A_t, R_t ,S_{t+1}>
Update Behaviour policy by data that it just collected

Fig. 3. Illustration of on-policy algorithm process such as PPO

Second, SAC or Soft Actor-Critic, is an off-policy model-free DRL algorithm which overall concept can be seen in Fig 4. Normally, many DRL algorithms will focus on maximizing reward function only, while SAC focuses on maximizing the entropy term as well[4]. It is the algorithm that will trade-off between explore and exploit which is suitable for the real-world task, plus, it can work on continuous action space. SAC can also reuse data that has been collected from other tasks, just adjusting some parameters and reward functions for the new task will make it able to learn faster.

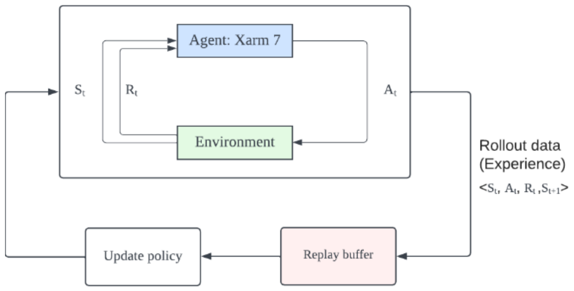


Fig. 4. Illustration of off-policy algorithm process such as SAC

3. Experiment and Result

We trained the agent 2 times using the process as mentioned in the previous section, the first time training with the PPO algorithm and the second time training with SAC algorithm. Moreover, we also presented other 2 results in the case of changing the observation space into the image and the result of using the grasping position to be the starting position.

3.1. *Training with PPO algorithm*

Firstly, agent, Xarm7, will take the action which is position and gripper width. The position will be normalized to be in the range [-1, 1] before the agent does the action. After that, it will observe the environment by using the gripper position, gripper width, left and right gripper position, plus the current object position. Only gripper width will be normalized to be in the range [0, 1] All of the observations will be used to calculate the reward using shaped reward function that has been described in above section. After that, it will update the policy before doing the next episode. Each episode contains 3,000 timesteps. Termination can be decided by 3 factors

- Reach 3,000 timesteps (reward -100)
- Object reach the goal (reward +1,000)
- Soft body crash (reward -100)

For the last factor, sometimes, soft object can crash by the gripper putting too much force on it which makes pybullet not able to calculate its physics. To test the environment and model, we use PPO as the DRL algorithm to test the result.

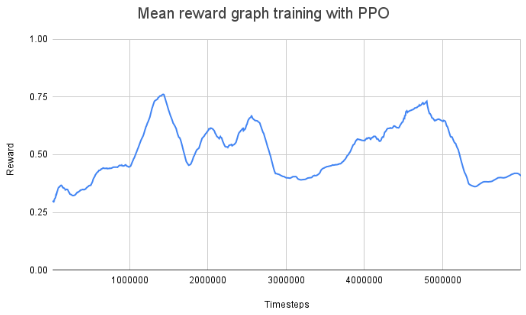


Fig. 5. Mean reward per timesteps graph, training with PPO

From Fig 5, we can see that the environment is basically working. Mean reward gradually increases while as decreases at some point such as when it reached 1.5 million timesteps or 2.5 million timesteps. After we have tested and observed, we noticed some interesting results.

- If it is able to grasp the object, the mean reward will be around 0.7
- From the testing result, it can stay at the grasping position perfectly when it reached 2.5 million timesteps.
- Training object manipulation task for soft object is slower than training with solid object around 35%

Still, even if the agent can reach the grasping position, it does not grasp the object up to the given position. We also continue further around 2 million timesteps but the result is not better.

3.2. *Training with SAC algorithm*

Training with SAC process is different from PPO in some parts, in order to make it faster for training.

- Move the starting position closer to the object.
- Change timesteps per episode, from 3,000 timesteps to 1,000 timesteps.
- Reduce terminal reward due to fewer timesteps per episode.
- Adjust clipping reward range due to less distance of starting position.

With nearer distance, agent training with SAC can reach grasping position in 300,000 timesteps, with mean reward of around 0.35. After that, as we can see in Fig 6, the mean reward sharply reduces after training around 1.3 million timesteps. It tries to grasp and move the object to various positions. Still, sometimes it is moving too much which damage the soft object and gets much minus termination reward, leading to the gripper trying to avoid the grasping position and move to the opposite side. So, there is a possibility that SAC exploitability is not suitable for object manipulation task for soft object.

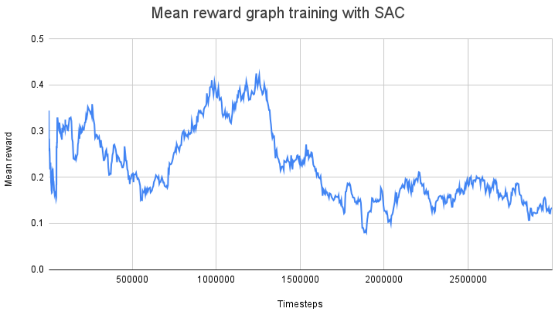


Fig. 6. Mean reward per timesteps graph, training with SAC

3.3. *Training by using image as the observation, with SAC algorithm*

Not only training with normal observation like using the current position of the gripper or object, but we also use a gym wrapper to change the observation space into the array of images. We have tested this environment with SAC algorithm and changed the policy from MlpPolicy to CnnPolicy to extract features from the image. This agent has been trained after changing some parts as mentioned in the previous section.

The result, as in Fig 7, is not different from training with the normal one. Still, one noticeable point is that using image observation for training takes 2 times longer than training with normal observation space.

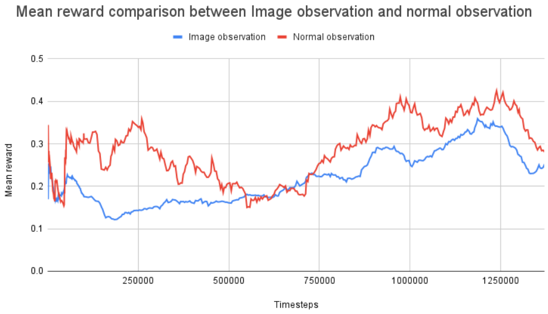


Fig. 7. Mean reward per timesteps comparison graph between image observation and normal observation

3.4. *Training with PPO algorithm by adjusting the starting position*

Lastly, to ensure that it is able to grasp the soft object by mainly considering only the gripper width, we trained the agent by changing the starting position, as we can see from Fig 8, the gripper stays at the grasping position from the beginning. We trained for 400,000 timesteps with PPO algorithm, using the same configuration with SAC.

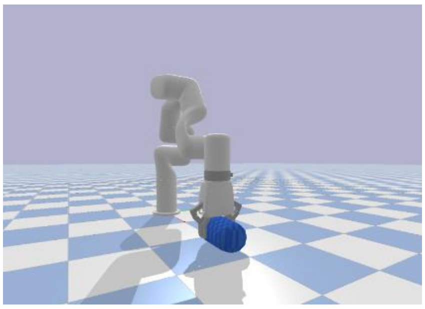


Fig. 8. New starting position

The result, according to Fig 9, is the agent can grasp the object and hold it in the air when it reached only 50,000 timesteps. It adjusts the gripper width correctly to hold the object, not too far to make the object fall and not too close to damage the soft object.

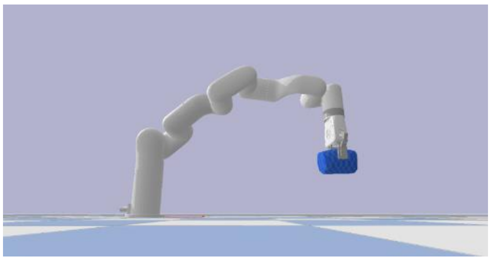


Fig. 9. Agent can grasp soft object into the air after training around 50,000 timesteps

4. Conclusion

In this research, we developed object manipulation task by focusing on soft object. The results show that the agent can hold soft object without damaging it if the gripper is in the correct position. Still, if the starting position is far from the object, it needs some period of time for training which is longer than training with solid object due to soft object physic calculation. Furthermore, using SAC algorithm for soft object manipulation task, the result is not good enough which the reason can be that its exploitability leads to action that harm soft object or the model needs to tune more hyperparameter to get the better result.

References

1. Pan, C. (2022). Gym implementation for Xarm. [online] GitHub. Available at: https://github.com/jc-bao/gym-xarm [Accessed 15 Dec. 2022].
2. stable-baselines3.readthedocs.io. (n.d.). Stable-Baselines3 Docs - Reliable Reinforcement Learning Implementations — Stable Baselines3 1.7.0a5 documentation. [online] Available at: https://stable-baselines3.readthedocs.io/en/master [Accessed 26 May. 2022].
3. huggingface.co. (n.d.). *Proximal Policy Optimization (PPO)*. [online] Available at: https://huggingface.co/blog/deep-rl-ppo [Accessed 1 Dec. 2022].
4. Yazici, B. (2021). Sample Efficient Robot Training on Pybullet Simulation with SAC Algorithm. [online] Medium. Available at: https://towardsdatascience.com/sample-efficient-robot-training-on-pybullet-simulation-with-sac-algorithm-71d5d1d4587f [Accessed 5 Dec. 2022].

Authors Introduction

Sornsiri Promma

She received bachelor degree in 2022 from Computer Engineering, King Mongkut's University of Technology Thonburi, Thailand. She is currently a Master student at Kyushu Institute of Technology and conducts research at Hayashi Laboratory.

Dr. Sakmongkon Chumkamon

Dr. Sakmongkon Chumkamon received Doctor of Engineering degree from Kyushu Institute of Technology in 2017. He was a postdoctoral researcher at Guangdong University of Technology in 2017-2019. Presently he is a postdoctoral researcher in Kyushu Institute of Technology since 2019. His research interests include factory automation robots and social robots.

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).

Prof. Abbe Mowshowitz

Prof. Abbe Mowshowitz received the Ph.D. degree from University of Michigan in 1967. He has been professor of computer science at the City College of New York and member of the doctoral faculty at the Graduate Center of the City University of New York since 1984. His current research interests lie in two areas are organizational and managerial issues in computing, and network science. In addition to teaching and research, He has acted as consultant on the uses and impacts of information technology (especially computer networks) to a wide range of public and private organizations in North America and Europe.

Practical Implementation of FastSLAM for Forestry Robot

Sylvain Geiser[*]
Department of Interdisciplinary Informatics, Graduate School of Computer Science and Systems Engineering,
Kyushu Institute of Technology
680-4, Kawazu, Iizuka-City, Fukuoka, 820-8502, Japan

Sakmongkon Chumkamon
Department of Mechanical Information Science and Technology, Kyushu Institute of Technology
680-4, Kawazu, Iizuka-City, Fukuoka, 820-8502, Japan

Ayumu Tominaga
National Institute of Technology, Kitakyushu College
5-20-1 Shii, Kokuraminamiku, Kitakyushu, Fukuoka, 802-0985 Japan

Takumi Tomokawa
Department of Mechanical Information Science and Technology, Kyushu Institute of Technology
680-4, Kawazu, Iizuka-City, Fukuoka, 820-8502, Japan

Tan Chi Jie
Department of Mechanical Information Science and Technology, Kyushu Institute of Technology
680-4, Kawazu, Iizuka-City, Fukuoka, 820-8502, Japan

Eiji Hayashi
Department of Mechanical Information Science and Technology, Kyushu Institute of Technology
680-4, Kawazu, Iizuka-City, Fukuoka, 820-8502, Japan
E-mail: geiser.nathan-sylvain778@mail.kyutech.jp, m-san@mmcs.mse.kyutech.ac.jp, tominaga@kct.ac.jp,
tomokawa.takumi163@mail.kyutech.jp, tan.jie-chi339@mail.kyutech.jp, haya@.mse.kyutech.ac.jp,
http://www.kyutech.ac.jp/

Abstract

As the Japanese forestry workforce is shrinking, field robots are gaining interest in performing dangerous tasks. This paper presents research conducted on the SOMA robot designed at Hayashi Laboratory. It focuses on issues encountered through the implementation of FastSLAM algorithm on this robot. In particular, the determination of the positions of trees from the raw pointcloud of the lidar, side effects occurring at the boundary of the lidar visibility scope, and the modelling of motion and observation noises are discussed.

Keywords: Field Robot, online SLAM, Forestry, Particle Filter

1. Introduction

In Japan, the decline of the number of workers in the forestry sector urges to use robots, especially for perilous tasks. The SOMA robot is a prototype developed at Hayashi Laboratory in order to address this need[1].

Since the ability of Simultaneous Localization and Mapping (SLAM) is essential for a mobile robot,

---

[*] 7 rue Peyron, Vienne (38200), France

FastSLAM is under implementation on it[2]. This online algorithm is based on a particle filter whose particles represent joint hypothesis of the pose of the robot and the feature map of the environment. Employing this solver in the forestry context requires many adjustments and this article shed light on the encountered issues and the remedies which help fixing them.

The first part of this article deals with the process needed to extract the coordinates of the trees relative to the robot from the raw pointcloud of the lidar. Then, the side effects occurring at the boundary of the visibility scope of the sensor are described and a solution is proposed. Thirdly, the characterization of motion and observation noises through statistical analysis of simulation data is detailed. Finally, a conclusion gives an insight of the remaining topics to be investigated.

2. Observation of trees

Since the SOMA robot evolves in the forest, trees fulfill the role of landmarks. The feature maps of the particles are made of a collection of the positions of these easily distinguishable objects.

In order to get their coordinates, trees must be found in the pointclouds generated by the lidar mounted on the robot. This first part of the process is performed by clipping the pointcloud in height to remove the ground and the canopy and then using euclidian clustering (Fig. 1). Once trunks are uniquely identified, a point has to be defined for each of them, involving polar coordinates in the local coordinate frame of the robot.

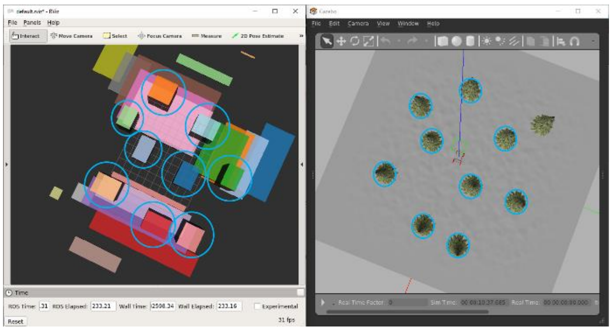


Fig. 1. Clusters made from lidar pointcloud (RViz on the left, Gazebo on the right)

2.1. *Centroids of clusters*

At first, the centroids of the clusters were chosen as the positions of the trees. However, a tree is never entirely captured by the lidar. In fact, only half of its perimeter is represented at the most. This leads to a mismatch between the geometric center of the points and the center of the tree projected on the plane. In addition, due to its dependence on the viewed part of the tree, the centroid varies with the pose of the robot relative to it[3]. This breaks the fundamental static world assumption, on which FastSLAM rests.

2.2. *Cylinder recognition*

In order to improve the accuracy and ensure the consistency of the positions of the trees, the use of pattern recognition has been proposed. The RANSAC algorithm is applied to fit a cylinder to each cluster and take the intersection of the axis of the cylinder and the plane to define the coordinates of the tree. All clusters are processed simultaneously in parallel, the synchronization being handled by the *message_filters* package of ROS.

2.3. *Comparison*

The use of cylinder recognition meets expectations, as it can be seen on Figure 2. Its addition to the poincloud processing leads to much lower final errors, divided by ten for the pose and by more than three for the map.

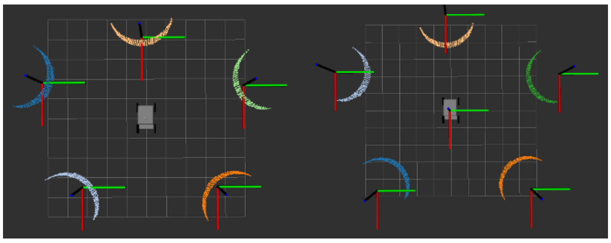


Fig. 2. Positions of trees found by using centroids of clusters (left) and cylinder recognition (right)

3. Side effects of visibility scope

Like every sensor, lidar has a maximum range of operation, which can be roughly defined by a circle of a certain radius around it. This boundary which separates the visible and the invisible is subject to side effects during data association.

Indeed, because of observation noise, the landmark in the map of a particle corresponding to the tree seen by the real robot can be out of the visibility scope but very close to the boundary. On the contrary, the real tree associated to a visible landmark in the map of a particle can be just beyond the limit (Fig. 3). If no action is taken, in the first case, a new landmark is created in the map, and in the second one, the visible landmark is deleted from the map.

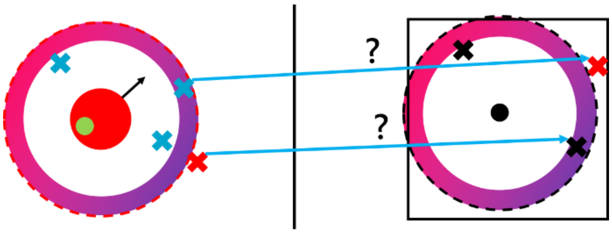


Fig. 3. Side effect occurring at the boundary of the visibility scope of the lidar (real robot on the left, particle on the right)

3.1. *Loose boundary*

One remedy to this issue is to define an area around the visibility boundary in which real trees and map landmarks are not taken into account. Using this conservative approach prevents from improper creations and deletions of features in maps, at the expense of an under-use of observation data.

This airlock area is located between the maximum range circle and a smaller one based on the correspondence threshold. The former is defined such that no tree closer to the robot could be associated with an invisible landmark, and vice versa. The radius of this circle is different for each observation and landmark, because its definition requires a distance to the robot pose to be compared to the maximum range.

More precisely, the procedure applied through multiple data association is the following. Each individual observation situated in the loose boundary or possibly associated with a landmark located in this area is thrown away. Then, the Gale-Shapley algorithm is executed to associate real trees to features in the map[3]. During this step, only visible features outside the airlock zone are considered. Finally, if a tree is not represented, a new landmark is created, and features which are orphan and could not correspond to unprocessed observations are deleted.

This procedure allows the final average map error to be divided by two.

4. Motion and observation modeling

In order to implement the FastSLAM algorithm, motion and observation noises have to be modelled. The movement of the robot is decomposed into three parts: a first rotation, a translation and a second rotation. Regarding observation, it consists of distance and azimuth between the viewed tree and the robot. The noisy values taken by these five variables are assumed to follow a gaussian distribution whose standard deviation is a linear function (Eq. 1).

$$\delta rot1 \sim \mathcal{N}\left(\underbrace{\overline{\delta rot1}}_{\boldsymbol{\mu_{rot1}}}, \underbrace{a_{rot} \cdot \left|\overline{\delta trans}\right| + b_{rot} \cdot \left|\overline{\delta rot1}\right| + c_{rot}}_{\boldsymbol{\sigma_{rot1}}}\right) \quad (1)$$

The coefficients of these linear functions need to be adjusted to explain experimental data. As a first step, simulation has been used to characterize them.

4.1. *Statistical analysis*

Since the same procedure has been applied for each variable, its description will only be made for the first rotation of motion.

First of all, the true value of all the parts of motion and the noisy value of first rotation are recorded during a simulation in which the robot is moved in a diversified manner. The time interval used for estimating motion has to be carefully chosen, otherwise the final results would not be usable. As a general rule, it needs to be approximately equal to the time interval between two executions of the FastSLAM algorithm. Similar movements are then grouped to compute the associated variance of the gaussian distribution they are drawn from. This is done for each group by using maximum likelihood estimation. Then, after having removed abnormal values, non-linear least squares are applied on these new points to find the best hyperplane minimizing residuals. The results are graphically represented by Figure 4.

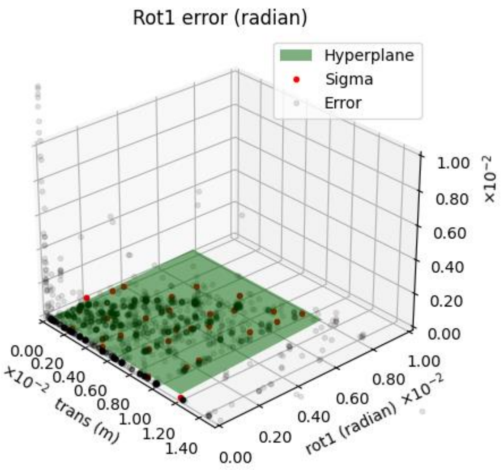


Fig. 4. Non-negative least squares estimation of the parameters of first rotation noise

Globally, the values of the parameters of motion and observation noises derived from simulation data are close to manually tuned ones. They are gathered in Table 1 and Table 2.

Table 1. Estimation of the parameters of motion noise

Errors / Factors	Translation	Rotation	Constant
First rotation	0.0	0.06	0.0002
Translation	0.02	0.006	0.00003
Second rotation	0.0	0.06	0.0003

Table 2. Estimation of the parameters of observation noise

Errors / Factors	Distance	Constant
Distance	0.005	0.0
Azimuth	0.0005	0.0

5. Conclusion

5.1. *Achievements*

Implementing FastSLAM on a real platform arouses numerous issues from which results can suffer.

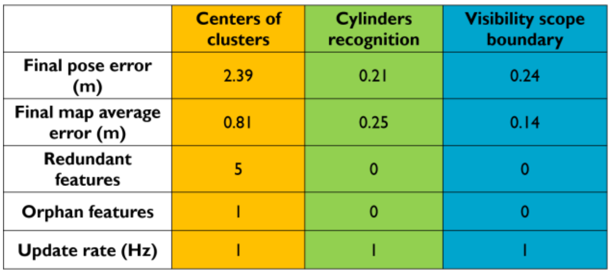
At the beginning of this article, the localization of trees, essential for the whole algorithm, has been addressed. Centroids of clusters identifying trees have proved to be inconsistent and inaccurate estimates. However, applying RANSAC on these clusters to fit cylinders to them generates satisfactory locations for landmarks.

A sharp boundary for the lidar visibility scope causes side effects at the source of wrong maps with orphan and twin features. In order to limit inopportune creations and deletions of landmarks, an airlock area relative to each observation and feature can be defined according to the correspondence threshold. All trees within this loose boundary are not considered for data association anymore, leading to a wasteful but safe method.

Finally, motion and observation models need to be tuned to reflect the real behavior of actuators and sensors. Instead of manual tuning, which can cause overestimation or underestimation of noise and thus distorts performance, advantage can be taken of simulation or experimental data by using statistical analysis. Non-linear least squares method has been proposed as the core of this approach to find optimized noise parameters for motion and observation.

Providing an answer to these problematics has improved stability and accuracy of pose and map estimates, as it can be seen in Table 3.

Table 3. Quantitative results of successive implementations of FastSLAM



	Centers of clusters	**Cylinders recognition**	**Visibility scope boundary**
Final pose error (m)	2.39	0.21	0.24
Final map average error (m)	0.81	0.25	0.14
Redundant features	5	0	0
Orphan features	1	0	0
Update rate (Hz)	1	1	1

5.2. *Future research*

There are other issues which still need to be handled. Among others, the excellent accuracy of the lidar compared to wheel odometry sometimes leads to divergence of FastSLAM, because particles are generated based on motion only. This is possible that no particles lie within the vicinity of real pose at some point, making observation ineffective at promoting best particles. This drawback can be mitigated by using a mixture approach, where particles are generated by either motion or observation in a fixed or variable ratio.

References

1. N. Takegami, E. Hayashi, R. Fujisawa, "Environment map generation in forest using field robot", Proceedings of International Symposium on Applied Science 2019
2. M. Montemerlo, S. Thrun, D. Koller, B. Wegbreit, "FastSLAM: A Factored Solution to the SLAM Problem", 2002
3. S. Geiser, S. Chumkamon, A. Tominaga, T. Tomokawa, E. Hayashi, "Online SLAM for Forestry Robot", Journal of Robotics, Networking and Artificial Life Vol. 9 (2), September 2022, pp. 177–182

Authors Introduction

Mr. Sylvain Geiser

He studied Engineering and Computer Science in Ecole des Mines de Nancy, France, between September 2018 and June 2020. He is currently a Master student at Kyushu Institute of Technology, Japan, and conducts research at Hayashi Laboratory.

Dr. Sakmongkon Chumkamon

Dr. Sakmongkon Chumkamon received Doctor of Engineering degree from Kyushu Institute of Technology in 2017. He was a postdoctoral researcher at Guangdong University of Technology in 2017-2019. Presently he is a postdoctoral researcher in Kyushu Institute of Technology since 2019. His research interests include factory automation robots and social robots.

Projected Assist. Prof. Ayumu Tominaga

Projected Assist. Prof. Ayumu Tominaga 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.

Mr. Takumi Tomokawa

He received bachelor degree in Engineering in 2021 from mechanical system engineering, Kyushu Institute of Technology in Japan. He is aquiring the Master degree in Kyushu Institute of Technology.

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).

Research on AR system for industrial robot introduction

Takuya Matsumoto
Department of Intelligent and Control Systems, Kyushu Institute of Technology, 680-4, Kawazu
Iizuka-city, Fukuoka, 820-8502, Japan

Eiji Hayashi[1],Sakmongkon Chumkamon[1], Ayumu Tominaga[2], Abbe Mowshowitz[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-city, Fukuoka, 802-0985, Japan
[3]Department of Computer Science, The City College of New York,
160 Convent Avenue, New York, NY 10031, USA

Group, Laboratory, Address
City, State ZIP/Zone, Country
E-mail: matsumoto.takuya130@mail.kyutech.jp
http://www.kyutech.ac.jp/

Abstract

In recent years, labor shortages in small and medium-sized enterprises (SMEs) have become a serious issue, and the demand for automation by robots is increasing. Another challenge is the high cost of introducing industrial robots. In order to reduce the introduction cost, we are developing an AR system with the aim of providing robot introduction support to SMEs. The AR system developed enables the display of the robot's movements on the AR screen of a smartphone by communicating with ROS while developing Google's ARCore in Unity. In this paper, we describe the system configuration and evaluate the application of an AR application that enables AR simulation and intuitive GUI operation to check the safety range at the time of robot introduction.

Keywords: AR, Factory Automation Robots, ARCore, Unity, ROS

1. Introduction

In recent years, labor shortages in small and medium-sized enterprises (SMEs) have become an increasingly serious issue. In the food service industry in particular, automation using robots is expected to improve productivity in order to solve labor shortages. However, the introduction of robots is costly in terms of design, parts, and other costs. For small and medium-sized enterprises (SMEs) that do not have abundant funds, the high cost is a major barrier to the introduction of robots. This study utilizes Augmented Reality (AR) technology, which displays virtual information that does not exist in real space on a display via a camera such as a smartphone or AR glasses. In other words, a virtual robot model is operated in AR to check the actual food preparation and safety range in a factory.

This eliminates the need to set up an actual robot in a factory and check its operation, and is expected to significantly reduce the cost of introduction, including design and system integration costs.

In this study, we developed an AR application for small and medium-sized enterprises (SMEs) to check the safety range of robots when they are introduced. In addition, we assumed that the users of this application would be people who are not familiar with robot development. Based on these assumptions, we developed an application that is easier to operate and more user-friendly, and we describe its evaluation and verification.

2. Software Overview

2.1. *AR System Configuration*

Three main technologies were used to develop the AR application. We used ARCore SDK for Unity to use ARCore, a software for AR development, on Unity for application development. Fig. 1 shows a schematic diagram of the AR system.

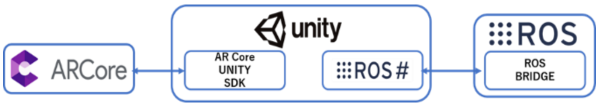


Fig. 1. AR System Configuration

2.2. *AR Application*

The application developed in this study is to operate a real robot after AR simulation. The flow of the application is shown in the following order (a), (b), (c), and (d) in Fig. 2.

(a) Object selection screen

Select the type of solidified food to be grasped, the number of pieces, and the location of the container to be used for serving.

(b) AR simulation screen

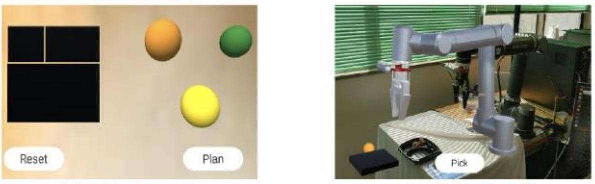
The coordinates of the object to be grasped are acquired from the RGB-D camera mounted on the tip of the actual robot, and projected onto the AR.

(c) Real robot start screen

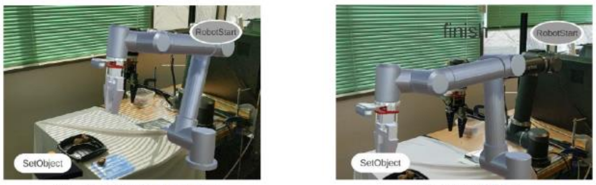
After the simulation is completed, a button for the actual robot operation appears. After pressing the button, the actual robot performs grasping and serving.

(d) End screen

After the actual robot completes grasping and serving, a text informing the end of the operation is displayed on the AR.



(a) Screen of select object (b) Screen of AR simulation



(c) Screen of real robot starting (d) Screen of finish

Fig. 2. Screen of AR application

3. Robot System

3.1. *Robot Configuration*

The appearance of the robot used in this study is shown in Fig. 3.1 and Fig. 3.2. Since this robot is required to perform the same tasks as a human in a food factory, a 6-axis vertically articulated robot with a shape similar to a human arm and a high degree of freedom of movement was used. An RGB-D camera for object recognition and a gripper for grasping are mounted on the tip of the robot. This gripper can be opened and closed by stepper motor control to realize the grasping of solidified food.

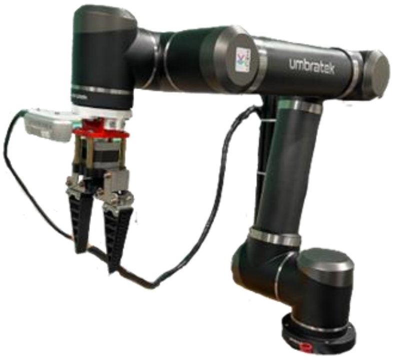


Fig. 3.1. Appearance of the robot

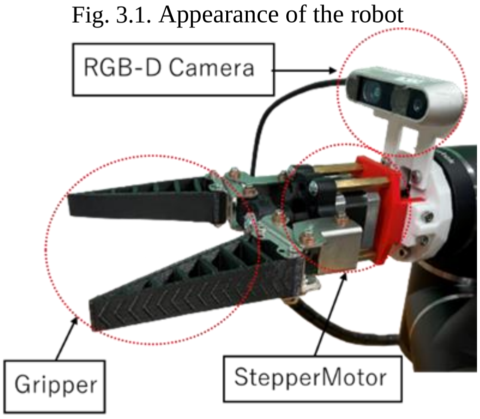
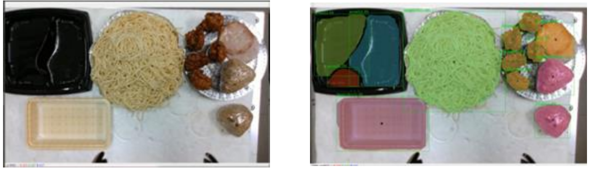


Fig. 3.2. Detail of the gripper

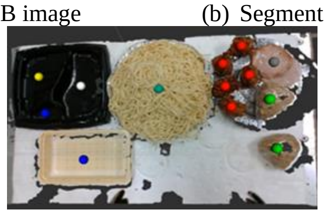
3.2. *System Configuration*

The position of the solidified food must be identified during the gripping and serving process. An object recognition system was used. First, the RGB information of the object is acquired from an RGB-D camera mounted on the tip of the robot, and instance segmentation is performed. The coordinates of the center of gravity in two dimensions are calculated for the generated segmentation image. Furthermore, 3D center-of-gravity coordinates are obtained by acquiring depth information for the object.

In addition, a microcontroller is used to control motors from a PC to open and close the grippers. Fig. 4 shows the image obtained from the RGB-D camera and the processed image.



(a) RGB image (b) Segmentation image



(b) Gravity center points

Fig. 4. Image from RGB-D camera

4. Experiment

4.1. *GUI Evaluation Experimen*[1]

Ten subjects were asked to use the application from (a) to (d) according to the application flow described in 2.2. The subjects were divided into two groups of five each. Group A was given detailed instructions on how to operate the application, while Group B was asked to use it without being told how to operate it. The subjects who completed the operation to the end were considered successful, and the percentage of successful subjects in each group was investigated. The results are shown in Table 1.

Table 1. Operation result of application user

	Successful	Total	Rate of successful [%]
Group A	5	5	100
Group B	3	5	60

4.2. *Virtual robot motion time measurement*

We considered the time aspect in the process of simulating the developed application. Compared to direct operation by code on the ROS that controls the virtual robot, operation by GUI requires a human operator to give operation instructions for each movement, resulting in a difference in operation time. In addition, the communication between the AR and the ROS also causes delays. In the process of (a) and (b), which simulate the robot motions in 2.2, we performed 10 measurements each for GUI operation in AR and direct operation in ROS, and the fastest, slowest, and average values are shown in Table 2.

Table 2. Measurement result of operation time

	GUI input	Directly input
Fastest[s]	20.95	13.13
Latest[s]	24.49	16.79
Average[s]	22.11	14.68

4.3. *Discussion*

Experiments 4.1 and 4.2 were conducted to evaluate the usability of the developed application.

From 4.1, it was found that although the application can be used by explaining how to operate it, it is difficult to use for people who use this application for the first time. In particular, the largest number of users failed in the process of serving the virtual grasping object on the simulation screen. It is expected that the number of failures in this process will be reduced by developing an application that does not accept the next operation command while the virtual robot is in operation.

In 4.2, since both PCs controlling the virtual robot are the same, the operation time is theoretically the same. However, in addition to the time between GUI operations and the communication time between AR and ROS, the time required to recognize the image to make the virtual robot appear on AR was found to cause a significant delay. The average time was 16.98 seconds when the above effects were suppressed. In particular, GUI operation on the object selection screen contributes to shorter operation time compared to direct input.

The robot can be operated by GUI, and the robot can be operated without being near the control PC.

5. Conclusion

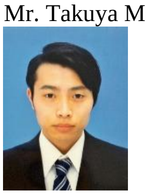
In this study, we developed an AR system that can visualize robot motions on a smartphone. In addition to the visualization, we also developed an application with higher usability. The effectiveness of this system in introducing robots was confirmed.

References

1. Kazuya Ikegami, Hidehiko Okada. [Toward Usability Quantification] Usability teiryouka ni mukete(in Japanese). NEC Technical Journal, Vol.61 No.2/2008, pp. 53-56.

Authors Introduction

Mr. Takuya Matsumoto



He received bachelor degree in Engineering in 2022 from intelligent and Control Systems, 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

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).

Dr. Sakmongkon

Dr. Sakmongkon Chumkamon received Doctor of Engineering degree from Kyushu Institute of Technology in 2017. He was a postdoctoral Guangdong Technology Presently he is a postdoctoral researcher in Kyushu Institute of Technology since 2019. His research interests include factory automation robots and social robots.

Prof. Abbe Mowshowitz

He received the Ph.D. degree from University of Michigan in 1967. He has been professor of computer science at the City College of New York and member of the doctoral faculty at the Graduate Center of the City University of New York since 1984. His current research interests lie in two areas are organizational and managerial issues in computing, and network science. In addition to teaching and research, He has acted as consultant on the uses and impacts of information technology (especially computer networks) to a wide range of public and private organizations in North America and Europe.

Projected Assist. Prof. Ayumu Tominaga

Projected Assist. Prof. Ayumu Tominaga 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

The BCRobo dataset for Robotic Vision and Autonomous Path Planning in Outdoor Beach Environment

Tan Chi Jie
Department of Mechanical Information Science and Technology, Kyushu Institute of Technology
680-4, Kawazu, Iizuka-City, Fukuoka, 820-8502, Japan

Takumi Tomokawa[1], Sylvain Geiser[1], Shintaro Ogawa[1], Ayumu Tominaga[2], Sakmongkon Chumkamon[1], Eiji Hayashi[1]
[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

E-mail: tan.jie-chi339@mail.kyutech.jp, m-san@mmcs.mse.kyutech.ac.jp, tominaga@kct.ac.jp, tomokawa.takumi163@mail.kyutech.jp, geiser.nathan-sylvain@mail.kyutech.jp, ogawa.shintaro553@mail.kyutech.jp, haya@mse.kyutech.ac.jp,
http://www.kyutech.ac.jp/

Abstract

Along with the universalization of autonomous driving and image segmentation, various datasets are available freely for anyone to use to train their own neural network which speeds up the growth of deep learning technology. However, most of the datasets target only urban environments and other offroad environments are still lacking in datasets. This paper presents a beach environment dataset, BCRobo with the aim to contribute to closing the gap of robotic visual perception in offroad environment, especially in beach.

Keywords: Dataset, Image segmentation, Deep Learning, Field Robotics, Computer Vision

1. Introduction

Deep learning is advancing at an extremely fast pace, especially in the field of image segmentation and object detection. This is because of the curation of largely labeled datasets such as CityScapes [1] and KITTI [2] dataset targeted to push forward the development of autonomous driving using computer vision and neural network. As said, dataset plays a huge part in training a neural network for deep learning regardless of how good or efficient is the algorithm, the characteristic of the trained network mainly depends on the dataset itself. This is the sole reason that there are a few varieties in the types of datasets such as TACO that focus on garbage detection, COCO dataset for indoor detection and KITTI or Cityscapes for outdoor detection.

The current challenge of deep learning is that every neural network tends to become less flexible after they are trained. Regardless of the problem of overfitting, neural networks are usually specialized at one thing only such as a network that is good in indoor detection will in turn have a substandard performance in outdoor detection. This is something that no neural network could escape from. In simpler words, with the tradeoff of flexibility, neural networks gain accuracy in specific territory. Although the problem of diversity in neural network is yet to be tackled but a simple workaround is just to have several types of datasets matching its applications.

This paper aims to push forward the advancement of image segmentation in offroad environment especially in beach environment. Offroad environment is one of the environments that still lacks exploration and dataset as urban environment is given much more focus due to the recent surge of autonomous driving cars. Nevertheless, the development of autonomous robots in offroad environments such as forest and beach exploration remain as one of the main targets of robotic researchers.

This paper begins with the description of the sensor setup, dataset setup and collection, statistics and evaluation of the dataset using 3 types of current state-of-art image segmentation network.

2. Sensor Setup

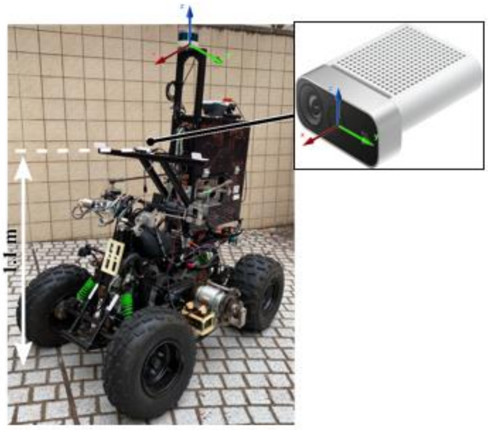


Figure 1: SOMA Sensor setup

Figure 1 shows the overall sensor setup used to capture BCRobo dataset. An autonomous forest and beach exploration robot that was built in Hayashi Laboratory, named SOMA [3] is used. SOMA is originally an All-Terrain Vehicle (ATV) which was then modified to be an autonomous driving exploration robot. On top of a creative fully automated steering mechanism, the robot is also equipped with distinct types of sensors as below:

2.1. *RGB-D sensor*

At the height of approximately 1.1m from the ground, an RGB-D sensor is installed in front of SOMA. The RGB-D camera used is Azure Kinect DK produced by Microsoft [4]. Wide Front of View (WFOV) mode is chosen for the depth camera which produces 1024x1024 depth images at the rate of 15 frames per second (fps). On the other hand, the color camera is set to produce 1280x720 MJPEG video stream at 30fps.

2.2. *Lidar sensor*

A Velodyne VLP-16 rotating 3D laser scanner is installed on top of SOMA as shown in Figure 1. This sensor is capable of recording 3D point clouds with a field of view of 360 degrees at the range from 1m to 100m. The rotation rate and accuracy are set at 5Hz and +/- 3cm [5].

2.3. *Global Positioning System (GPS)*

Emlid Reach RS+ is installed on SOMA as well which means that GPS data is also included in the form of National Marine Electronics Association (NMEA) messages. The GPS is taken with the precision up to +/- 5cm with Real-time Kinematic Positioning (RTK).

3. Dataset setup

BCRobo dataset is a beach environment dataset that consists of video sequences and lidar point clouds that were recorded by SOMA in Figure 1. The operation of the robot is manually controlled by human to explore and record the environment of several beaches in Munakata City and Kitakyushu City, Japan. A total of 6850 video frames are captured while ground truth annotations are provided for every tenth frame of a video sequence. If the tenth frame is blurred, the frame before or after might be used instead. The exploration data could be categorized based on the location below:

- Jinoshima Island – An island in Munakata City with a port area and rock bed environment.
- Agawa Hosenguri Seaside Park – A typical beach for vacation and sea bathing.

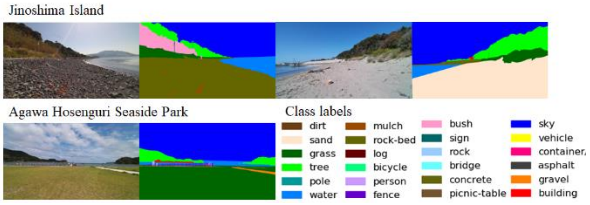


Figure 2: Example of video sequences frame and its corresponding ground truth image with class labels

Figure 2 shows the sample images taken in Jinoshima and Agawa Hosenguri Seaside Park with its corresponding ground truth image. For every location, approximately 10 minutes of data are recorded where the video frame rate is 15Hz and lidar rate is 1Hz. Note that not all video sequences frames are included in this dataset. SOMA moves at 1.6m/s at all times except during start up or climbing across some obstacles. The ground truth images in this dataset consists of 24 classes derived from KITTI and RUGD dataset.

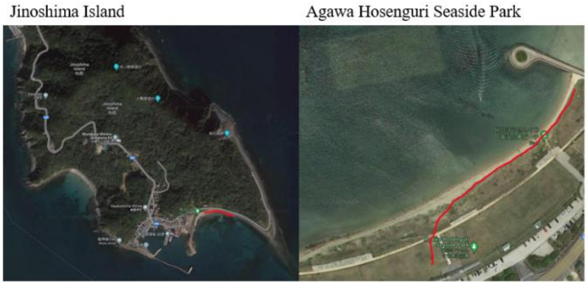


Figure 3: Predicted route of SOMA based on GPS data

With the RTK GPS sensor, we are able to pinpoint the location of the robot throughout the whole recording but there are times with bad connection with the base of the GPS and hence, the routes taken by the robot are predicted using the GPS coordinates as shown in Figure 3. SOMA is controlled and follows the red route in Figure 3 back and forth.

3.1. *BCRobo Class Labels and Statistical Analysis*

Figure 4 shows the breakdown of class annotations for BCRobo dataset. As expected, this dataset is skewed towards the sky, grass, sand and water labels. This is the expected outcome as this dataset is made specially to tackle the issue of lack of dataset for image segmentation in beach environment. For autonomous robot to transverse in a beach environment, the robot needs to have the ability to recognize the terrain accurately which is to differentiate between the traversable terrain such as sand, grass, gravel and mulch with the untraversable terrain like water and rock bed depending on the dynamic ability of the robot whether it is a wheeled or tracked robot.

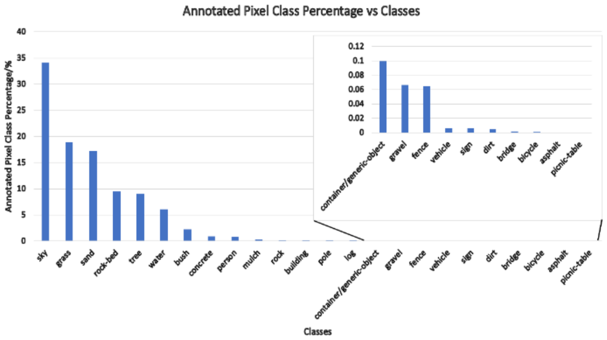


Figure 4: Annotated Class Pixel Percentages

4. Experiment and Evaluation

In order to evaluate the quality and actual usage of the dataset, three types of state-of-the-art semantic segmentation approaches are selected to be trained with the BCRobo dataset. In this experiment, the three selected semantic segmentation models are chosen based on their backbone structure which is ResNet50 [6]. The reason that ResNet50 is chosen as the fixed variable in this experiment is that despite being the very first working very deep feedforward neural network, ResNet50 still remains as the most used backbone and one of the most cited neural networks in image segmentation approaches since winning the ImageNet competition in 2015. The three selected semantic segmentation approaches are:

- PSPnet [7] – ResNet50 – d8 backbone
- OCRnet [8] – ResNet50 – d8 backbone
- UPerNet [9] – ResNet50

PSPnet is one of the earliest approaches to include global context information in scene parsing for image segmentation which eventually become the winner for PASCAL VOC and Cityscapes benchmark in 2016. This ability of examining global scene category is proven to be useful in complex scene parsing scenarios especially in beach environments where much attention is needed to different sub-regions that contain some remarkably small or large objects.

OCRnet is the most recent approach among the three approaches chosen for this experiment due to its ability to differentiate not only the same-object-class contextual pixels but also the different-object-class contextual pixels. In addition, the multi-scale context using dilated convolutions also benefits from high-resolution, large-scale contexts of this dataset.

On the other hand, UPerNet is trying to combine object classification, scene recognition, pixel-level scene parsing and texture recognition in a single neural network along with a novel learning method. Having the same Pyramid Pooling Module (PPM) as PSPnet, UPerNet applies one PPM on scene, object, part and material recognition each to achieve the said unified perceptual parsing.

4.1. *Experimental Setup*

Like every other semantic dataset, the video frames are separated into train, validation and test sets for this experiment. 80% of the annotated ground truth video frames are partitioned into train set while the leftover 20% split evenly between validation and test sets. This dataset includes two different beach environments and to ensure that the semantic segmentation models are trained on both environments, the splitting ratio mentioned before this is applied on the two beach environments individually and then combined to form the final train, validation and test sets as in Table 1.

Table 1. Train, Validation and Test sets.

	Jinoshima	Agawa	Total	%
Train	315	233	548	80.00
Validation	39	30	69	10.07
Test	39	29	68	9.93

The training environment for all three models are setup as below:

- Ubuntu LTS 20.04
- AMD Ryzen Threadripper 3960X 24-Core
- Nvidia RTX 3090 – 3 units
- MMSegmentation v0.29.1 [10]

The images are first downsized to 688x550 before passing into the neural network for training and the crop size is set at 300x375. Batch size is set at 6 per GPU making it 18 since 3 GPUs are used. Stochastic Gradient Descent (SGD) optimizer with momentum [11] is selected with the parameter of learning rate 0.015 and momentum 0.9. The weight decay is set at 0.0004. “Polynomial learning rate” policy with warmup is chosen to avoid over training at the start of training. In other words, the learning rate will increase linearly for 1000 iterations until it reaches 0.015 then decay in a polynomial fashion until it reaches the minimum learning of 0.0001 for the whole training. The modals are trained for 2000 epochs which is around 60000 iterations using MMSegmentation which is an open-source semantic segmentation toolbox.

4.2. *Experimental Evaluation*

The performance of all three models is evaluated using the standard semantic segmentation metrics which are mean Intersection-over-Union (mIoU) and mean pixel-wise classification accuracy (mAcc). mIoU is the mean of IoU of each class given that, IoU is TP/(TP+FP+FN) [12] where TP is true positive, FP is false positive, and FN is false negative. Besides, mAcc is just the mean of pixel classification accuracy (aAcc) of all classes. The evaluation is first performed by inferring the test and validation sets using the trained modals as shown in Table 2. Another evaluation is also performed on all three sets as well as shown in Table 3.

Table 2. Evaluation on Test + Validation sets

	PSPnet, %	OCRnet, %	UPerNet, %
mIoU	73.90	74.64	75.34
mAcc	81.74	83.26	84.22
aAcc	98.09	98.06	97.83

Table 3. Evaluation on all sets

	PSPnet, %	OCRnet, %	UPerNet, %
mIoU	71.76	72.32	71.70
mAcc	78.22	79.71	79.14
aAcc	98.20	98.12	97.86

Overall, it is observed that despite the sets used in evaluation, mIoU for all modals achieves a reasonably good rate above ~70% which means that all the models are learning the visual classes correctly. The aAcc are high as well around ~98% but we do observe some degree of degrade at mAcc. This is probably due to the irregular boundaries which is very common in beach environments due to the constant changing water tide, shape of sands, changing position of tree branches and leaves due to the windy condition in beach.

5. Conclusion and future work

In the nutshell, BCRobo dataset is a highly specialized dataset that contains high resolution beach environment images captured by a field exploration robot, SOMA. For this reason, any image segmentation model trained with this dataset would expect higher mIoU compared to other major dataset as this dataset is skew towards the major class labels which is sky, sand, water and tree.

As a conclusion, this dataset is proven helpful in performance image segmentation for beach environment from the experimental data evaluated using PSPnet, OCRnet and UPerNet that achieve over ~70% mIoU. However, the high mIoU would also mean that the models trained using this dataset might perform badly in other environments besides beach. Nevertheless, this is the common downside of current neural network models which lose diversity as they gain accuracy. Therefore, it is recommended to use this dataset with other dataset if the scene for image segmentation is not limited to beach only.

As for now, BCRobo dataset is focus only on beaches around the Southern part of Japan, namely Kyushu area. In the future, we will continue to expand this dataset by adding more images and its annotated ground truth images of different beach environments in other part of Japan. The dataset is available for download at https://github.com/chijie1998/BCRobo-dataset with the 685 annotated ground truth images and its corresponding video frames. Full video sequences and lidar point cloud would be provided upon request due to its large sizes.

6. References

1. M. O. S. R. T. R. M. E. R. B. U. F. S. R. a. B. S. M. Cordts, "The cityscapes dataset for semantic urban scene understanding," in *IEEE Computer Vision and Pattern Recognition*, 2016.
2. P. L. C. S. a. R. U. A. Geiger, "Vision meets robotics: The kitti dataset," in *International Journal of Robotics Research*, 2013.
3. E. H. R. F. N. Takegami, "Environment map generation in forest using field robot," in *Proceedings of International Symposium on Applied Science*, 2019.
4. S. M. B. T. T. E. S. W. O. A. A. P. J. G. M. F. V. R. e. a. C. S. Bamji, "Impixel 65nm bsi 320mhz demodulated tof image sensor with 3μm global shutter pixels and analog binning", in *IEEE International Solid - State Circuits Conference - (ISSCC)*, 2018.
5. J. R. Kidd, "Performance Evaluation of the Velodyne VLP-16 System for Surface Feature Surveying," University of New Hampshire, 2017.
6. X. Z. S. R. a. J. S. K. He, "Deep residual learning for image recognition.," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2016.
7. J. S. X. Q. X. W. J. J. Hengshuang Zhao, "Pyramid Scene Parsing Network," in *Computer Vision and Pattern Recognition*, 2017.
8. A. G. N. A. a. J. G. V. Gupta, "OCRNet - Light-weighted and Efficient Neural Network for Optical Character Recognition," in *IEEE Bombay Section Signature Conference (IBSSC) 2021*, 2021.
9. T. a. L. Y. a. Z. B. a. J. Y. a. S. J. Xiao, "Unified perceptual parsing for scene understanding," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018.
10. M. Contributors, "OpenMMLab Semantic Segmentation Toolbox and Benchmark," 10 7 2020. [Online]. Available: https://github.com/open-mmlab/mmsegmentation
11. Y. G. Y. Yanli Liu, "An Improved Analysis of Stochastic Gradient Descent," in *34th Conference on Neural Information Processing Systems (NeurIPS)*, Vancouver, Canada, 202.
12. A. Rosebrock, "Intersection over Union (IoU) for object detection," pyiamgeserach, 7 Novemeber 2016. [Online]. Available: https://pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection/

Authors Introduction

Mr. Tan Chi Jie

He received his Bachelor of Engineering Electronics Majoring in Robotics and Automation from the Faculty of Engineering, Multimedia University, Malaysia in 2020. He is currently a Master student at Kyushu Institute of Technology and conducts research at Hayashi Laboratory.

Mr. Takumi Tomokawa

He received bachelor degree in Engineering in 2021 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.

Mr. Sylvain Geiser

He studied Engineering and Computer Science in Ecole des Mines de Nancy, France, between September 2018 and June 2020. He is currently a Master student at Kyushu Institute of Technology, Japan, and conducts research at Hayashi Laboratory.

Mr. Shintaro Ogawa

He received bachelor degree in Engineering in 2022 from intelligent and Control Systems, Kyushu Institute of Technology in Japan. He is currently a Master student at Kyushu Institute of Technology and conducts research at Hayashi Laboratory.

Projected Assist. Prof. Ayumu Tominaga

Projected Assist. Prof. Ayumu Tominaga 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.

Dr. Sakmongkon Chumkamon

Dr. Sakmongkon Chumkamon received Doctor of Engineering degree from Kyushu Institute of Technology in 2017. He was a postdoctoral researcher at Guangdong University of Technology in 2017-2019. Presently he is a postdoctoral researcher in Kyushu Institute of Technology since 2019. His research interests include factory automation robots and social robots.

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).

The research about editing system of performance information for player piano.
-Develop inference methods using machine learning -

Takaaki Ueno
Graduate School of Creative Informatics, Kyushu Institute of Technology
Castail Iizuka 318, 268-1, Kawazu, Iizuka-City, Fukuoka, 820-0067, Japan

Sakmongkon Chumkamon
Department of Mechanical Information Science and Technology, Kyushu Institute of Technology
680-4, Kawazu, Iizuka-City, Fukuoka, 820-8502, Japan

Eiji Hayashi
Department of Mechanical Information Science and Technology, Kyushu Institute of Technology
680-4, Kawazu, Iizuka-City, Fukuoka, 820-8502, Japan

E-mail:
ueno.takaaki568@mail.kyutech.jp

Abstract

In 1996, Professor Eiji Hayashi of this laboratory developed an automatic piano performance system that aims to reproduce performance skills equivalent to those of humans.[1] However, this system does not have a function to infer performance expressions such as the strength and weakness of notes, and is not capable of expressing emotions such as intonation for each phrase, as is done by actual human performers. Therefore, the goal of this research was to reproduce the performances of world-famous pianists under the objective of "pursuit of more human-like performance expression. In the future, when this research matures and the human expression of the piano is clarified mathematically, the data will be used to reproduce human performances, which will fulfill the original purpose of this research. In this paper, we attempted to reproduce "Prelude Op.4 No.10" composed by Fryderyk Franciszek Chopin, using the world-famous pianist "Vladimir Davidovich Ashkenazy" as the research subject.

Keywords: Neural Network, Machine Learning, Long Short-Term Memory, Recurrent Neural Network, Automatic Piano, Computer Music

1. Introduction

In order for an automatic piano to perform in a human-like manner, it needs data on the intonation of each note. However, existing automatic pianos do not have the ability to infer such data. Therefore, until now, data inference has been performed manually. Therefore, we attempted to construct a system that can perform data inference automatically by using machine learning. In this paper, we describe the actual performance information inference system we developed using machine learning.

2. Editing Support System

2.1 Performance Information

The songs shown in Table 1 of CrestMusePEDB, created from MIDI data of a performance by Vladimir Davidovich Ashkenazy (Russia), were used as performance information. Hereafter, the song names will be A to F in Table 1.

Table 1. Information of songs

Sign	musical composition title	Tonality
A	Prelude Op.28 No.4	E minor
B	Prelude Op.28 No.7	A major
C	Prelude Op.28 No.20	C minor
D	Waltz Op.69 No.1	A flat major
E	Waltz Op.64 No.2	C sharp minor
F	Prelude Op.28 No.15	D flat major

2.1 Sheet Music Information

Table 2 shows the variables used in the MIDI (Musical Instrument Digital Interface) standard for musical notation information, where Key represents keyboard position, Velo represents note strength, Gate represents note length, and Time represents playing time.

The information on the strength and weakness of the note, such as forte and piano, is stored in the Dyn variable, and the interval between the next note is stored in the Step variable.

Table 2. Used variable (MIDI)

Parameter	Key	Gate	Velo	Time
Unit	—	ms	—	ms
Reference	21~108	—	1~127	—

2.2 LSTM Model Construction

Long Short-Term Memory (LSTM) was adopted as the inference system for this study. This is an improved model of Recurrent Neural Network (RNN), which is a kind of general circular neural network.

RNN has a continuous relationship between input and output by adding the output of one time to the input of the next time, and can learn data as time-series information by adding the output of each time to the input of each time as memory information. The ability to learn as time-series data is mainly used in the field of natural language processing, such as machine translation of languages. This is because machine translation requires not only word-by-word translation, but also judgment based on the context before and after the translation. The model of LSTM, which is an improved model of RNN, is shown in Figure 1 below.

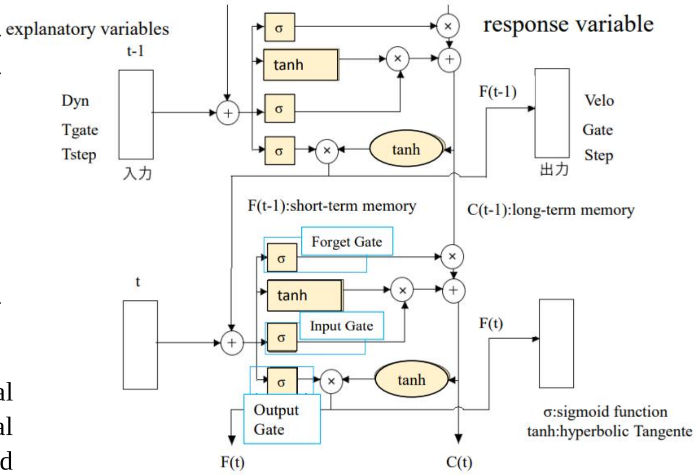


Fig. 1. Abstract of LSTM

LSTM is equipped with a function that divides the given memory information into long-term and short-term memory, and by selecting information at any time and adding important information as long-term memory, it is possible to learn a longer time series than a normal RNN.

3. Inference Experiment

Major headings should be typeset in boldface with the first letter of important words capitalized.

3.1. experimental procedure

As a preliminary step to train the machine learning model, we mapped the created music data to the performance data. 5 songs from A to E were mapped to each other, and then each song was trained for 1000 epochs with 4 songs from A to D as teacher data and E as evaluation data, and inference of F was performed.[2] In learning, Velo, Gate, and Step were placed as objective variables, and Dyn, Tgate, and Tstep, which correspond to the score information of each song, were placed as explanatory variables, judging that they were the most correlated values. After training, the music data of song F, the object of inference, was input to the model, and the performance data obtained as output was evaluated.

3.2. experimental results

Figure 2, Figure 3, and Figure 4 show the results of comparing the inferred values for each of Velo, Gate, and Step with the actual performance (MIDI data).

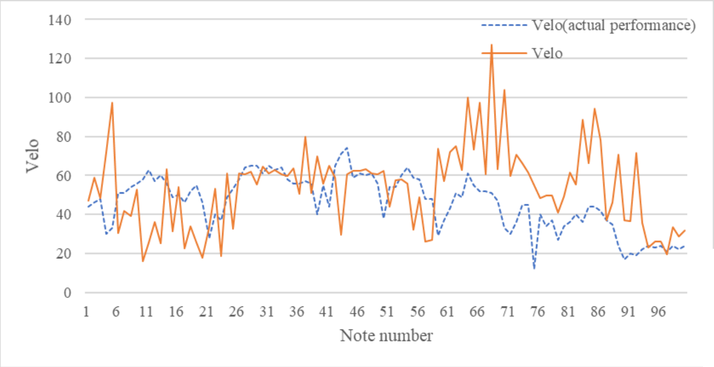


Fig. 2. Velo comparison of the right hand

T3

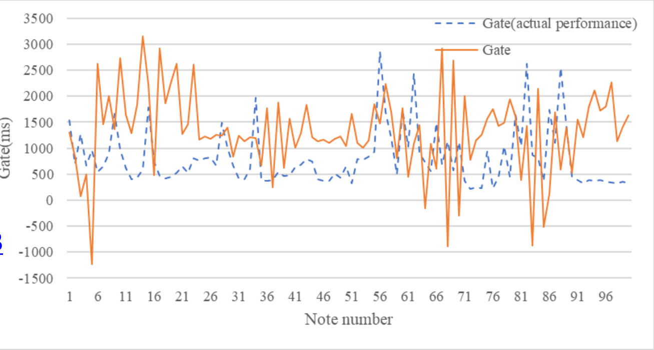


Fig. 3. Gate comparison of the right hand

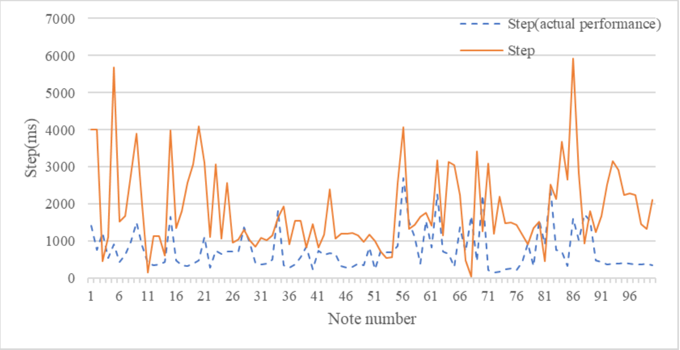


Fig. 4. Step comparison of the right hand

The correlations between the inferred values of Velo, Gate, and Step and the performance (MIDI data) are shown in Table 3, and the loss trends for each piece are shown in Figure 5. The loss function is the mean square error.

Table 3. Correlation with actual performance

	correlation
Velo	0.127
Gate	-0.110
Step	0.245

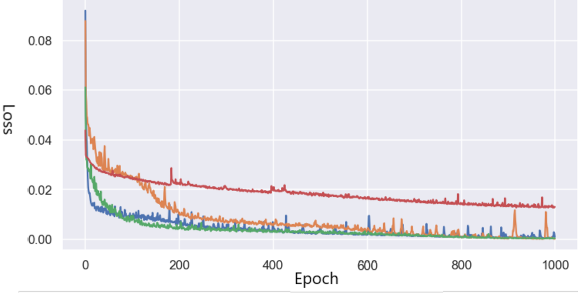


Fig. 5. Loss of 4 songs

4. Consideration

Figure 5 shows that the learning is correct, but Table 3 indicates that the accuracy of inference is low. The fact that few positive correlations were observed suggests that the combination of explanatory and objective variables used as the axis of evaluation for inference was inappropriate.

In addition, in Figure 2, the value of Velo exceeded the upper limit of 127, and in Figure 3, the value of Gate was negative in some cases, indicating that the inference results sometimes included values that were impossible to perform. In addition, abnormal values were output for Step as a performance of about 6000 (ms) in some cases. This may be due to the fact that the model was trained with data that had no output threshold in the model training process, and that data with sharp ups and downs was used as the teacher data.

5. Conclusion

. This year, we constructed a system using machine learning to replace the conventional inference system and conducted inference experiments. This year's experiments made it possible for the system to learn performance information (MIDI data). However, some of the values in the performance inference exceeded the limit of expression, and few positive correlations were observed.

As a future prospect, we aim to construct a system that is capable of sufficient inference. For this purpose, we will review the combination of variables considered in this study, and in addition, we will consider performing inference on other songs and comparing the results.

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.
2. Y. Isomichi, T. Ogawa, Pattern matching by using dynamic programming, J. Inf. Process Soc. Japan 16 (1975), 15–22 (in Japanese).

Authors Introduction

Mr. Takaaki Ueno

He received his B.S. degree in Information Engineering from the Department of Intelligent Systems Engineering, Kyushu Institute of Technology in 2022. He is currently enrolled in the master's program at Kyushu Institute of Technology, where he is conducting research in the Hayashi Laboratory.

Dr. Sakmongkon Chumkamon

Dr. Sakmongkon Chumkamon received Doctor of Engineering degree from Kyushu Institute of Technology in 2017. He was a postdoctoral researcher at Guangdong University of Technology in 2017-2019. Presently he is a postdoctoral researcher in Kyushu Institute of Technology since 2019. His research interests include factory automation robots and social robots.

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).

Automated Random Simulation for Checking a Behavioral Model of Systems Based on Extended Place/Transition Net with Attributed Tokens

Sho Matsumoto
Department of Engineering and Design, Faculty of Engineering and Design, 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 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
E-mail: s19t332@kagawa-u.ac.jp, kat@cs.miyazaki-u.ac.jp, takagi@eng.kagawa-u.ac.jp

Abstract

Extended Place/transition Net with Attributed Tokens (EPNAT) is one of formal modeling languages, and it enables system engineers to construct an executable and abstracted behavioral model of multiple software systems. In this paper, we propose an automated random simulation technique of an EPNAT model in order to detect failures in the model. In the simulation, input is randomly selected for model execution. When a constraint given for each system or multiple systems is violated through the model execution, a failure is revealed. The simulation is terminated by the detection of a failure or the satisfaction of a criterion focusing on the combination of marking, data writing and reading between different systems. A prototype tool of the simulation technique was developed and applied to a trial model to discuss its effectiveness. Two failures inserted into the trial model were successfully detected, and a few challenges were found in the experiment.

Keywords: formal modeling, place/transition net, VDM, simulation, failure detection

1. Introduction

The collaborative behavior of multiple software systems can provide advanced services and functions, but also can be a contributory factor in the occurrence of failures. Systematic quality control techniques need to be constructed so that system engineers can detect such failures and improve the reliability effectively. Therefore, in this paper, we propose an automated random simulation technique for checking an abstracted collaborative behavior of multiple systems based on an Extended Place/transition Net with Attributed Tokens (EPNAT).

EPNAT [1] is a relatively new formal modeling language developed from PN (Place/transition Net), and includes the following additional elements written in VDM++ [2]:

- Attributes (variables for each token to characterize the behavior of systems or subsystems)
- Actions (procedures of data processing in each transition of systems or subsystems)
- Basic constraints (conditions that should be satisfied in each variable and transition, that is, pre-conditions, post-conditions, and invariants)

By using EPNAT, engineers will be able to construct an executable and unambiguous model for the abstracted collaborative behavior of multiple systems, which is called EPNAT model or simply model in this paper. In our previous study [3], a semi-automated simulation technique was proposed to visualize the behavior of a given EPNAT model and arbitrarily search its common failures. In the semi-automated simulation, engineers need to manually select input to execute a model, and need to manually evaluate some constraints. Note that basic constraints can be automatically evaluated, but there are complex requirements among systems that are not so easy to be defined as basic constraints. On the other hand, in the automated random simulation we propose in this paper, the selection of input and the evaluation of non-basic constraints are automatically executed to reduce engineers' effort. Another study [4] shows a test stopping criterion to cover successive markings and transitions based on coverability trees of a PN model, but it is not suitable for the automated random simulation of an EPNAT model. Therefore, a new stopping criterion for the model execution is introduced to concentrate on the detection of failures on the collaborative behavior. It focuses on the combination of marking, data writing and reading between different systems. This idea is related to data flow testing [5].

This paper is organized as follows. In section 2, we propose the automated random simulation technique that consists of a model execution algorithm and a stopping criterion of the model execution. Section 3 shows the results of a preliminary experiment using our prototype tool and a trial model, and gives discussion about the effectiveness of the proposed technique. Finally, section 4 describes conclusion and our future work.

2. Automated Random Simulation Technique

In this section, we propose the automated random simulation technique that consists of a model execution algorithm and a stopping criterion of the model execution.

2.1. *Model Execution Algorithm*

The input to the algorithm is a model to be checked, non-basic constraints, and the initial state of the model. The non-basic constraints are based on the engineers' aim of checking, and are defined by the engineers for markings, variables, and fire of transitions. They are classified into type A and B according to the timing of evaluation. Type A means that the constraints should be satisfied among systems at all times or at the timing of specific execution, and they can often be replaced with basic constraints. Type B means that sometime the constraints should be satisfied among systems. A state of a model is represented by a marking, and values of variables in this study. The initial state is defined by the engineers according to their aim of checking.

The output from the algorithm is a failure of the model. If the stopping criterion has been satisfied without any failures, the output is null. Also, the process of model execution can be outputted as animation, since EPNAT is a visual language.

The algorithm consists of the following six steps:

[Step 1] The given model is initialized to the given initial state.

[Step 2] A set of transitions that may be fireable is created according to only the current marking. Hereinafter, the set is referred to as *T*. Note that the fireability is determined by not only the current marking but also pre-conditions of each transition that are addressed in a later step.

[Step 3] If *T* is empty and the current state is not any of final states that were defined as a part of the given model, or if *T* is not empty and the current state is one of the final states, then this situation is reported as a failure and this algorithm is terminated. If *T* is empty and the current state is one of the final states, then this algorithm returns to Step 1. Otherwise, a transition is randomly selected from *T*. Hereinafter, the selected transition is referred to as *t*.

[Step 4] The pre-condition of *t* is evaluated. If the pre-condition includes any arguments of *t*, their values are randomly selected from a set of valid values that makes their invariants and the pre-condition true. If the set of valid values is empty or if the result of the evaluation is false, then *t* is removed from *T* and this algorithm returns to Step 3.

[Step 5] *t* is fired with valid values that are randomly selected for its arguments. The action of *t* is executed, and then the current state is changed. All the related invariants, the post-condition of *t*, and the given non-basic constraints (type A) are evaluated to detect failures. If there are any constraints whose evaluation results are false, then this situation is reported as a failure and this algorithm is terminated.

[Step 6] If the stopping criterion discussed in the next subsection is not satisfied, this algorithm returns to Step 2. Otherwise, the given non-basic constraints (type B) are evaluated to detect failures. If there are any constraints whose evaluation results are false, this situation is reported as a failure. This algorithm is terminated.

2.2. *Stopping Criterion*

In EPNAT models, glue transitions play an important part in the collaborative behavior of multiple systems, and thus we focus on the glue transitions in order to construct the stopping criterion.

One model consists of multiple sub-models that correspond to multiple systems. The sub-models are connected with each other by glue transitions. When a glue transition is fired, the writing and/or reading of attributes among connected sub-models is executed in its action. A result of the reading of an attribute depends on the preceding writing of the attribute. Also, the results of the writing and reading are related to markings in which they are executed.

According to the above, the combinations of markings, writing and reading that are related to glue transitions should be covered in the model execution. A set of the combinations, which is referred to as C, is defined as follows:

$$C = \bigcup_{a\in A} \left(\bigcup_{w\in W_a} (M_w \times \{w\}) \times \bigcup_{r\in R_a} (M_r \times \{r\}) \right) \qquad (1)$$

where A expresses a set of attributes. W_a and R_a express sets of the writing and reading of attribute a ($a \in A$), respectively. One of or both of w ($w \in W_a$) and r ($r \in R_a$) in each element of C should be included in action of a glue transition. M_x ($x \in W_a \cup R_a$) expresses a set of markings in which x is executed.

Each element of C is called check object in this study. Not all check objects defined by equation (1) are usually feasible, and the strict evaluation of feasibility on software is known as a hard problem to solve. Therefore, when the number of check objects that have been actually reached through the model execution converges, Step 6 concludes that the stopping criterion is satisfied.

3. Discussion

This section shows the results of a preliminary experiment using our prototype tool and a trial model.

In this study, we are developing a prototype of a tool that has the functions to edit an EPNAT model and to execute the automated random simulation of the model. A screenshot of the prototype tool is shown in Fig. 1. Also, we created a trial model that represents the collaborative behavior of a simple student management system and book management system. The model satisfies the following two non-basic constraints:

[C1] Each student cannot borrow multiple books.

[C2] Sometime each student can log out.

C1 and C2 are type A and B, respectively.

We executed the automated random simulation by applying the trial model, C1, C2 and an initial state of the trial model to the prototype tool. The overview of its result is shown in Fig. 2. No failures had been detected, and it met our expectations. 64 check objects had been successfully reached on the trial model. The number of the reached check objects converged when transitions had been fired about 2,500 times. It grows rapidly in the earlier stage of the model execution, since the algorithm can easily find new execution paths on the given model. There are check objects that the algorithm failed to reach, which will be due to its random search approach.

Additionally, we created two faulty models F1 and F2 that do not satisfy C1 and C2 respectively, and then similarly executed the automated random simulation of

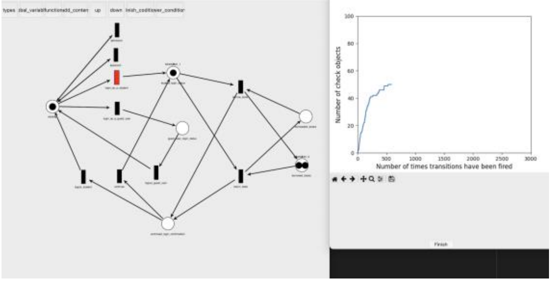


Fig. 1. A screenshot of our prototype tool.

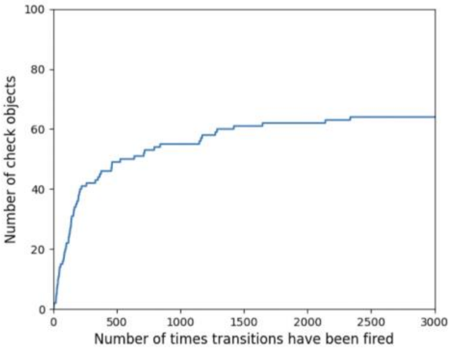


Fig. 2. Overview of an experimental result.

them five times. As their results, when transitions had been fired an average of 17.6 times in F1, the simulation was automatically terminated, and successfully reported a failure that is the violation of C1. The simulation of F2 was terminated after the number of reached check objects had been converged. It reported that C2 had not been satisfied, and implied the existence of a failure. Note that type B constraints cannot reveal failures as established facts. It is expected that the proposed technique will be able to detect failures to some extent. However, the faulty models and non-basic constraints used in this experiment are relatively small and simple, and thus additional experiments are needed to strictly evaluate its effectiveness.

4. Conclusion and Future Work

In this paper, we proposed an automated random simulation technique of an EPNAT model in order to detect failures in the model. Also, we developed a prototype tool of the technique, and applied a trial model to it. Two failures inserted into the trial model were successfully detected, and a few challenges were found in the experiment.

In future study, we plan to (1) improve the model execution algorithm so as to reach check objects efficiently, (2) develop a technique to express more complex non-basic constraints, and (3) evaluate the effectiveness by using large and complex models.

Acknowledgements

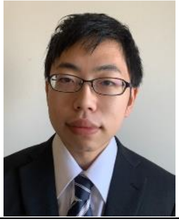
A part of this work was developed from the graduation thesis written by Mr. Hikaru Morimoto at Kagawa University. This work was supported by JSPS KAKENHI Grant Number JP22K11976, and Young Scientists Fund of Kagawa University Research Promotion Program 2021 (KURPP).

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. J. Fitzgerald, P.G. Larsen, P. Mukherjee, N. Plat and M. Verhoef, *Validated Designs for Object-Oriented Systems*, Springer-Verlag London, 2005.
3. T. Takagi and R. Kurozumi, Simulation and Regression Testing Technique for Software Formal Specifications Based on Extended Place/Transition Net with Attributed Tokens, *Journal of Robotics, Networking and Artificial Life*, Vol.8, No.2, pp.112-116, Sep. 2021.
4. 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.
5. B. Beizer, *Software Testing Techniques*, Van Nostrand Reinhold, 2nd edition, 1990.

Authors Introduction

Mr. Sho Matsumoto



He is an undergraduate student in the Faculty of Engineering and Design 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.

Training of Software Formal Modeling Using Visual Blocks for Actions and Guards of Extended Place/Transition Net

Akio Usuda
Division of Reliability-based Information Systems Engineering, Graduate School of Engineering, Kagawa University
2217-20 Hayashi-cho, Takamatsu-shi, Kagawa 761-0396, Japan

Ryoichi Ishigami
Department of Engineering and Design, Faculty of Engineering and Design, Kagawa University
2217-20 Hayashi-cho, Takamatsu-shi, Kagawa 761-0396, 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
E-mail: s21g457@kagawa-u.ac.jp, s19t301@kagawa-u.ac.jp, takagi@eng.kagawa-u.ac.jp

Abstract

Extended Place/transition Net (EPN) that is one of software formal modeling languages consists of the parts of PN and VDM++. The PN part deals with state transitions of software. On the other hand, the VDM++ part deals with actions and guards on the transitions, and requires skills of system engineers who construct EPN models. We propose an extended training technique of EPN-based modeling using visual blocks for the VDM++ part. The visual blocks implemented by Blockly will be useful to accelerate the trainees' understanding of syntactical aspects of VDM++, and thus they are introduced into each step of the existing training technique for EPN-based modeling. The effectiveness of the proposed technique is discussed through a preliminary experiment using our prototype tool.

Keywords: software modeling language, place/transition net, VDM, training

1. Introduction

Extended Place/transition Net (EPN) [1] is one of formal modeling languages that can be used to analyze, design, and test software from the abstracted viewpoint of its behavior. It consists of the parts of PN [2] and VDM++ [3]. The PN part deals with state transitions of software. On the other hand, the VDM++ part deals with actions and guards on the transitions.

System engineers will sometimes face the difficulty of using EPN, since it includes special syntax. To address this problem, some studies of training/learning support techniques for software modeling have been conducted. As the first of a series of the studies, [4] describes a training support technique for bug fixing in EPN-based modeling. In the technique, animated graphics were introduced so that trainees can intuitively understand the meaning of given faulty EPN models and the results of their bug fixing. After that, in order to focus on modeling itself, Ue et al. [5] modified the previous technique, and constructed a learning support technique for PN-based modeling. Additionally, in order to focus on EPN that has higher representation power than PN, we extend the previous technique, and proposed a training technique for EPN-based modeling [6]. The technique was designed for personal on-demand training, and consists of the following three steps (overview):

[Step 1] A trainer creates an exercise that includes software requirements (that is, questions in the exercise), correct EPN models that satisfy the software requirements (that is, correct answers in the exercise), and components that the correct/incorrect EPN models consist of.

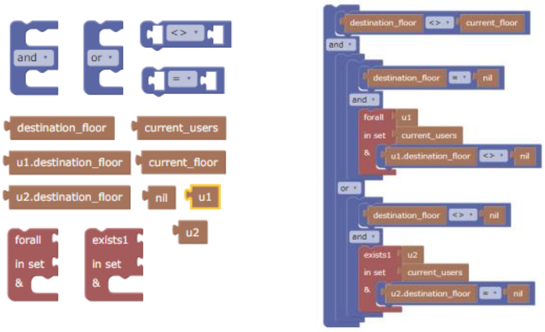
[Step 2] A trainee reads the software requirements, and tries to assemble his/her EPN model from the components.

[Step 3] The trainee's EPN model is checked according to the correct EPN models.

In order to support the trainee, Step 2 provides (i) VDM++ specifications that are generated from the trainee's EPN model and (ii) reference documents of VDM++. However, there is still room to improve the training for VDM++ part of EPN.

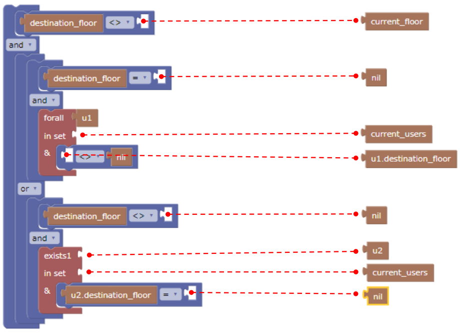
We propose an extended training technique of EPN-based modeling using visual blocks for the VDM++ part. The visual blocks that look like the pieces of a jigsaw puzzle are well-known and well-used in the field of programming education, and play an important role in visual programming. Existing visual programming support tools such as Blockly [7] can be used in our technique. The visual blocks will be useful to accelerate the trainee's understanding of syntactical aspects of VDM++, and thus they are introduced into each step of the previous training technique for EPN-based modeling.

This paper is organized as follows. The extended training technique is proposed in section 2, and then the effectiveness is discussed in section 3. Finally, we show conclusion and future work in section 4.



(a) (b)

Fig. 1. Example of visual blocks and a block-structured guard.



(a) (b)

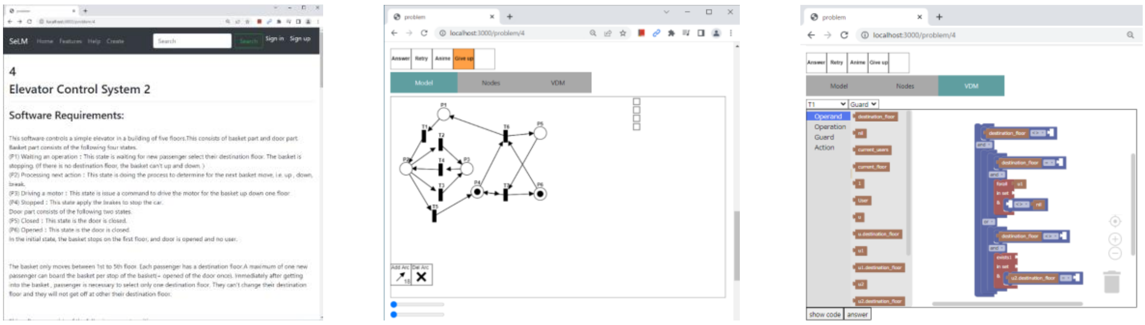
Fig. 2. Example of block-based components
(a fill-in-the-blank component and choice components).

2. Extended Training Technique

In this section, we propose an extended training technique of EPN-based modeling using visual blocks for the VDM++ part (hereinafter, the extended technique). Each step of the previous technique discussed in section 1 is extended as follows.

In Step 1, a trainer creates visual blocks, and then block-structured actions and guards (that is, the VDM++ parts of the correct EPN models). The visual blocks are chiefly classified into statements/expressions, operators, and operands. An example of the visual blocks that were implemented by Blockly is shown in Fig. 1 (a). The block-structured actions and guards are assembled from the visual blocks. An example of a block-structured guard that were implemented by Blockly is shown in Fig. 1 (b). After that, the trainer creates block-based components to give a trainee fill-in-the-blank questions with multiple choices. The block-based components are classified into fill-in-the-blank components and choice components. They are created by disassembling the block-structured actions and guards. The way of disassembling determines the degree of difficulty of the exercise. An example of a fill-in-the-blank component and choice components, which was created from Fig. 1 (b), is shown in Fig. 2 (a) and (b), respectively. Additionally, in order to increase the degree of the difficulty, the trainer will be able to create faulty (mutated) block-based components that lead a trainee to construct his/her incorrect EPN models. Note that the PN part and other materials of the exercise in the extended technique are the same as ones in the previous technique.

In Step 2, the way of constructing the PN part in the extended technique is the same as the way in the previous technique. If a trainee selects a transition to construct its action or guard in his/her EPN model, he/she receives a complete set of the block-based components. The trainee



4

Elevator Control System 2

Software Requirements:

This software controls a simple elevator in a building of five floors.This consists of basket part and door part.
Basket part consists of the following four states.
(P1) Waiting an operation : This state is waiting for new passenger select their destination floor. The basket is stopping. (If there is no destination floor, the basket can't up and down.)
(P2) Processing next action : This state is doing the process to determine for the next basket move, i.e. up , down, break.
(P3) Driving a motor : This state is issue a command to drive the motor for the basket up down one floor.
(P4) Stopped : This state apply the brakes to stop the car.
Door part consists of the following two states.
(P5) Closed : This state is the door is closed.
(P6) Opened : This state is the door is closed.
In the initial state, the basket stops on the first floor, and door is opened and no user.

The basket only moves between 1st to 5th floor. Each passenger has a destination floor.A maximum of one new passenger can board the basket per stop of the basket(= opened of the door once). Immediately after getting into the basket , passenger is necessary to select only one destination floor. They can't change their destination floor and they will not get off at other their destination floor.

(a) Section of software requirements (b) Editor for the PN part (c) Editor for the VDM++ part

Fig. 3. Partial screenshots of the prototype tool.

tries to select an appropriate fill-in-the-blank component from the set. Additionally, the trainee tries to choose appropriate choice components from the set, and then tries to fill in the blanks with them. After that, the action or guard of the selected transition is immediately checked according to one in the correct EPN models. In this study, we use Blockly and are developing a prototype of a tool that supports the extended technique. Fig. 3 shows partial screenshots of the prototype tool that supports this step.

In Step 3, the PN part in the trainee's EPN model is checked according to one in the correct EPN models. Note that the constructing of the PN part, the constructing and checking of the VDM++ part are concurrently performed in Step 2 in the extended technique.

3. Discussion

We performed a preliminary experiment using prototype tools to discuss the effectiveness of the extended technique.

We created two types of exercises on the subject of an elevator control system, and set them on the old prototype tool (discussed in [6]) and the new prototype tool (discussed in section 2), respectively. After that, three trial users, that is, three undergraduate students in our laboratory tackled the exercises in the order of using the new one and the old one. In the three, the two are beginners of EPN, and the rest is researching on an EPN-related field. Before the tackling of the exercises, we made oral explanations about EPN and how to use the prototype tools to the three. During the tackling, we observed the behavior of each of the three. Also, we added oral explanations, if one had any questions. After the tackling, the three answered our questionnaire. After the first trial user finished tackling, we fixed some bugs and improved the description of software requirements. As the result of this preliminary experiment, we found the following:

- In the new one, the trial users seemed to need more detailed software requirements, and seemed to think that the degree of the difficulty of the exercises is relatively high. In other words, the new one will give more practical training, since the trial users seemed to need a deeper understanding of software requirements to construct their EPN models.
- In the new one, the trial users seemed to be able to intuitively understand the syntactical rules and structures of VDM++ to some extent by the visual blocks that were implemented by Blockly. Also, Blockly seemed to have improved the usability of our prototype tool.
- In the new one, the early checking of each action and guard in Step 2 seemed to help the trial users to understand and construct.
- The preparation of the exercises in the old one needed a great amount of our efforts, but it was smaller than the amount of our efforts to prepare the exercises in the new one.

We will need the improvement of the prototype tool and additional experiments with many trial users to reveal the effectiveness more clearly.

4. Conclusion and Future Work

In this paper, we proposed an extended training technique of EPN-based modeling using visual blocks for VDM++ part. In the step of creating an exercise, a trainer creates block-based components for actions and guards in order to give a trainee fill-in-the-blank questions with multiple choices. In the step of tackling the exercise, each action or guard of a selected transition in a trainee's EPN model is immediately checked according to one in trainer's correct EPN models. We performed a preliminary experiment using prototype tools, and found that the extended technique will give more practical training, will help trainees to understand to some extent, will require a greater amount of trainer's efforts to prepare exercises, and so on.

In future study, we need to discuss (1) the effectiveness of giving more detailed software requirements to trainees, (2) training techniques based on abstracted software requirements, and (3) techniques of supporting the preparation of exercises. Furthermore, the visual blocks will be useful also to construct EPN models in actual software development, and may be introduced into EPN-based software modeling tools.

Acknowledgements

A part of this work was developed from the graduation thesis written by Mr. Yuya Morinishi at Kagawa University. This work was supported by JSPS KAKENHI Grant Number JP17K00103.

References

1. T. Takagi and R. Kurozumi, Prototype of a Modeling Tool to Convert between Extended Place/Transition Nets and VDM++ Specifications, *Proceedings of International Conference on Artificial Life and Robotics*, ALife Robotics, Oita, Japan, pp.157-160, Jan. 2019.
2. N.G. Leveson and J.L. Stolzy, Safety Analysis Using Petri Nets, *IEEE Transactions on Software Engineering*, Vol.13, No.3, IEEE, United States, pp.386-397, Mar. 1987.
3. J. Fitzgerald, P.G. Larsen, P. Mukherjee, N. Plat and M. Verhoef, *Validated Designs for Object-Oriented Systems*, Springer-Verlag London, 2005.
4. T. Takagi, S. Morimoto, Y. Ue and Y. Imai, Animated Graphics-based Training Support Method and Prototype Tool for Bug Fixing of Extended Place/Transition Nets, *Journal of Robotics, Networking and Artificial Life*, Vol.5, No.4, pp.278-282, Mar. 2019.
5. Y. Ue and T. Takagi, Learning Support Technique of Software Visual Modeling Using Place/Transition Nets, *Proceedings of International Conference on Artificial Life and Robotics*, ALife Robotics, Oita, Japan, pp.751-754, Jan. 2020.
6. T. Takagi and A. Usuda, A Technique for Learning Software Modeling Using Extended Place/Transition Net and Its Prototype Tool, *Journal of Robotics, Networking and Artificial Life*, Vol.9, No.1, pp.81-86, June 2022.
7. Blockly, https://developers.google.com/blockly (accessed in Dec. 12, 2022).

Authors Introduction

Mr. Akio Usuda

He received the B.S. degree from Kagawa University in 2021. He is a master's student in the Graduate School of Engineering at Kagawa University. His research interests are in software engineering, particularly software design.

Mr. Ryoichi Ishigami

He is an undergraduate student in the Faculty of Engineering and Design at Kagawa University. His research interests are in software engineering, particularly software quality control.

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.

Proposal of a Framework to Improve the Efficiency of the Implementation Step in Test Driven Development (TDD)

Takeaki Miyashita[*], 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: miyashita@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

TDD is a development methodology that brings us closer to better implementation and testing by repeating a series of steps: test design, implementation that satisfies the tests, and refactoring. This paper proposes a framework aimed at supporting the implementation steps in TDD. The proposed framework generates source code that passes tests while retaining refactoring by the developer. The prototyped framework reduced the time required for the implementation process by 94.22% and the generation time by 66.17% compared to manual work.

Keywords: TDD, Boundary Value Analysis, syntax analysis, automatic generation.

1. Introduction

Test Driven Development (TDD) is a software development methodology. In TDD, developers repeat a series of steps to get closer to better test cases and implementation: creating tests that fail the existing source code, implementing the minimum source code that will pass the tests, and refactoring the implemented source code. One of the disadvantages of TDD is that it wastes more time in some cases due to repeated test failure[1]. In recent years, many researches on automatic source code generation from UML and natural language have been reported[2][3]. However, most of them focus only on one-time generation and not on maintaining continuous refactoring.

This study proposes a framework aimed at supporting the implementation process in TDD. The proposed framework automatically generates new source code by modifying the original source code, which fails to pass tests by test code, to pass tests. In addition, refactoring to the generated source code is retained and reflected in the next development cycle.

2. Proposed Framework

In this chapter, we present the structure and behavior of the proposed framework. The framework takes as input the test code and the source code S_{old} that cannot pass the test by the test code, and outputs the source code S_{new} that can pass the test. The structure of the framework is shown in Fig. 1. The behaviors of the five processing parts shown in Fig. 1 are described as below.

2.1. *Test Code Analyzer*

Test Code Analyzer parses the given test code and extracts test case data from each test case. The test case data is defined as four elements: the names of the class and member function to be tested, the arguments given to the member functions to be tested, and the expected output of the function under test for the given arguments.

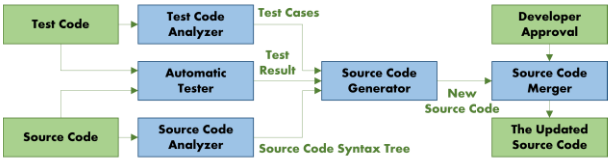


Fig. 1 The Structure of the Proposed Framework

The following is a flow of Test Code Analyzer's behavior.

(i) Test Code Analyzer parses the test code and generates an abstract syntax tree.
(ii) It searches for assertions in the generated abstract syntax tree and identifies functions that contain assertions.
(iii) For each variable in the function, create the variable table by extracting type and value. The value means either a literal value or a pair of the function and the value of the arguments given to the function. If the variable is assigned the return value of a function call, record the pair of the function and the argument values given to the function as the value of the variable in the variable table.
(iv) For values used in assertion, get the type and value from the variable table. If the value is a pair of the function and the arguments, the function shall be the function to be tested.
(v) It sends the extracted test case data to Source Code Generator (described in section 2.4).

2.2. *Source Code Analyzer*

Source Code Analyzer parses source code S_{old} and extracts existing class data. Specifically, the name and type of the argument of a member function are extracted. The class data is sent to the Source Code Generator (described in section 2.4). If source code S_{old} does not exist, Source Code Analyzer does not send data.

2.3. *Automatic Tester*

Automatic Tester automatically runs tests and extracts failed test case data from the test results. Specifically, it is the names of the test case and the member function that the test case target. The failed test case data is sent to the Source Code Generator (described in section 2.4).

2.4. *Source Code Generator*

Source Code Generator generates source code S_{gen} that satisfies the test case based on data received from Test Code Analyzer, Source Code Analyzer, and Automatic Tester. The Source Code Generator operates only if the existing source code S_{old} fails to pass the test. Otherwise, it terminates the process.

2.4.1. Integration of Input Data

Source Code Generator first extracts the test cases in the test case data received from the Test Code Analyzer that match the names of the failed test case data received from the Automatic Tester. In addition, the Source Code Generator also extracts test case data for the function that is the target of the failed test case in the same way. If the class data received from the Source Code Analyzer has data for a class that has a function targeted by the failed test case, extract this class data.

2.4.2. Estimation of Expected Output based on Boundary Value Analysis

Source Code Generator sorts test cases based on the arguments and member variables of the extracted test case data. Then, assume that the expected outputs from the arguments and member variables located between adjacent test cases from sorted test cases with the same expected outputs are equivalent. This assumption is based on the concept of boundary value analysis. Boundary value analysis is a test design method that uses test cases around boundary values. If the test cases are well designed, outputs by parameters between test cases that are adjacent on the sorted test cases and that expect the same value of output have a high probability of expecting the same value as well.

2.4.3.Source Code Generation to Satisfy Test Cases

conditional expression and a return statement with the expected output as the return value. The function intermediate data has a function name, a return type, and multiple blocks. The class intermediate data has a class name, member functions, and member variables.

Member intermediate data are generated based on the class data extracted in section 2.4.1 and the generated function. The Source Code Generator generates source code S_{gen} that satisfies the test case by recursively transforming the class intermediate data and sending it to the Source Code Merger (described in Section 2.5).

List. 1 Example of Test Code

```
#include "gtest/gtest.h"
#include "FeeCalculator.h"
namespace foolish_coder{
   TEST(CalcFeeTest, CalcFeeTestCaseYoungestChild){
      FeeCalculator fee_calculator;
      EXPECT_EQ(fee_calculator.calcFee(0), 500);}
   TEST(CalcFeeTest, CalcFeeTestCaseOldestChild){
      FeeCalculator fee_calculator;
      EXPECT_EQ(fee_calculator.calcFee(17), 500);}
   TEST(CalcFeeTest, CalcFeeTestCaseYoungestAdult){
      FeeCalculator fee_calculator;
      EXPECT_EQ(fee_calculator.calcFee(18), 800);}
   TEST(CalcFeeTest, CalcFeeTestCaseOldestAdult){
      FeeCalculator fee_calculator;
      EXPECT_EQ(fee_calculator.calcFee(60), 800);}
   TEST(CalcFeeTest, CalcFeeTestCaseYoungestOld){
      FeeCalculator fee_calculator;
      EXPECT_EQ(fee_calculator.calcFee(61), 500);}
   TEST(CalcFeeTest, CalcFeeTestCaseOldestOld){
      FeeCalculator fee_calculator;
      EXPECT_EQ(fee_calculator.calcFee(120), 500);}
   TEST(CalcFeeTest, CalcFeeTestCaseTooYoung){
      FeeCalculator fee_calculator;
      EXPECT_EQ(fee_calculator.calcFee(-1), -1);}
   TEST(CalcFeeTest, CalcFeeTestCaseTooOld){
      FeeCalculator fee_calculator;
      EXPECT_EQ(fee_calculator.calcFee(121), -1);}}
```

2.5. *Source Code Merger*

Source Code Merger merges the refactoring data from the previous cycle with the S_{gen} generated by Source Code Generator and overwrites the existing source code.

2.5.1. Integration of Refactoring Data

Source Code Merger merges the existing source code S_{old} into the source code S_{gen} generated by Source Code Generator and generates a new source code S_{new} with the refactoring data included in S_{old}.

Source Code Merger first generates the source code S_{merge}, which merges S_{gen} and S_{old}. If conflicts exist between the parts in S_{gen} changed due to the addition of test cases and the parts in S_{old} changed due to refactoring, Source Code Merger generates source code $S_{ahead\ old}$ of S_{old} and source code $S_{ahead\ gen}$ merged ahead of S_{gen}. Then, S_{merge}, $S_{ahead\ old}$, and $S_{ahead\ gen}$ are tested in this order, and the one that can pass the test is the new source code S_{new}. If all source code cannot pass the test, notify the developer that it cannot be generated source code and terminate the process.

List.2 The Header File for Test Code is shown in List. 1

```
#ifndef FEE_CALCULATOR_H
#define FEE_CALCULATOR_H
class FeeCalculator{
  public:
              int calcFee(int param1);
};
#endif
```

List. 3 The Source Code File for Test Code is shown in List. 1

```
#include "FeeCalculator.h"
int FeeCalculator::calcFee(int param1){
  if(param1 <= -1 || 121 <= param1){
    return -1;
  } else if((0 <= param1 && param1 <= 17) || (61 <= param1
&& param1 <= 120)){
    return 500;
  } else if((18 <= param1 && param1 <= 60)){
    return 800;
  }}
```

2.5.2. Review and Refactoring Requests

Source Code Merger requires a developer to review and refactor S_{new}. After receiving approval from the developer, Source Code Merger overwrites the existing source code with the refactored S_{new}.

3. Application Example

We have prototyped the framework and generated C++ source code from the test code. It has been implemented in Python, with Antlr (ANother Tool for Language Recognition) used for parsing and Google C++ Testing Framework used for testing. The output source code is generated in two files: a header file and a source code file.

An example of the test code is shown in List. 1. The generated source code for the test code shown in List. 1 is shown in List.2 and List. 3. The header file is shown in List.2. The source code file is shown in List. 3. The test code is for the member function getFee() of the FeeCalculator class. The function getFee() takes the user's age as an integer argument and returns 500 for ages 0-17, 800 for ages 18-59, 500 for ages 60-120, and -1 for all other values.

The generated source code passed all test cases. Also, List. 3 shows that the range of conditionals generated for the example is reasonable. In addition, the following refactoring was done to the source code shown in List. 1.

- Delete the conditional expression that returns -1 and add "return -1;" at the end of the function.
- Change argument name to age.

List. 4 Extension with Retained Refactoring

```
#include "FeeCalculator.h"
int FeeCalculator::calcFee(int age){
  if((0 <= age && age <= 17)){
    return 0;
  } else if((18 <= age && age <= 60)){
    return 800;
  } else if((61 <= age && age <= 120)){
    return 500;
  }
  return -1;
}
```

Table 1 Time per Test Case

Method	Implementation Time	Refactoring Time	Total Time
Manual	3m28s	29s	4m45s
The Proposed Framework	12s	45s	1m36s

Then, the test case was extended to return 0 for 0-17 years old, and the source code was generated again. The generated source code is shown in List. 4. This source code passed all test cases after the extension. List. 4 shows that the refactoring could be retained and the program extended.

4. Discussion

The time required for the implementation process in TDD was compared between manual and using the proposed framework. The results are shown in Table 1.

A total of six subjects participated in the experiment: four graduate students and two fourth-year undergraduate students. They will use TDD to solve two tasks. Half of them used the framework only for the first of two tasks, while the others used the framework only for the second of two tasks.

The following is a description of the tasks.

(i) Member function calcFee() of the FeeCalculator class: Member function calcFee() of the FeeCalculator class: Function that takes the user's age as an integer argument and returns 100 for ages 0~17, 500 for ages 18~60, 200 for ages 61~120, and -1 for all other values

(ii) Member function getLastDayMonth() of the DatePicker class: Receives the month as an integer argument and returns 31 for January, March, May, July, August, October, and December, 30 for April, June, September, and November, 28 for February, and 0 for all other values.

The experimental flow is shown below.

(i) Make one test case that cannot be passed by the existing source code.

(ii) Implement source code that can pass testing, either manually or using the framework.

(iii) Review generated code and refactor as needed.

(iv) Repeat (i)-(iii) until the subject determines that task is complete.

For the average time per test case, the time taken for (ii) is the implementation time, the time taken for (iii) is the refactoring time and the time taken for (i)-(iii) is the total time. Table 1 shows that the proposed framework reduced the implementation time in TDD by 3m16s (94.22%). Refactoring time was 16s longer. The cause is the time taken to review the generated source code. Although, the total time was reduced by 3m9s (66.17%). In addition, the proposed framework retained the refactoring for the argument names. Therefore, the proposed framework is useful to improve the efficiency of the TDD implementation step.

5. Conclusion

We have proposed a framework to improve the efficiency of the implementation process in TDD. In the applied example, the prototype framework reduced especially the time required for the implementation process by 94.22% and the development time in TDD by 66.17%. It also retains refactoring by the developer.

Hence, the proposed framework is useful to improve the efficiency of the implementation process in TDD.

Future issues are as follows.

- Addition of supported syntax
- Improving refactoring retention

References

1. F. Anwer, S. Aftab, U. Waheed, Muhammad, S. S., “Agile Software Development Models TDD, FDD, DSDM, and Crystal Methods: A Survey”, International Journal of Multidisciplinary Sciences and Enginnering, Vol. 8, No. 2, pp. 1-10, 2017.
2. Mukhtar, M. I., Galadanci, B. S., “AUTOMATIC CODE GENERATION FROM UML DIAGRAMS: THE STATE-OF-THE-ART”, Science World Journal, Vol. 13, No. 4, pp. 47-60, 2018.
3. P. Yin, G. Neubig, “A Syntactic Neural Model for General-Purpose Code Generation”, Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, Vol. 1, pp. 440-450, 2017.

Authors Introduction

Takeaki Miyashita

Takeaki Miyashita received the Bachelor's degree in engineering (computer science and systems engineering) from the University of Miyazaki, Japan in 2022. He is currently a Master's student in Graduate School of Engineering at the University of Miyazaki, Japan. His research interests software testing and software quality.

Tetsuro Katayama

Tetsuro Katayama 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

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.

Hisaaki Yamaba

Hisaaki Yamaba 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

Kentaro Aburada 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

Naonobu Okazaki 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.

Continuance Intention Factor of Online Learning Management System in Case on Faculty of Computer Science at Brawijaya University in Indonesia

Mochamad Chandra Saputra[1], Tetsuro Katayama[2], Yoshihiro Kita[3], Hisaaki Yamaba[2], Kentaro Aburada[2], and Naonobu Okazaki[2]
[1]Faculty of Computer Science, Brawijaya University, Malang-Indonesia,
[2]University of Miyazaki, Japan, [3]University of Nagasaki, Japan
andra@ub.ac.id, 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

The outbreak of the COVID-19 pandemic changed the model of the learning process. Online learning became one of the best solutions for many educational institutions, including the University, especially the Faculty of Computer Science, Brawijaya University. The continuance intention factor for using an online learning management system is important to ensure learning sustainability. To understand students' continuance intention this study proposes self determination model as a research model to find the factors affecting students' continuance intention toward online learning management systems. The proposed model is useful for investigating continuance intention factors. The study finds that intrinsic regulation, external regulation, identified regulation, and introjected regulation significantly positive impact on continuance intention in learning management systems. Finally, this study provides suggestions for the Faculty to improve the continuance intention of the student in using an online learning management system.

*Keywords:*Learning Management System, Self Determination, Continuance Intention

1. Introduction

Early in the Covid-19 pandemic, the disease's spread across many nations is unpredictable, and thus university educational programs should keep going. The covid-19 pandemic changes the model of the learning process from traditional to online. This condition enforces that the universities and all members of them are adaptive to improving the learning process. The students and lecturer are forced to learn online with limited facilities and strongly different learning processes from previous. In this situation, the capability for student and lecturer adaption is necessary with many problems such as experiencing difficulties to manage their resources and engaging in self-regulated learning, different profiles of adaptation emerged: the overwhelmed, the surrenderers, the maintainers, and the adapters[1].

Learning Management Systems, referred to as LMS, is a platform that assists in the delivery of content online for learning purposes[2]. LMS is used in both public and private educational institutions to implement learner-centered instruction as well as to facilitate innovation processes. Using an online LMS to achieve higher education institution aims has been shown to enhance the efficacy and efficiency of facilitating student learning. Support from the university improves the relationship between the perceived impact of COVID-19 on degree completion and the future employment prospects of the students[3]. Faculty of Computer Science, the Brawjaya University implementing an online learning model on the learning process. During the Covid-19 Pandemic, the Brawijaya University, especially the Faculty of Computer science, uses online LMS to improve its learning process named ELING. ELING provides virtual classrooms that focus on quality online learning at the Faculty of Computer Science, Brwaijaya University. Until now, many lecturers and students are continuing to use the LMS for their daily courses.

Nowadays, the study of the human motivation of using technology is an interesting field of study. A study on human motivation investigates how humans feel more motivated to take action in the case of how they think and affect the outcome. Self Determination Theory is one of the methods that study human motivation and personality that represents several viewpoints of factors or variables from internal and external humans[4].

This paper focuses on investigating the contribution of the variables on self determination model to continuance intention for the learning management system during the pandemic covid 19 and after in the case of the Faculty of

Computer Science, Brawijaya University, Malang-Indonesia. The finding of this study suggests a deeper understanding of Self Determination Factor that improves the policy-makers to better improve learning management system services.

2. Literature Review

2.1 Self Determination Theory

Self Determination Theory (SDT) is a theory of motivation that is well established[4]. SDT has two types of motivation such as intrinsic motivation and extrinsic motivation. Intrinsic motivation is related to doing something for its own purpose and extrinsic motivation refers to doing something for an external effect on the activity itself.

Self determination theory consists of four factors such as Intrinsic Regulation, Identified Regulation, Introjected Regulation, and External Regulation. Intrinsic Regulation defines as the activity with satisfaction that is motivated because of external prods, pressures, or rewards. Identified Regulation is defined as acceptance of regulation as being one's own. Introjection regulation remains as controlling people on perform such actions under pressure to avoid guilt or anxiety or to achieve ego enhancements or pride. Introjected regulations are distinguished by the individual's internalization of external regulations. External Regulation is defined as behaviors influenced by external conditions such as compliance, threat, punishment, or external rewards, etc.

2.2 Continuance Intention

Continuance intention is defined as a user's decision to continue using an Information Technology (IT) that they have previously used. Continuance intention is to explain people's decisions and motivation to use information technology in daily activities that help people solve their problems. Continuance intention is determined by user satisfaction and perceived usefulness to continue using the system[5]. User satisfaction related to information quality, system quality, service quality, perceived usefulness, perceived ease of use, and communication quality toward the learning management system[6]. Perceived usefulness is one of the main determinants of faculty members' behavioral intention to use a Learning Management System[7]. In the case of the learning management system in this research, continuance intention is important to improve the learning process. The continuance intention in this study refers to the behavioral motivations and attitudes of students to engage in the online learning management system.

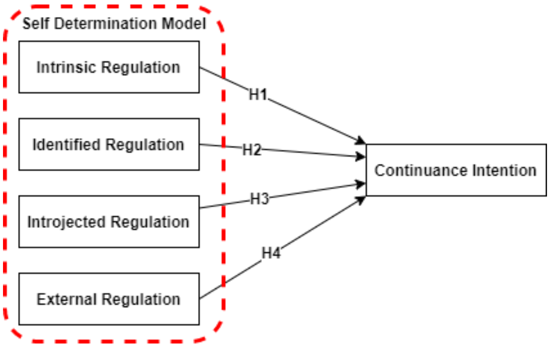


Fig. 1 The Research Hypothesis Model

3. Research Model and Hypothesis

Fig. 1Fig. 1 shows the research hypothesis model in this research. Current experiments focused on self determination models with continuance intention. The research model explains the result of investigating the factors from self determination model that have a strong relationship or impact on continuance intention. Investigating the factor by considering the hypothesis accepted or rejected.

This research drives the following hypotheses and every hypothesis will measure by the questionnaire that should answer by the student.

1. H1: The Intrinsic Regulation of the Learning Management System has a positive influence on user continuance intention of using the Learning Management System.
2. H2: The Identified Regulation of the Learning Management System has a positive influence on user continuance intention of using the Learning Management System.
3. H3: The Introjected Regulation of the Learning Management System has a positive influence on user continuance intention of using the Learning Management System.
4. H4: The External Regulation of the Learning Management System has a positive influence on user continuance intention of using the Learning Management System.

4. Methodology

4.1 Instrument for Data Collection

Collecting the data on this study uses a questionnaire on self determination theory related to the research on self determination and continuance intention [8]. The questionnaire gets the data by requesting the respondents to answer a set of questions. The questionnaire collected the demography of the respondent such as age, gender, department, and year of study. The next information on the questionnaire is identified the learning process such as online, offline, or hybrid, and the application for online learning.

The respondent of the study are students in the Faculty of Computer Science at the Brawijaya University. Students answer the questionnaire with choose the Likert scale with a maximum score of five (strongly disagree to strongly agree). Intrinsic Regulation, Identified Regulation, Introjected Regulation, External Regulation, and Continuance Intention are among the five variables included in the questionnaire that is related to the hypothesis model.

4.2 Data Collection and Analysis

Data collected focus on the student who has used online learning in the learning process during the Covid-19 Pandemic. The questionnaire distributes online using Google Forms. The number of respondents is 423 which 23 respondents are answered offline learning process. The number of respondents uses on this 400, respondent which answer offline learning process was not included. Data cleaning for the questionnaire deleted the response which not use online learning and the data analysis uses the statistical approach.

The research of self-determination and continuance intention on learning management systems which type of research in the field of social science tries to investigate and explain the social condition by using quantitative methods which use statistical analysis techniques to evaluate data collected with questionnaires.

5. Data Analysis

The demography of respondent on this survey consist of 65.7% male and 34.3% female. Respondents in this study also had an age range from 18 years to 25 years from various study programs at the Faculty of Computer Science, Brawijaya University. This research does not perform the validity and reliability test because the questionnaire that using is standardized for self determination theory. This research uses normality test and homogeneity test to validate the sample of the respondents.

The Kolmogorov-Smirnov test produces test statistics that are used to test for normality[9]. The Kolmogorov-Smirnov test is used to test the null hypothesis that a set of data comes from a Normal distribution. The result of the normality test by using one sample Kolmogorov Smirnov shows that the score of significance is 0.507 that larger than 0.05 which means passed the normality test.

Table 1. Result of Regression Measurement

Hypotheses	Beta	Interpretation
Intrinsic Regulation to Continuance Intention	0.29	Positive Contribution (Hypothesis Accepted)
Identified Regulation to Continuance Intention	0.15	Positive Contribution (Hypothesis Accepted)
Introjected Regulation to Continuance Intention	0.06	Positive Contribution (Hypothesis Accepted)
External Regulation to Continuance Intention	0.26	Positive Contribution (Hypothesis Accepted)

6. Result and Discussion

Table 1 shows the result of statistical analysis using regression approach. By using the research model hypothesis, the research successfully investigated the contribution of self determination factor to continuance intention of the learning management system at the Faculty of Computer Science, Brawijaya University. The result shows that all hypothesis is accepted. The value of the contribution of the factor is related to the beta value. These findings suggested that a deeper understanding of intrinsic regulation, external regulation, identified regulation, and Introjected regulation will help policymakers to better design learning management system services. All percentage impact from self determination with continuance intention is lower than 50% and the interpretation such as follows.

H1: The intrinsic regulation of the Learning Management System has a positive influence on user continuance intention of using the Learning Management System. This hypothesis is accepted with 29% of the impact on continuance intention. To improve the percentage of intrinsic regulation the organization needs to focus to improve the service learning management system aspects such as user pleasure, satisfaction, and personal interest to continue using it.

H2: The identified regulation of the Learning Management System has a positive influence on user continuance intention of using the Learning Management System. This hypothesis is accepted with 15% of the impact on continuance intention. To improve the percentage of identified regulation the organization needs to ensure that the service learning management system improve the individual psychological feelings for acknowledgment and being owned to continue using it.

H3: The Introjected regulation of the Learning Management System has a positive influence on user continuance intention of using the Learning Management System. This hypothesis is accepted with 6% of the impact on continuance intention. To improve the percentage of introjected regulation the organization needs to focus on the service learning management system aspects such as user confidence, feeling of worth,

H4: The external regulation of the Learning Management System has a positive influence on user continuance intention of using the Learning Management System. This hypothesis is accepted with 26% of the impact on continuance intention. To improve the percentage of external regulation the organization needs to collaborate with external entities such as Student Organizations to inform the benefit of continue using the learning management system.

7. Conclusion

This research confirms the proposed self determination model successfully investigated the contribution of self determination factors to the continuance intention of the learning management system at the Faculty of Computer Science, Brawijaya University. The result shows that all hypothesis is accepted. All percentage impact from self determination factor to continuance intention is still lower than 50% which means that need to improve the service to increase the motivation of continuance intention on learning management systems.

Future research may implement and evaluate the suggestion for each self determination factor to know the improvement results.

References

1. Biwer, F. et al. “Changes and Adaptations: How University Students Self-Regulate Their Online Learning During the COVID-19 Pandemic”. Front. Psychol. 12, 2021, pp. 1–12.
2. Barreto, D., Rottmann, A. & Rabidoux, S. “Learning Management Systems: Choosing the Right Path for Your Organization”. (Ed Tech Books.org, 2020).
3. Plakhotnik, M. S. et al. “The Perceived Impact of COVID-19 on Student Well-Being and the Mediating Role of the University Support: Evidence From France, Germany, Russia, and the UK. Front”. Psychol. 12,2021 , pp. 1–13.
4. Ryan, R. M. & Deci, E. L. “Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being”. Am. Psychol. 55, 2000 , pp. 68–78.
5. Bhattacherjee, A. “Understanding Information Systems Continuance: An Expectation-Confirmation Model”. MIS Q. 25, 2001 , pp. 351.
6. Ohliati, J. & Abbas, B. S. “Measuring Students Satisfaction in Using Learning Management System”. Int. J. Emerg. Technol. Learn. 14, 2019, pp. 180.
7. Lavidas, K., Komis, V. & Achriani, A. “Explaining Faculty Members’ Behavioral Intention to Use Learning Management Systems”. J. Comput. Educ. 9, 2022, pp. 707–725.
8. Rahi, S. & Abd.Ghani, M. “Integration of DeLone and McLean and self-determination theory in internet banking continuance intention context”. Int. J. Account. Inf. Manag. 27, 2019, pp. 512–528.
9. Massey, F. J. “The Kolmogorov-Smirnov Test for Goodness of Fit”. J. Am. Stat. Assoc. 46, 1951, pp. 68.

Authors Introduction

Mochamad Chandra Saputra

Mochamad Chandra Saputra received the Master's Degree from the University of Miyazaki, Japan, and Brawijaya University, Indonesia on Double Degree Program 0n 2014. In 2021 received the Ph.D. degree from the University of Miyazaki. Since 2015 has been a lecturer on the Faculty of Computer Science, Brawijaya University. The research interest includes software testing, software quality, and software project management.

Tetsuro Katayama

Tetsuro Katayama received the 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 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

Yoshihiro Kita received a PhD 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.

Hisaaki Yamaba

Hisaaki Yamaba 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

Kentaro Aburada 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 network and security. He is a member of IPSJ and IEICE.

Naonobu Okazaki

Naonobu Okazaki 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.

Design of Intelligent Crutch System Based on STM32 and Raspberry Pie

Zongxuan Zhang, Jianhao Hu, Yizhe Sun, Yizhun Peng*
College of Electronic Information and Automation,Tianjin University of Science and Technology, Tianjin, China

*E-mail: *pengyizhun@tust.edu.cn*

www.tust.edu.cn

Abstract

In order to strengthen the development of China's elderly care industry, the project has designed an intelligent crutch based on stm32 and raspberry pie. This product adds many functions on the basis of general intelligent crutches. It is equipped with real-time GPS monitoring and uploading APP. When the elderly fall, they can immediately and automatically alarm. When they fall, they can simultaneously emit an alarm sound of about 100 decibels to the surrounding. It provides a touch screen, supports voice recognition, and can achieve multiple functions. It is equipped with a step counting function, temperature and humidity display and other functions. The product uses 4G modules to communicate with the APP.

Keywords: Elderly care industry; Intelligent crutch; STM32; Raspberry pie; APP

1. Introduction

At present, the trend of aging in China is getting worse. In order to protect the due rights and interests of the elderly in China and comply with the development of AI, an intelligent crutch with AI elements has been designed. At present, the society is developing rapidly. Young people go out and leave the elderly at home alone, unable to contact the elderly in time. If the elderly go out and have an accident, their children cannot get the news at the first time, which makes the elderly accidents happen frequently. The appearance of this product has ingeniously solved this problem. First of all, this product is equipped with GPS technology for real-time monitoring and uploading to the supporting APP[1]. There is no need to worry about the elderly getting lost when they go out. Second, it is equipped with a one button alarm: after the crutch falls to the ground, it will send an alarm at about 100 decibels and upload it to the APP at the same time to remind users[2]. Many elderly people can't adapt to smart phones. This product can be used to call their children by voice one button, making communication between children and elderly people more convenient and fast[3]. This product also provides a touch screen and speech recognition[1,4], which can display time and current position, and can play music through the touch screen. You can also listen to plays on your daily trip[5]. Many elderly people have inconvenient legs and feet and are not suitable for too much walking and physical exercise. Through the step counting function of this product, they can get the exercise time of the elderly in real time. When the exercise reaches the standard, they can be prompted to have a proper rest. The steps of the elderly and the ambient temperature and humidity will be transmitted to the screen in real time. Finally, this product is equipped with video camera and flashlight functions, so that users can fully monitor the elderly to ensure that their rights and interests are not violated, so that the fraudsters can not escape, and night travel can also be guaranteed [6-8]. Fig.1 is a real product diagram, and Fig.2 is a 3D modeling diagram.

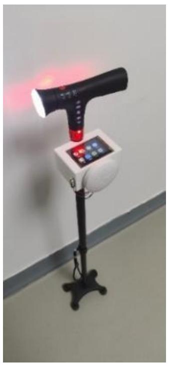


Fig.1 Real product diagram

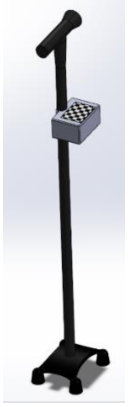


Fig.2 3D modeling diagram

2. The System Design

2.1. Embedded system design

First of all, the sensor data collection part is dominated by stm32. We distribute the controller in the upper part of the crutch and the sensors around the controller. When the sensor data collection is completed, the result is sent to stm32. After stm32 completes receiving and processing, it is sent to the 4B end of Raspberry Pie again. Raspberry Pie 4B is responsible for the output and interaction part as well as the communication with the APP. On the LCD resistance touch screen, we carry a self-designed UI interface to achieve the interaction design of various functions, supplemented by a voice control system. Through voice recognition, we can quickly retrieve all functional modules and find satisfactory choices. Fig.3 is the design framework.

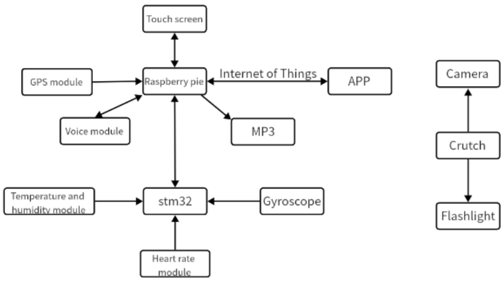


Fig.3 The design framework

2.2. APP system design

APP application is designed on Android system, which is mainly divided into three interfaces. The main interface is the status interface, with step counting display, heart rate and blood pressure display, temperature and humidity display, connection status display, and music switching function. The second interface is the alarm communication interface, which has the functions of receiving alarm, dialing with one button, and manual shutdown. The third interface is the GPS system, which calls the Gaode navigation API to monitor the crutch positioning information in real time. The data is transmitted from Raspberry Pie to the APP through 5G network, so that the APP can be updated in real time. Fig.4 is an APP display.

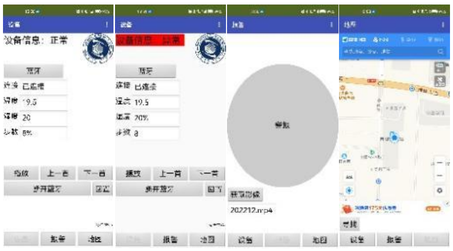


Fig.4 APP display

3. The Hardware Design

The hardware system is roughly composed of raspberry pie 4B, stm32f103, sensor, 3.5 inch LCD resistance touch screen, 5G module, MP3 module and crutch body. The raspberry pie 4B and stm32f103 are used as the core hardware system, in which stm32 receives and processes signals from various sensors, such as temperature and humidity sensors, gyroscopes, acceleration sensors, and sends them to raspberry pie. The raspberry pie part is connected to the LCD resistance touch screen, 5G module, MP3 module, and stm32 sends data to the raspberry pie through the serial port, then the raspberry pie is used as a transit, and then sent to the client.

Camera: load a small camera under the main control part of the crutch, automatically store the captured image to the SD card, automatically clean it every day, and upload it to the APP for backup.

Light: A small flashlight and manual switch are installed at the front end of the crutch grip.

Alarm: It is set on the upper end of the cane, which can be triggered automatically or controlled by manually pressing the switch.

4. The Software Design

4.1. stm32

The main function of the program is to read and process the data of each sensor, and communicate with the serial port of Raspberry Pie. The stm32 uses the query method to read the relevant sensor data. However, considering the priority of the fall alarm, we need to use the interrupt processing for the gyroscope module to ensure that it takes effect at the first time. After data collection, send the data to Raspberry Pie to complete a cycle structure.

Step counting function: through the data acquisition of the gyroscope and accelerometer, the stm32 completes the data analysis. When the gyroscope senses the vibration amplitude higher than the threshold value, it sends a decision command. At the same time, when the accelerometer senses the change of the center of gravity, it also sends a decision command. When both commands are true, the decision is successful, and the steps are increased by one. Otherwise, it skips the command and the program moves forward. Fig.5 is the principle of step counting.

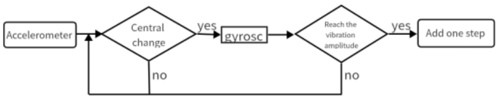

Accelerometer
Central change
yes
no
gyrosc
Reach the vibration amplitude
yes
no
Add one step

Fig.5 The principle of step counting

Temperature and humidity sensor: after the temperature and humidity module collects data, it is converted into Celsius and relative humidity units through ADC A/D conversion, and collected by stm32. Fig.6 is the schematic diagram of temperature and humidity module.

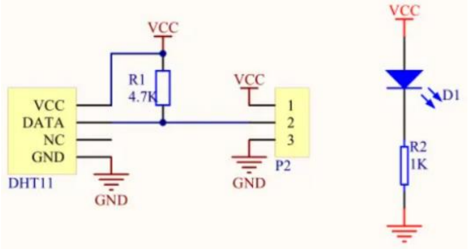

VCC
R1
4.7K
VCC
DATA
NC
GND
DHT11
GND
VCC
1
2
3
P2
GND
VCC
D1
R2
1K

Fig.6 schematic diagram of temperature and humidity module

GPS positioning: GPS module can realize satellite positioning, receive the position information from the satellite, upload the information to stm32, and use stm32 to process the information and send the information. Fig.7 is the schematic diagram of GPS.

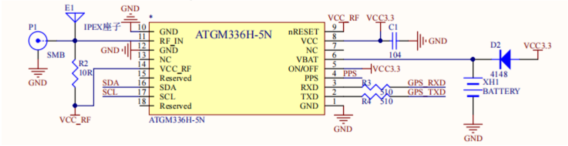

E1
GND
P1
IPEX座子
SMB
GND
R2
10R
VCC_RF
GND
SDA
SCL
10 GND
11 RF_IN
12 GND
13 NC
14 VCC_RF
15 Reserved
16 SDA
17 SCL
18 Reserved
ATGM336H-5N
nRESET 9
VCC 8
NC 7
VBAT 6
ON/OFF 5
PPS 4
RXD 3
TXD 2
GND 1
ATGM336H-5N
VCC_RF
VCC3.3
C1
104
GND
PPS
VCC3.3
R3
510
R4
510
GPS_RXD
GPS_TXD
GND
D2
4148
VCC3.3
XH1
BATTERY
GND

Fig.7 The schematic diagram of GPS.

Data exchange: STM32 and raspberry pie exchange data through serial port. STM32 and raspberry pie cross connect TXD and RXD ports. STM32 sends the data of all sensors to raspberry pie by serial transmission, allowing raspberry pie to complete data collection. Fig.8 is the principle of data communication.

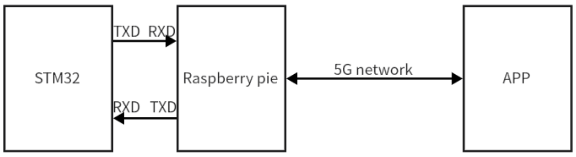

STM32
TXD RXD
RXD TXD
Raspberry pie
5G network
APP

Fig.8 the principle of data communication

Fall alarm: When the amplitude of vibration collected by the gyroscope exceeds a certain threshold or the data collected by the accelerometer exceeds a certain threshold, the alarm will immediately send an alarm, and the signal will be sent to the Raspberry Dispatch controller using an external interrupt. Fig.9 is the alarm principle.

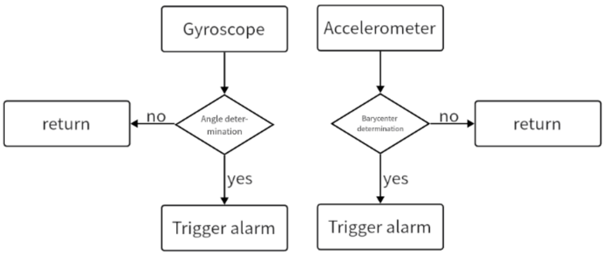

Gyroscope
Accelerometer
return
no
Angle determination
Barycenter determination
no
return
yes
yes
Trigger alarm
Trigger alarm

Fig.9 the alarm principle

4.2. Raspberry pie

The program mainly completes data receiving and sending, voice recognition, and smart screen functions. Raspberry Pie communicates with stm32 through serial port, and communicates with APP through 5G network. With 5G network, communication function and alarm signal transmission function can be realized in addition to data transmission. With the touch screen, GPS, time, MP3 and alarm call functions will be freely selected and interacted. The voice recognition module shall also be equipped. After the voice mode is opened on the screen, the above functions can be called by using voice.

Music playing: After connecting USB module and mp3 module, Raspberry Pie can read the mp3 format file of the U disk, and use the touch screen to complete music demand, volume and pause functions. The external speaker can realize voice output, music playing and alarm. Fig.10 is the MP3 module PCB diagram.

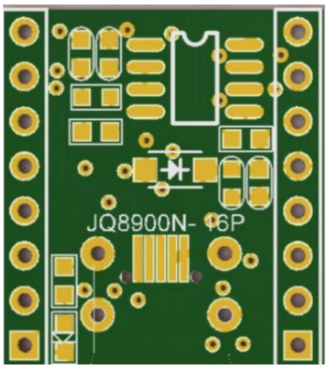


Fig.10 the MP3 module PCB diagram

Network: raspberry pie uses the 5G wireless network communication module to connect to the local network base station, realize the Internet of Things function, and complete some functions and remote communication with the APP. Fig.11 is a 5G module.

Fig.11 5G module

Screen and voice recognition: Connect to Raspberry Pie with the touch screen, draw and import UI icon by computer, and realize UI interaction function by Raspberry Pie. After the voice module is connected, the recognition command will be imported to the module in advance. After debugging to the best performance, voice control can replace the operation of the smart screen. Fig.12 is the schematic diagram of the voice module.

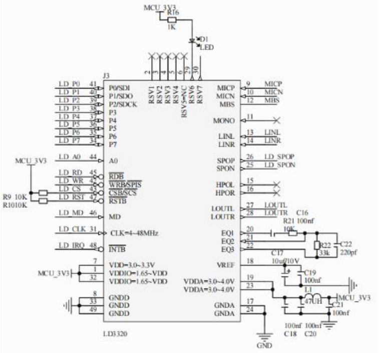


Fig.12 the schematic diagram of the voice module

4.3.APP

The APP program is mainly composed of three interfaces. The first is the device information part, which consists of the connection status of the device, the control of MP3 music, the temperature and humidity, steps, heart rate and blood pressure of the sensor. The sensor data comes from the collection and processing of stm32. After the cane falls, there will be color changes to indicate warnings. The second part is the alarm part, which can control the opening and closing of the alarm of the crutch, and the lower part is also equipped with a call answering module. The third part is GPS positioning, which is connected to the navigation API of Gaode, and the fast navigation function is set. The GPS module transmits the longitude and latitude information collected to the APP through the raspberry pie, and then the APP searches for the position of the crutch according to the data.

5. Conclusion

After testing, the product can operate normally without obvious problems in each module.

References

1.Li Xie, Zheng Zhilan, Lin Yuanqiu ,et al.,Intelligent multi-function crutches based on NB IoT and UWB. *Network Security Technology and Application*, 2022 (08): 43-45

2.Liu Yifei, Meng Shihui, Lv Peihua, et al.,Design of a multi-functional intelligent crutch system. *China Science and Technology Information*, 2022 (06): 69-71

3.Zheng Chengwei, Ma Liang, Wang Xing, et al., Design and Implementation of Intelligent Crutches for the Elderly. *Computer Knowledge and Technology*, 2021,17 (20): 137-138. DOI: 10.14004/j.cnki.ckt.2021.2000

4.Liu Juan, Huang Zhong, Guo Yu, et al.,Design of intelligent crutches for the elderly based on AT89S52 microcontroller. *Light Industry Science and Technology*, 2021,37 (02): 59-60

5.Mao Lifeng, Lin Junhao, Yan Hengyang, et al., Design of intelligent crutches based on the elderly. *Electronic Production*, 2021 (01): 31-33+8. DOI: 10.16589/j.cnki.cn11-3571/tn.2021.01.010

6.Sun Zhengyang, Qu Zhifeng, Wan Hongqiang et al., Innovative design of multi-functional intelligent crutches. *Electronic Testing*, 2020 (14): 17-18. DOI: 10.16520/j.cnki.1000-8519.2020.14.006

7.Wu Jin, Fu Yafei. Intelligent crutch based on STM32.*China Integrated Circuit*, 2020,29 (Z3): 77-80

8.Qu Pengyu, Li Yiqiang, Cao Jingqi, et al.,Scheme design of multi-functional full-automatic intelligent crutches. *Hubei Agricultural Mechanization*, 2020 (03): 182

Authors Introduction

Mr.Zongxuan Zhang

He is currently an undergraduate of Tianjin University of Science and Technology. His research field is software design.

Mr.Jiahao Hu

He is currently an undergraduate of Tianjin University of Science and Technology. His research field is embedded system.

Mr.Yizhe Sun

He is currently an undergraduate of Tianjin University of Science and Technology. His research field is embedded system.

Dr. Yizhun Peng

He is an Associate Professor in Tianjin University of Science &Technology. He recevived 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 Intelligent Fish Box Based on Machine Vision and Internet of Things Technology

Suqing Duan, Jiangyu Wu, Shuai Chen, 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 fish box based on machine vision and Internet of Things includes many fields such as machine vision, Internet of Things technology, single-chip microcomputer control and so on. It is an intelligent system that receives user instructions or voice control through the Internet to achieve a series of operations. At the same time, there is a temperature sensor in the fish box, which can transmit the temperature to the mobile phone APP in real time, and intelligent control the water temperature in the fish box. In addition, the fish box granaries send information to the user when the fish food is insufficient, and the camera in the fish box can monitor the situation of the fish box in real time, when a fish belly notifies the user to deal with.

Keywords: Raspberry Pi, machine vision, neural network, Internet of things, machine learning

1. Introduction

At the request of China's Ministry of Industry and Information Technology, many places across the country should be based on the positioning of "seeking reform, promoting application and guiding experience", further strengthening application orientation, accelerating the application of new technologies and products such as artificial intelligence, fostering the formation of new economic growth points and exploring more new models, and now the demand for intelligent fish feeding equipment is increasing. The huge demand for intelligent fish boxes in society has also prompted the development and manufacture of intelligent aquarium controllers of various types and applications within manufacturers in different countries around the world.[1] As a result, some controllers with simple circuits, stable operation, and automatic intermittent operation have appeared on the market. However, since different devices operate independently and each has its own environmental parameters, several independent controllers must be installed in the aquarium. This individual control device is relatively expensive, so the research and development of multi-purpose fish tank controller has been increasingly important to major manufacturers, and a variety of multi-functional controllers have been introduced.[2]

The rest of this paper is organized as follows. The second section carries out the analysis of smart vision fish box functions. The third section carries out the design of the intelligent visual fish box controller. The fourth section carries out the design of the smart vision fish box sensor, including both peripheral sensors and actuators. The fifth part is the software design, which illustrates the control algorithm of each part of the fish box and introduces the neural network vision algorithm to identify the health condition of fish. The sixth part is the system test, which carries out the debugging and preparation of each part. The seventh part summarizes the main contents of this paper.

2. Analysis Of Intelligent Visual Fish Box Functions

2.1. Overall system scheme design

The system is composed of Arduino development board as the main control board, Raspberry Pi as the visual recognition part, plus peripheral circuits. The important components are mechanical structure, Raspberry Pi, camera, main control board, switch, relay, temperature sensor, WIFI module, voice control system, various motors, etc.

The system is controlled by the cell phone side of the APP network to send signals to the microcontroller, or voice control part and the main control board for serial communication. The mechanical part models the automatic feeding part with the electric control box, and then connects directly with the box.

The system adopts 5V for control electricity, and it can work normally by accessing 220v household socket. The fish box can be automatically connected to the Internet after the power is turned on, and the water pump, oxygen pump, and lights can be switched on and off through the cell phone; and the data collected by the temperature sensor can be transmitted to the cell phone, and the historical water temperature change curve can be viewed at the cell phone. In addition, the voice control part recognizes the sound collected by the radio device and sends the corresponding hexadecimal data to the main control board, which can realize voice control. The fish tank is left with a 4-day fish food bin, so users can open the APP for feeding anywhere; after the fish food bin is empty, the fish tank sends a signal to the user to remind the user to add fish food and reset the food bin to zero position. In addition, the camera can monitor the fish health in real time and send a notification to the user when there is a fish death, so that the user can clean up in time to prevent further pollution of the water body.[2]

2.2. Introduction of external structure

The electric control box body is built with white PVC foam board, which is more sturdy. The circuit wiring is placed inside to ensure the normal function of the system. The design of the electric control box is shown in Fig1.

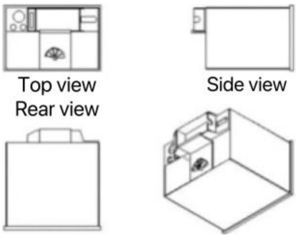


Fig.1 electric control box

The feeding mechanism is obtained by 3D printing instead. The design of the printing feeding mechanism is shown in Fig2.

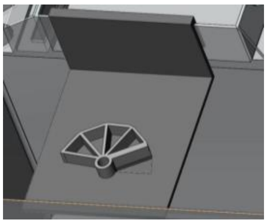


Fig.2 printing feeding mechanism

After each part is prepared, the fish box is built. The design of the smart fish box effect is shown in Fig3.

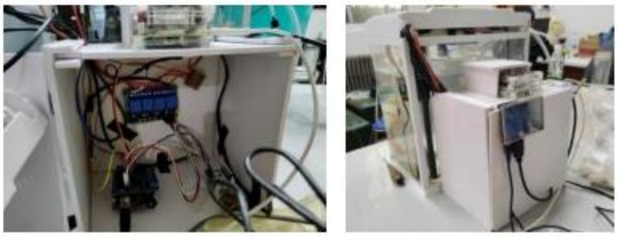


Fig.3 smart fish box effect

3. Design Of Intelligent Visual Fish Box Controller

3.1. Main control board

Arduino UNO uses AtmelAtmega328 chip, which can develop projects easily and quickly. It is an open source controller, which can be programmed, compiled and burned through IDE and USB cable without additional programming tools. The main control board has a total of 0~13 digital inputs/outputs. A total of 0~5 analog inputs/outputs. Internet service provider download is provided. USB power supply is available, as well as external power supply from 5V to 9V. This type of development board can be used as the main control board to realize the control requirements of this smart fish box with good stability and high efficiency, which can reduce the development cost of this system design and improve the efficiency of program development. The PCB design is carried out in the software. The design of the arduino schematic diagram completed using Altium Designer is shown in Fig4.

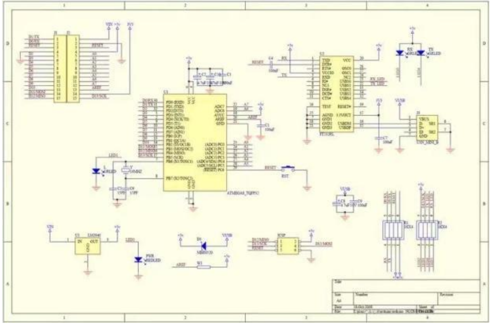


Fig.4 The arduino schematic diagram completed using Altium Designer

3.2. Vision part and camera

(1) Raspberry Pi 4B

This design uses the ARM-based microcomputer motherboard Raspberry Pi 4B to implement the vision part of the smart fish box algorithm, the Raspberry Pi can access devices such as mouse, network cable and keyboard to operate the Raspberry Pi for programming and other work, and can output high-definition video through the HDMI interface. Its memory hard drive uses a MicroSD/SD card. There is also an Ethernet port on the motherboard, and several hardware ports for USB plugging in.

The Raspberry Pi4B was chosen because of its excellent performance. A 64-bit quad-core processor running at 1.5Ghz ensures that the Raspberry Pi4B has a high processor speed, as well as good memory and an excellent desktop system. The original camera can well meet the requirements of this design for the identification of ornamental fish survival or non-survival. The design of the Raspberry Pi4B motherboard is shown in Fig5.

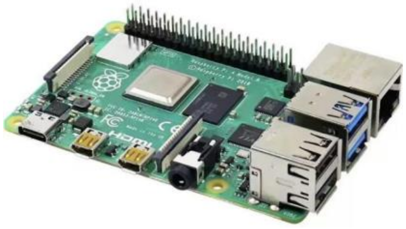


Fig.5 Raspberry Pi4B motherboard

(2) Camera

The camera in this design is the original Raspberry Pi camera with five megapixels. The camera in this scheme is connected to the CSI interface in the motherboard to obtain real-time video and detect the survival of fish in the fish box by the neural network run by the Raspberry Pi. The design of the Raspberry Pi supporting camera is shown in Fig6.

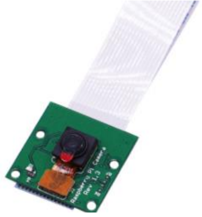


Fig.6 Raspberry Pi supporting camera

(3) Networking part

The networking part of this smart fish box is implemented by ESP8266, a low-cost and highly integrated wireless network MCU, which is mainly used for the development of IoT products. This system has stable performance, high integration, a wide range of operating temperature, stable performance during operation, and a built-in 32-bit Tensilica processor, standard digital peripherals, suitable for different working environments.

Hardware, only a small number of peripheral circuits such as interface, RF BLUN, power management module, etc., can reduce the PCB occupation.

With 32-bit Tensilica processor, the power consumption is low. ESP8266 adopts various patented technologies to achieve extremely low power consumption. Its energy-saving approach is suitable for a wide range of low-power applications.

The built-in ultra-low power TensilicaL106 32-bit RISC processor with 160 MHz CPU clock rate and support for Real Time Operating System (RTOS) and Wi-Fi protocol provides 80% of the processing power for program design and development. The design of the ESP8266-12E minimum system is shown in Fig7.

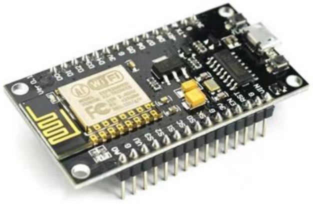


Fig.7 ESP8266-12E minimum system

4. Design of smart vision fish box sensor

4.1. Voice module

Su-03T is a speech recognition chip with small size, low power consumption and low cost, which can be easily applied to all kinds of smart products that need voice control, such as various smart homes. The design of the SU-03T application circuit schematic is shown in Fig8.

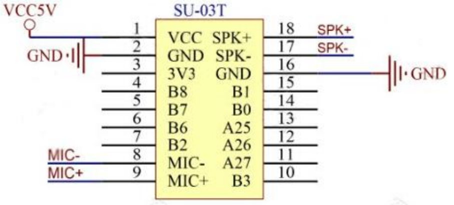


Fig.8 SU-03T application circuit schematic

The on-board microphone is used as the radio device of the voice module. It adopts electret capacitor φ with the size of 4.0*1.2mm. The working voltage is 1.1V-10V and the sensitivity is 42dB. The design of the on-board microphone is shown in Fig.9. The design of the system function block diagram is shown in Fig.10.

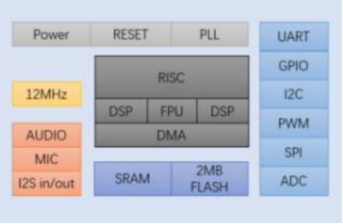


Fig.9 on-board microphone

Fig.10 system function block diagram

The module collects voice information through the microphone, analyzes the spectrum, extracts features, and compares them with the pre-set command library, and finally sends the recognition results to the on-board processor for the next step, such as serial transmission, etc. The SU-03T module can also be developed using an online platform to improve the accuracy and speed of voice recognition. Its low power consumption and high development efficiency can meet the needs of the voice control part of this design.

4.2. Temperature sensor

The design uses DS18B20 type digital temperature sensitive element as the temperature sensor, and the water temperature detection line of the probe is sealed with stainless steel and other waterproof measures.

DS18B20 is a new type of thermal digital sensing element, which ensures sensitivity to temperature, data transmission through a one-wire bus transmission method, and has a very low delay. [3] DS18B20 type digital thermometer are a unique number, and according to the number to identify the corresponding sensor. The length of the sensor wires can be increased or decreased as required by the user. In this design, water temperature detection in the fish tank is carried out using a waterproof treated DS18B20 probe. The design of the DS18B20 temperature sensor circuit schematic is shown in Fig.11.

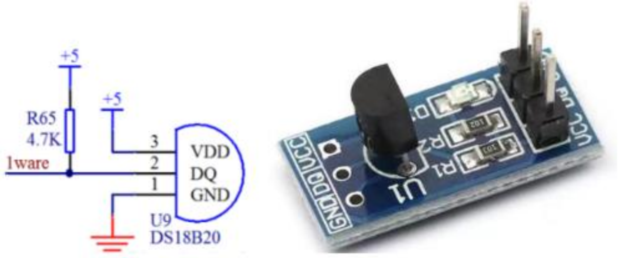


Fig.11 DS18B20 temperature sensor circuit schematic

4.3. Servo

The 180° digital servo servo DS3235 is used for the servo. The servo control panel within the digital servo used in the design of this paper is controlled by MCU, which provides one PWM pulse width and can lock the directional angle. It has high control accuracy, good linearity, and close integration with the control protocol to achieve a minimum control angle of 0.9° or less. The built-in digital circuit board, metal material gears, CNC aluminum mid-case, double ball bearings, fast heat dissipation, and higher durability.

In the application of this design, the servo is directly connected to the feeding mechanism, and the servo turns through 36° for every feeding command received by the main control board. With a 7.4V power supply, this servo can reach a torque of 35 kg*cm and takes only 0.11 seconds to turn through 60° [8]. It allows precise control of the angle, fast heat dissipation and long life, which is good for the automatic feeding function of the fish box. The design of the Servo Interface is shown in Fig.12.

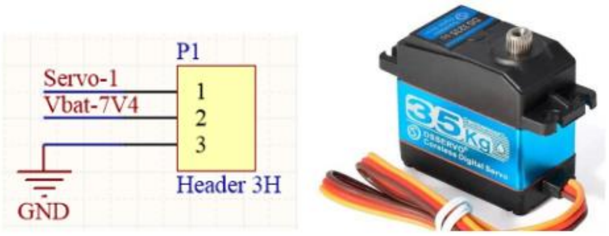


Fig.12 Servo Interface

4.4. Motor relay

This design uses JQC-3FF-S-Z relay. The common terminal and normally open contact are connected to the fire wire at one end and the appliance at the other end, and the zero wire is directly connected to the appliance. The control end is connected to 5V control power, GND and microcontroller IO. When the command is received, when the corresponding pin becomes low voltage, the normally open contact of the relay will be closed, thus turning the appliance on. The design of the JQC-3FF-S-Z relay is shown in Fig.13.

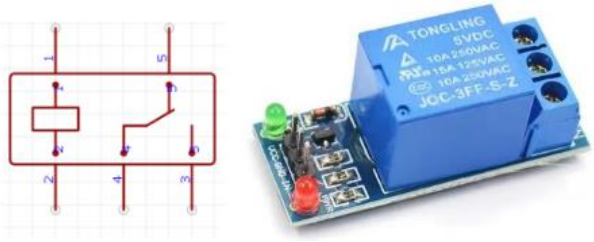


Fig.13 JQC-3FF-S-Z relay

4.5. Buzzer

This design uses a 5V buzzer 12065. the negative terminal of the buzzer shares ground with the microcontroller, and the positive terminal is connected to the IO of the main control board. when the fish food bin is empty the microcontroller will set the pin to high level, thus triggering the buzzer to remind the user of the purpose. The design of the Buzzer is shown in Fig.14.

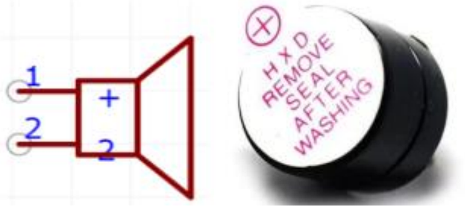


Fig.14 Buzzer

5. Software Design

5.1. Control algorithm of each part of the fish box

(1) General part

The main control board code development environment is arduino IDE, which is convenient and fast to develop and burn, and reduces the development difficulty. The overall idea of the code is that the arduino serial port receives the signals sent by the voice module and ESP8266, and judges the functions to be realized according to the signal type.

The first initialization: configure the pins and serial port baud rate, set some variables to be used later, and make each structure in the initial state.

The second step is the main loop, read the serial data, enter the condition judgment nesting, when the serial port has no data, the system standby; when the serial port receives data, enter the next level for condition judgment, when the condition is met, the corresponding operation will be executed. The design of the Control flow diagram of main control board is shown in Fig.15.

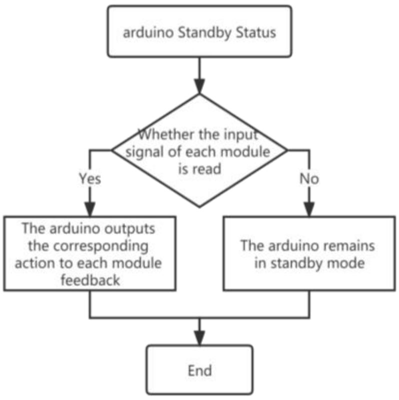

arduino Standby Status
Whether the input signal of each module is read
Yes
No
The arduino outputs the corresponding action to each module feedback
The arduino remains in standby mode
End

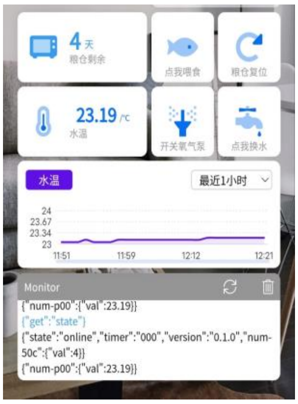
Fig.15 Control flow diagram of main control board

In the design, the lighting, oxygen pump, pumping pump and heating rod are all controlled by relays, which only need to output the corresponding pins to low level when the corresponding command is received or the corresponding conditions are met. The idea of the automatic feeding mechanism is slightly different. The feeding mechanism is controlled by the servo, and each time it receives a command, the servo will turn through 36°, i.e., one-tenth of the circumference. When the bin reaches the last frame, i.e. there is no fish food left, the bin is reset to the initial position, the attitude variable is set back to the initial value, and the buzzer pin is output high to prompt the user.

(2) ESP8266 networking part

1. Temperature sensor data processing

According to the analysis of the principle of DS18B20 implementation, the temperature sensor stores a temperature complement value and inverts its low 8 bits (LS Byte) and then performs +1 operation to restore to the original code. The Hex code data of the temperature value acquired by the temperature sensor is stored in RAM, and then the Hex code is converted to BCD code and stored in RAM. Write sensors.requestTemperatures() to send the command to get the temperature, sensors.getTempCByIndex(0) to return the water temperature measured by the sensor. esp8266 through WIFI and through the Temp component to display in the cell phone APP. The design of the Mobile phone APP in the numerical components and historical temperature effect is shown in Fig.16.


4 天
粮仓剩余
点我喂食
粮仓复位
23.19 /℃
水温
开关氧气泵
点我换水
水温
最近1小时
24
23.67
23.34
23
11:51
11:59
12:12
12:21
Monitor

{"num-p00":{"val":23.19}}
{"get":"state"}
{"state":"online","timer":"000","version":"0.1.0","num-50c":{"val":4}}
{"num-p00":{"val":23.19}}

Fig.16 Mobile phone APP in the numerical components and historical temperature effect

2.Mobile APP interface

With the help of blinker platform secondary development cell phone APP operation interface, connected to the third party server to achieve network control. Users can switch on/off the device, send the set temperature to the fish tank, watch the real-time situation and historical temperature curve of the fish tank, check the water temperature and the remaining days of fish food through this interface. The design of the Mobile phone APP interface is shown in Fig.17.

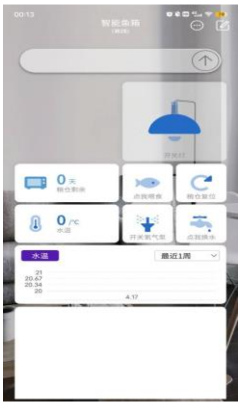

00:13
智能鱼缸
0 天
0 /℃
水温
最近1周
21
20.67
20.34
20
4.17

Fig.17 Mobile phone APP interface

When the button of the interface is pressed, ESP8266 sends the corresponding control command to arduino according to the switch status, and the function written corresponds to the components of the cell phone interface one by one, and the corresponding code is executed when the corresponding component is pressed. The numerical component is used to display the numerical value, which is the usage of Temp in the previous section.[4]

(3) Voice control part

The voice control part is written in C. The keywords are set during development, and the voice chip extracts the features from the voice and compares them with the keywords. The code is written and the serial port sends the corresponding data to the main control board when the corresponding command is recognized, thus realizing voice control. the SU-03T module can also be intelligently learned by other platforms to improve the recognition accuracy.

The voice recognition chip is connected to MCU and transmits the data using JSON format. After receiving this data, the serial port analyzes it to get the recognition code corresponding to the current command and executes the corresponding operation accordingly.

5.2. Neural network vision algorithm for identifying the health status of fish

(1) Data collection and processing

Firstly, a crawler is used to crawl images of goldfish from the web when they are healthy and when they are dead. When crawling with the crawler, we found that a large part of the collected data set was inaccurate (e.g., the background was too complicated), so we first reduced the number of data sets and selected some of them that were relatively pure and met the requirements of this design as data . The design of the Data set part of the data is shown in Fig.18. The design of the 3D diagram of data features is shown in Fig.19.

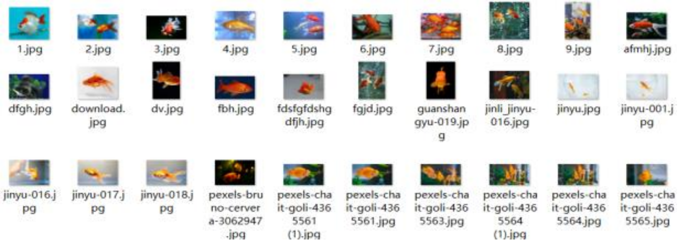

1.jpg
2.jpg
3.jpg
4.jpg
5.jpg
6.jpg
7.jpg
8.jpg
9.jpg
afmhj.jpg
dfgh.jpg
download.jpg
dv.jpg
fbh.jpg
fdsfgfdshgdfjh.jpg
fgjd.jpg
guanshangyu-019.jpg
jinli_jinyu-016.jpg
jinyu.jpg
jinyu-001.jpg
jinyu-016.jpg
jinyu-017.jpg
jinyu-018.jpg
pexels-bruno-cervera-3062947.jpg
pexels-chait-goli-4365561 (1).jpg
pexels-chait-goli-4365561.jpg
pexels-chait-goli-4365563.jpg
pexels-chait-goli-4365564 (1).jpg
pexels-chait-goli-4365564.jpg
pexels-chait-goli-4365565.jpg

Fig.18 Data set part of the data

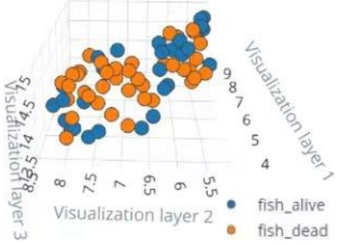


Fig.19 3D diagram of data features

The images were set in RGB form, and a normalization operation was performed on the images to convert each pixel channel of the image to a floating point value between 0 and 1. Before training all the images are resized t o 320*320px using the shortest axis fit method. after finishing the data processing work, the dataset is split into test set and training set in a certain ratio (two to eight) and labeled with the correct label for each object in the image, and finally the features are extracted. Most of the surviving data features are located in the upper right of the coordinate space, while most of the features of the dead data are concentrated in the lower right.

(2) Object detection model based on MobileNetV2

In order to achieve a good detection of fish survival in this design, this paper uses the MobileNetV2-based object detection model. The object detection algorithm is run on the Raspberry Pi to acquire images and output information about the class and number of objects in the images and their locations.

The effort to build a working computer vision model from scratch is significant, as a wide variety of input data is required to make the model generalize well, which makes the training time of the model long, which can take several days. To make this process more tolerable and faster, the fishbox is designed to use migration learning. In turn, well-trained models can be piggybacked and only the upper layers of the neural network need to be retrained, resulting in more reliable models that are trained in a very short period of time and use smaller data sets.

This design uses the MobileNetV2 SSD FPN-Lite pre-trained model that has been trained on the COCO 2017 dataset with images scaled to 320x320px. in MobileNetV2 SSD FPN-Lite, we have a basic network (MobileNetV2), a detection network (single detector or SSD) and a feature extractor (FPN-Lite).

The base network MobileNet, such as VGG-Net, LeNet, AlexNet and all other networks, are based on neural networks. The base networks provide advanced functionality for classification or detection. If fully connected layers and softmax layers are used at the end of these networks, a classification network is obtained.[5]

Also, it is possible to remove the fully connected layers and softmax layers and replace them with detection networks such as SSD, Faster R-CNN, etc. , to perform object detection. The design of the Schematic of classification network is shown in Fig.20.

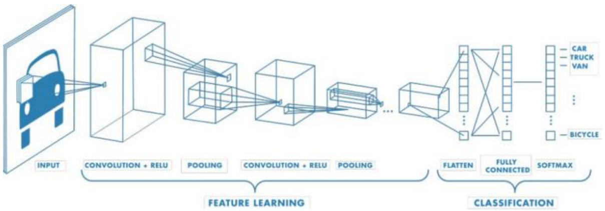


Fig.20 Schematic of classification network

The most common detection networks are SSD (Single Shot Detection) and RPN (Region Proposal Network). When using SSD, we only need to take one shot to detect multiple objects in an image. On the other hand, Region Proposal Network (RPN) based methods, such as the R-CNN series, require two shots, one for generating region proposals and one for detecting objects for each proposal. As a result, SSDs are much faster compared to RPN-based methods, but usually swap accuracy with real-time processing speed. They also tend to have problems in detecting objects that are too close or too small. Feature pyramids are a challenging approach in targets of different sizes, especially on smaller targets.

This method, Feature Pyramid Network (FPN), uses the concept of feature pyramid design, which can improve the recognition accuracy and processing speed.

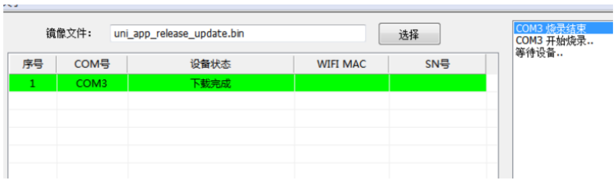
6. System operation test

6.1. The commissioning and preparation of each part

Feeding mechanism adjustment, use the cell phone APP for feeding mechanism debugging, after debugging, the initial angle of the servo is set to 42 °, feeding mechanism and chassis hole position with a better.

After adjusting the feeding mechanism, the cell phone APP was used to control the fish box to complete all kinds of actions. After testing, the pumping, oxygen pump, automatic feeding, lighting and other functions can operate normally. The difference between the processed data of the temperature sensor part and the actual water temperature is very small, and the heating equipment turns on normally when the water temperature is lower than the user's set temperature to realize the automatic control of water temperature.

Voice part: hardware construction is completed, start the program writing and burning. sU-03T module program burning needs to be completed with the help of burning tool UniOneUpdateTool. Burning firmware needs to be burned first, and then power supply to the module, sometimes it is necessary to carry out multiple power supply and power failure to ensure the burning of the firmware. The design of the Firmware burn-in is shown in Fig.21.



镜像文件： uni_app_release_update.bin

序号	COM号	设备状态	WIFI MAC	SN号
1	COM3	下载完成		

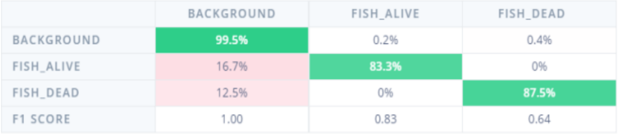
COM3 烧录结束
COM3 开始烧录..
等待设备..

Fig.21 Firmware burn-in

The number of training rounds is gradually increased from 20 to 50, so the number of rounds is set to 50 to ensure the training efficiency.

6.2. Run and derive results

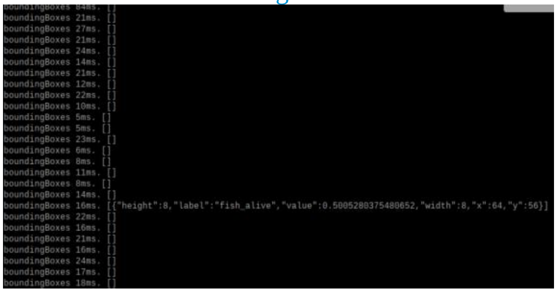
After the control code is written, the smart fish box is woken up by the "Little Fish" command, and the fish box will complete the corresponding action when the command "lighting" is said to the fish box; the fish box can be put back into the standby state by using the end command, and it needs to be woken up again to realize the command The end command can put the fish tank back into standby state, and it needs to wake up again to realize the command. For the vision part, after the parameters are set, the training of the model begins. The design of the Confusion matrix is shown in Fig.22.



	BACKGROUND	FISH_ALIVE	FISH_DEAD
BACKGROUND	99.5%	0.2%	0.4%
FISH_ALIVE	16.7%	83.3%	0%
FISH_DEAD	12.5%	0%	87.5%
F1 SCORE	1.00	0.83	0.64

Fig.22 Confusion matrix

Run the vision part of the algorithm, deploy the model to identify the survival status of fish to Raspberry Pi4, run the command at the command line to automatically compile the model with full hardware acceleration, download the model to Raspberry Pi, and then start detection. The object coordinates and labels are output when the object is detected, and both dead and healthy fish are correctly labeled in the interface. [6] The design of the Code running interface is shown in Fig.23. The design of the Target detection effect is shown in Fig.24.



```
boundingBoxes 84ms. []
boundingBoxes 21ms. []
boundingBoxes 27ms. []
boundingBoxes 21ms. []
boundingBoxes 24ms. []
boundingBoxes 14ms. []
boundingBoxes 21ms. []
boundingBoxes 12ms. []
boundingBoxes 22ms. []
boundingBoxes 10ms. []
boundingBoxes 5ms. []
boundingBoxes 5ms. []
boundingBoxes 23ms. []
boundingBoxes 6ms. []
boundingBoxes 8ms. []
boundingBoxes 11ms. []
boundingBoxes 8ms. []
boundingBoxes 14ms. []
boundingBoxes 16ms. [{"height":8,"label":"fish_alive","value":0.5005280375480652,"width":8,"x":64,"y":56}]
boundingBoxes 22ms. []
boundingBoxes 16ms. []
boundingBoxes 21ms. []
boundingBoxes 16ms. []
boundingBoxes 24ms. []
boundingBoxes 17ms. []
boundingBoxes 18ms. []
```

Fig.23 Code running interface

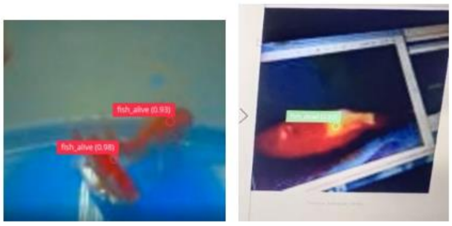


Fig.24 Target detection effect

7. Conclusion

Nowadays, smart home and Internet of Things technology is developing rapidly, and the design of smart fish box in this paper is following the trend of technology development, and various kinds of knowledge such as microcontroller, machine vision, artificial intelligence and voice control are cooperated with each other to form a complete system, and it is believed that various kinds of smart fish boxes will be more widely used in the automated fish rearing in the future. This design combines the knowledge of several disciplines to improve the human-computer interaction experience, and it is believed that there will be more and more demand for this type of product in the market.

References

1. Yuan Chunmei, Liu Sinian, ZhouJie. Design of intelligent monitoring system for fish tank based on mobile Internet, *Information Technology and Informatization*,2020(09):216-218.
2. Xu Haoming,Wang Yunlong,Yang Jun,Chen Kaijie. Intelligent fish tank system based on Internet of Things, *Internet of things technology*,2021,11(10):89-90+96.
3. Cao Yihao. Design of intelligent fish tank temperature control system based on microcontroller, *Mechatronics Information*,2019(11):11-12.
4. Kang Rongxian,Wang Xun,Wang Xiaotong et al. Research on intelligent control system of fish tank based on Internet of Things, *Information Technology*,2022,12(09).
5. Zhang Jing, Nong Changrui, Yang Zhiyong. A review of target detection algorithms based on convolutional neural networks, *Journal of Arms and Equipment Engineering*,2022,05-07:1-12.
6. Liu Rujun, Zheng Zekai, Zhao Min. Design and implementation of an intelligent aquaculture system based on Raspberry Pi, *Internet of things technology*,2022,12(04):67-70.

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. Jiangyu Wu

He received his Academic degrees from the School of Electronic Information and Automation, Tianjin University of Science and Technology in 2021.

Mr. Shuai Chen

He is currently pursuing his undergraduate degree at the School of Electronic Information and Automation, Tianjin University of Science and Technology. He research field software design.

Dr. Yizhun Peng

He is an Associate Professor in Tianjin University of Science &Technology. He recevived 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 Implementation of Internet of Things Planting System Based on esp32 MCU

Cuiying Ji , Yizhun Peng*

College of Electronic Information and Automation,Tianjin University of Science and Technology, Tianjin, China
*E-mail: *pengyizhun@tust.edu.cn*
www.tust.edu.cn

Abstract

Agriculture is the country's fundamental industry, but there is still a lot of hard work because of climate pests. It is the development direction of agriculture to integrate traditional agriculture into modern Internet of Things technology and realize visualization and digital management of agricultural production. This study combined modern information technology and used esp32 as the master to design and produce a plant planting system model for the Internet of Things.This research product uses a variety of sensors, Monitor the essential elements required for crop growth, including temperature, humidity, nutrient solution, soil moisture, etc.In the process of crop growth, automatic irrigation, nutrient solution supplement, ventilation cooling, light and other operations, greatly reduce the intensity of manual labor.

Keywords: Internet of Things; Soilless culture, Remote control, Automatic control

1. Introduction

With the rapid development of science and technology, mobile communication is changing from people to people to things, and then to things. The interconnection of everything has become an inevitable trend of mobile communication development.[1] This Internet of Things technology came into being. And it is considered to be the third wave of the world information industry after the computer and Internet industries. The Internet of Things technology is a new technology that senses and identifies external user information through various sensors, connects any object to the network according to the agreed protocol, analyzes and transmits information, and at the same time combines Internet technology to realize intelligent cognition, transmission, management and regulation of users.[2] The rest of this article is organized as follows. The rest of the article is organized in order. The second section introduces the hardware facilities of this design product. The third section introduces the overall structure of the system. Section IV introduces the overall content summary.

2. Main control board

In this design, ESP-WROOM-32 is selected as the main control board, which integrates traditional Bluetooth, low-power Bluetooth and Wi Fi, and has a wide range of applications: Wi Fi supports a wide range of communication connections, and also supports direct connection to the Internet through a router; Bluetooth can

To enable users to connect mobile phones or broadcast BLE Beacon for signal detection. The data transmission rate supported by ESP-WROOM-32 is up to 150Mbps. After passing through the power amplifier, the output power can reach 22dBm, enabling a wide range of wireless communication. Therefore, this chip has leading technical specifications, and has excellent performance in terms of high integration, wireless transmission distance, power consumption and network connectivity.[3] The design of theMain control board is shown in Fig.1 .

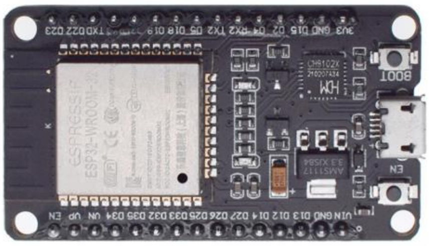


Fig.1.The design of the Main control board

2.1. Camera module

The right camera module is ESP32-CAM. It is a small size camera that can be released by Anson and is equipped with an OV2640 camera. The size is only 27x40.5x4.5mm. The main frequency is up to 240MHZ. It has a built-in 520KB SRAM and an external 8MB PSRAM. The board contains flash and has a TF card interface (which can be used to save pictures taken). It can fully display videos in real time on the web page or app.[4] ESP32-CAM is shown in the Fig.2.

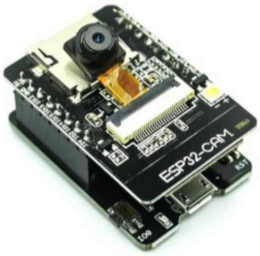


Fig.2. ESP32-CAM

2.2. Water sensor

The water sensor is a simple and cost-effective water level Sensor. Its working principle is to determine the water level by measuring its Water volume through a series of exposed parallel wire traces. The conversion from water volume to analog signal can be easily completed, and the output analog value can be directly read by the esp32 development board to achieve the effect of real-time monitoringThe Water sensor is shown in Fig.3.

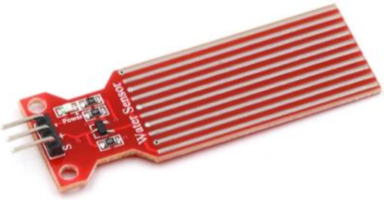


Fig.3. Water sensor

2.3. Temperature and humidity sensor

The temperature and humidity sensor adopts DTH11, which is a temperature and humidity composite sensor with calibrated digital signal output. It applies special digital module acquisition technology and temperature and humidity sensing technology to ensure that the product has reliability and excellent long-term stability, low cost, relative humidity and temperature measurement, fast response, strong anti-interference ability, long signal transmission distance, digital signal output, and accurate calibration.It has the following characteristics:

- Humidity measurement range: 20% ~ 95%.
- Humidity error: -5% ~ +5%
- Temperature measurement range:0 ºC ~ 50 ºC
- Temperature error:-2 ºC ~ 2 ºC
- working voltage:3.3V~5V

The design of the main control chip is shown in Fig.4.

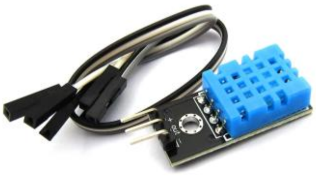


Fig.4. Temperature and humidity sensor

2.4 Display module

0.96 inch oled display screen is selected, and IIC communication protocol is adopted. Voltage stabilizing circuit design。

2.5 Photoresistor module

In automatic mode, XH-M131 photoresistor module is selected to turn on and off the fill light, which can be regarded as a combination of relay and photoresistor Its working principle is that when the photosensitive detector detects that the light becomes bright, the switch automatically closes, the load works, and the blue light of the working indicator lights up; When the light is dim, the relay will be disconnected, the load will stop working, and the working indicator will go out. The design of the Photoresistor module is shown in Fig.5.

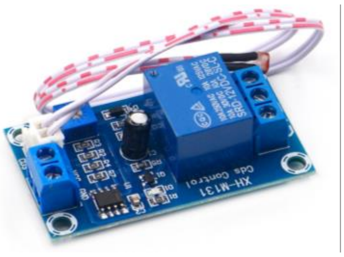


Fig.5. Photoresistor module

2.6 Electronic switch

The hardware electronic switch of this design product uses a relay. This component can be connected and disconnected through the control of high and low levels. That is to say, I can easily control whether the circuit is connected or not by just using the pin for high and low levels. Very simple and convenient Fig.6.

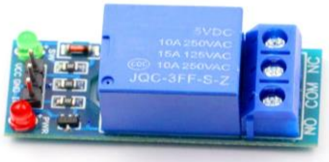


Fig.6. Photoresistor module

2.4 Other components

The fan is a small 5V power supply fan, the water pump is a submersible small water pump with a built-in brush motor, The 3.3V power supply LED lamp board, power supply part are No. 5 battery and 2596 voltage reduction module.

3. System framework

The Internet of Things technology is an important part of modern information technology. The Internet of Things realizes the fast and accurate connection between things.[5] As a new generation of information technology, the Internet of Things still takes the network as the core link to extend and expand user information through the network. Second, the Internet of Things realizes the transmission, communication and exchange of information between goods through the network. The Internet of Things is widely used in various fields through the combination of sensing technology, identification equipment and ordinary computer network technology, impacting the traditional operation mode. Taking traditional agriculture as an example, planting, management and harvesting are carried out manually, while the Internet of Things technology uses sensors, the Internet and the application layer to achieve user intelligence, information management and application.[6]

3.1. The Internet of Things. Basic framework

The Internet of Things is composed of three layers: the perception layer, the network layer and the application layer. The sensing layer consists of temperature and humidity sensors, cameras and other identification information terminals. The perception layer is an extremely important part of the development of the Internet of Things, and an important carrier and way to identify and collect external information. Therefore, the perception layer is called the eye of the Internet of Things. The network layer is responsible for transmitting the information acquired by the perception layer to the application layer through the network, which is the core part of the Internet of Things technology. Similar to the brain and central system of higher animals, the network layer is responsible for analyzing and processing information, and transmitting the information to the application layer; Then the application layer reacts. The application layer is the Internet of Things connected to the user control terminal, responsible for information regulation and management.[7]

3.2. System framework design of this product

The Internet of Things planting system for soilless culture designed in this study can enable the plant growth process to automatically supply nutrient solution, supply light, open ventilation settings, real-time monitoring, etc. according to the compiled work program. This design is divided into two modes, one is intelligent mode, the other is manual mode.[8]

In the intelligent mode, various sensors in the system conduct all-weather and uninterrupted real-time monitoring, make intelligent judgments based on real-time data and control models, and make corresponding adjustments according to actual needs, so as to achieve a better environment for plant growth in an unmanned environment. Realize small-scale agricultural automation. In the other manual mode, the plant growth can be monitored remotely and in real time, and some adjustments can be made according to the real-time data, such as nutrient solution supply, ventilation, etc. The manual mode is built on the basis of the intelligent mode, which means that we are generally in the intelligent mode. The manual mode is set for the failure of the intelligent mode due to uncertain factors.

The system integrates sensor, automatic monitoring, automatic control, communication, calculation and other technologies with expert system to realize automatic monitoring and control of temperature, humidity, nutrient solution and other factors in the strawberry soilless cultivation greenhouse.

4. Conclusion

In this paper, the design and implementation of the internet of things planting system based on esp32 master microprocessor are described and the basic working principle is designed. The framework of the system is also designed in detail. In the next stage, we will be committed to detailed system design on this basis.

References

1. Ye Yun. Research on the Application of Internet of Things Technology in the Age of Big Data [J]. Industrial Innovation Research, 2022 (22): 76-78
2. Huang Lili. Application Analysis of Internet of Things Technology in Smart Home [J]. Digital Technology and Application, 2022,40 (09): 39-41. DOI: 10.19695/j.cnki.cn12-1369.2022.09.12
3. Liu Qing, Yan Jinkui. Design and implementation of environmental monitoring system for vegetable greenhouses based on ESP32 [J]. Industrial Control Computer, 2021,34 (09): 137-138+140
4. Cui Jintao, Tong Ruidong. Design of visual tracking robot based on ESP32 and MQTT protocol [J]. Journal of Liaoning Teachers' College (Natural Science Edition), 2021,23 (04): 67-71
5. Lin Ning, Zhang Liang. Application of Internet of Things Technology in Soilless Cultivation [J]. Computer Knowledge and Technology, 2015,11 (25): 153-155. DOI: 10.14004/j.cnki.ckt.2015.2504
6. Zhou Zijun, Liu Xiaobo, Ye Zhichan, Li Huaiyuan, Yang Haijun. Research on the application of the Internet of Things cultivation mode of elevated strawberry [J]. Agricultural Science and Technology Communication, 2021 (04): 169-171
7. Huang Lili. Application Analysis of Internet of Things Technology in Smart Home [J]. Digital Technology and Application, 2022,40 (09): 39-41. DOI: 10.19695/j.cnki.cn12-1369.2022.09.12
8. Lei Yan, Wei Congqi. Design of intelligent agricultural environmental monitoring system based on Internet of Things technology [J]. Electronic Technology and Software Engineering, 2022 (18): 246-249

Authors Introduction

Ms. Cuiying JI

She is an undergraduate of Tianjin University of science and technology, her research field is Intelligent manufacturing.

Dr. Yizhun Peng

He is an Associate Professor in Tianjin University of Science &Technology. He recevived 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 Customized Dispensing Robot Based on OpenMV Visual Recognition

Hongze Liu , 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

Dispensing medicine in hospital wards is a mechanized process, but it is difficult to develop on a large scale due to the uncertainty of the environment. In this study, a customized dispenser robot was designed to automatically deliver drugs to the designated ward according to the given instructions. The Visual identity module (OpenMV) is designed to collect images from the progress and then compare them with pre-stored images to achieve automatic pathfinding. Among them, the tracking part uses PID open-loop operation to increase the accuracy of movement. The ward number enables the camera to recognize the number in different scenes for many times, which increases the accuracy of recognition. The unique structure of this study provides an innovative and effective way to customize ward dispensing.

Keywords: OpenMV, STM32; image recognition; multi-template matching

1. Introduction

Modern society cannot do without medicine and the intelligence of hospitals and other health institutions has become a current trend in development. The development of intelligence has taken the management, services and research of healthcare to a new level, helping hospitals to provide better services and offer higher quality healthcare to society.[1] With the expanding needs of the medical industry in recent years and the emergence of innovative technologies and systems for medical robotics, medical automation is emerging as an emerging industry with great promise.[2] But health institutions around the world are all different, and this requires robots that can have the ability to customise recognition to handle it.

Next we will discuss the Custom Delivery Robot (CDR). This is a robot based on OpenMV visual recognition, which can recognise external room numbers based on visual input and can automatically plan routes. In addition, it can perform automatic localisation and recognition to ensure effective obstacle avoidance of objects in front of it.

2. Digital recognition solutions

OpenMV forms the main vision recognition module (VPU) for CDR, an open source vision recognition system that can be used to develop and deploy custom CDR[3]. The system consists of an ov7725 camera and a 51 microprocessor that can be used to detect objects and identify their position in the environment. With its ability to recognise objects and their positions, OpenMV can be used to create bespoke CDR that are tailored to the exact needs of a given application. The design of OpenMV is shown in Fig.1.

The OpenMV system has been successfully tested in many different scenarios, including object recognition and navigation[4]. In addition, the system has proven to be reliable and robust in a wide range of environments. This makes the OpenMV system ideal for the development and deployment of CDR. In addition, the system provides a cost-effective solution for those wishing to create efficient and reliable robotic recognition systems.

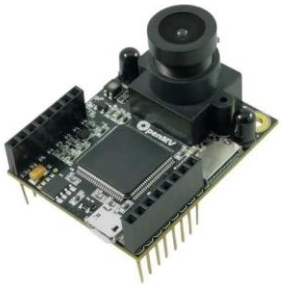


Fig. 1 The design of OpenMV

3. Tracing and junction identification solutions

Grey sensors are analogue sensors. The grey scale sensor uses the principle that different colours reflect light differently on the detection surface and that the photoresistor has a different resistance to the light returned from the different detection surfaces to detect the shade of colour. The design of the infrared homing module is shown in Fig.2.

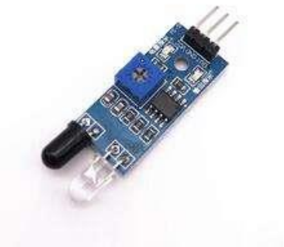


Fig. 2 The design of the infrared homing module

It is used to distinguish between black and other colours when the ambient light interference is not very serious. It also has a relatively wide operating voltage range and can work normally in the event of large fluctuations in the supply voltage. The output is a continuous analogue signal, which makes it easy to determine the reflectance of an object by means of an A/D converter or a simple comparator, making it a practical sensor for robotic line inspection. The grey scale sensor is equipped with an analogue size adjuster for the return of the detected colour. To detect a given colour, the transmitter/receiver head can be placed at the given colour and the appropriate return analogue can be adjusted with the regulator. Turn the regulator counterclockwise to increase the return analogue amount; turn the regulator clockwise to decrease the return analogue amount; and continue adjusting until the desired value is reached. If the exact analogue quantity is required, it can be programmed to be displayed on the LCD screen and the exact analogue quantity can be adjusted in conjunction with the regulator. Because different areas require different colours and thresholds to be recognised, the CDR can automatically distinguish and save the corresponding colour selection according to the size of the set threshold.

4. Microcontroller unit selection options

The STM32F1 microprocessor is a powerful and efficient alternative to the CDR. This microprocessor offers a high level of computing power with a wide range of features such as a 32-bit ARM Cortex-M3 core, DMA controller and multiple on-chip peripherals. In addition, it offers a wide range of on-chip memory options, including up to 512 KB of Flash memory and up to 64 KB of SRAM. These features make the STM32F1 ideal for a variety of applications, such as robots that need to process data quickly and accurately.

Secondly, the STM32F1 microprocessor is low power, with multiple low-power modes that can be used to reduce power consumption while maintaining performance. In addition, the processor is able to run from a low-voltage power supply, allowing the robot to be powered by a battery or other low-voltage power source. This feature is particularly beneficial in applications with limited or unreliable power supplies. The design of the MCU is shown in Fig.3.

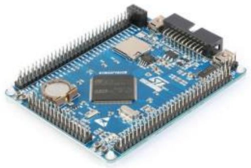


Fig. 3 The design of the MCU

Finally, the STM32F1 microprocessor is highly flexible and scalable. The processor can be used in combination with a wide range of peripherals and external memory.[5] This allows the robot to be customised to meet the specific needs of its application. In addition, the processor is able to run multiple tasks in parallel, allowing the robot to perform complex tasks more efficiently.

5. Software design and theoretical analysis

This summary will consider the functionality, architecture and performance of the system in order to determine the best solution. In order to achieve this goal, this paper will use computer vision techniques to build the CDR.

Firstly, a systematic analysis of the OpenMV framework will be carried out to identify its possible application scenarios and functional characteristics.[6] A live image is acquired by the camera and then the numbers are matched to the stored template and returned by a model recognition algorithm. The microcontroller processes the numbers and then controls the DC motor to drive in the corresponding direction, during the driving process the camera then matches the numbers captured during the driving with the stored template through the model recognition algorithm and returns the numbers. The microcontroller processes the numbers and makes the correct direction of travel, stops after identifying the corresponding ward, waits for the medicine to be taken away and then reverses and returns in the same direction according to the route recorded by the CDR on its way, and prompts after returning to the pharmacy.

6. Openmv model recognition algorithm

The basis for number recognition is the need to configure the use of NCC template matching.[7] The NCC template matching allows the template image of the numbers to be recognised to be saved to an SD card, which can then be used for the next step of recognition. During the programming process it is realised that pictures of different numbers are stored for model entry and then during the CDR run the models are compared with the signage ward number plates of the nearby walls and the corresponding numbers are recognised for the correct route selection.

Based on this algorithm, multiple sets of images are then entered, making it possible for the identified images to be found in the image library for comparison under different lighting and other external conditions. The comparison is shown in Fig.4.

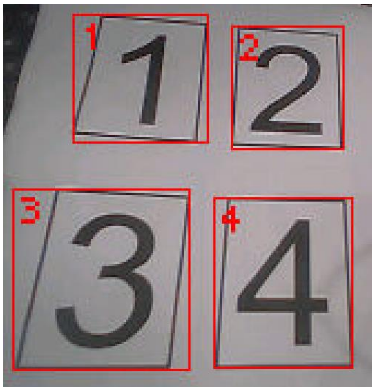


Fig. 4 Vision modules recognise numbers in different lighting conditions

7. PID algorithms

In process control, the PID controller (also known as PID regulator) is the most widely used automatic controller according to the proportional (P), integral (I) and differential (D) control of deviations. It has the advantages of simple principle, easy implementation, wide application, independent control parameters and simple selection of parameters; and in theory it can be proved that for the typical objects of process control - "first order hysteresis + pure hysteresis" and "second order hysteresis + pure hysteresis "The PID regulation law is an effective method for dynamic quality correction of continuous systems, with simple parameterisation and flexible structure changes. The use of the PID algorithm during the CDR travel can make the CDR travel more stable.[8]

8. Hardware circuit design

The hardware circuitry should be designed in such a way that all components of the CDR work together to achieve the required functionality. The hardware circuit consists of an STM32f103rct6 microcontroller, a power supply module, an image recognition module, a motor drive module, an optoelectronic module and a greyscale sensor module.

The STM32f103rct6 microcontroller is the core part of this hardware circuit, the power supply module provides sufficient power for the entire circuit, the image recognition module recognises images, the motor drive module translates commands into motor movements, the optoelectronic module detects light intensity and the greyscale sensor module detects the colour and shape of objects. Openmv is connected to the main control through the serial port, and the openmv serial port is connected to the serial port of the microcontroller. The block diagram of the scheme structure is shown in Fig.5. And the circuit schematic for the specific design is shown in Fig.6.

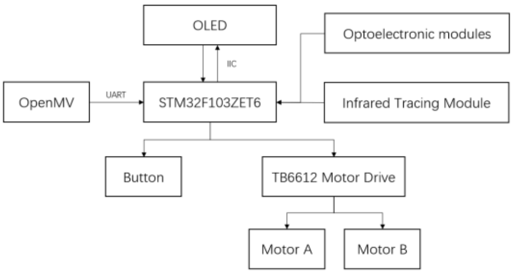


Fig. 5 The block diagram of the scheme structure

The DC regulated power supply is a device that converts AC power to DC power and can be connected directly to the motor drive module, which is then connected to the motor drive module to enable the motor to be driven.[9],[10] The greyscale sensor and the optoelectronic module are also each connected to the gpio of the microcontroller to enable data transfer. The operation of the entire system relies on the interplay between these devices and the functions they provide.

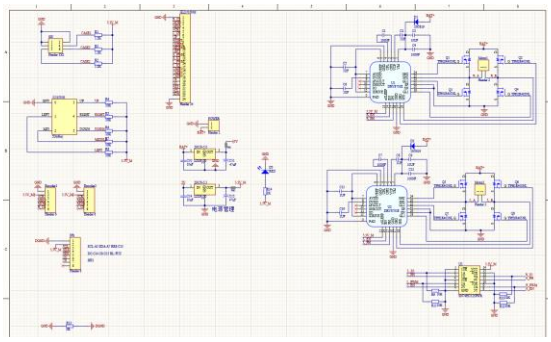


Fig. 6 Circuit schematic

9. Conclusion

This thesis presents an OpenMV vision recognition-based custom drug dispensing robot that enables fast and accurate drug dispensing, an intelligent drug transport process and reliable drug dispensing services. This novel custom drug dispensing robot greatly improves the efficiency of drug dispensing and provides a reliable drug dispensing service for clinical purposes.

References

1. Tan, M. & Wang, SH. (2013). Advances in robotics research. *Journal of Automation* (07), 963-972.
2. Zhao Xinguang, Duan Xingguang, Wang Qining & Xia Zeyang. (2021). Research perspectives on medical robotics. *Robotics* (04), 385.
3. Wang Congzheng & Li Qiguang. (2018). Opencv-based multi-target template matching algorithm for images. *Electronics Technology and Software Engineering* (05), 57-59.
4. Ding, Heng-Wang, Zheng, Jin-He & Yan, Shao-Min. 2011 A new species of the genus Phyllostachys (Coleoptera, Staphylinidae) from China. (2022). Research on drug delivery system based on neural network combined with template matching algorithm. *Automation Instrumentation* (09), 52-55.
5. Chen, Yan, Hu, Guang, Luo, Houcheng, Wan, Shuhan, Chen, Chuang & Zou, Min. (2022). Design of an STM32 chip-based intelligent guide cane for the blind. *Electronic Technology*(10),10-11.
6. Zhang Lei. (2016). Design of intelligent trolley control system (Master's thesis, Jilin University).
7. Yue Youjun, Huo Xiaodong, Wang Hongjun & Zhao Hui. (2023). Research on dynamic identification and tracking algorithm of apple. *Agricultural Mechanization Research* (06), 41-46.
8. Zhang Bo. (2021).Application of PID algorithms in control systems. *Information Record Materials* (07), 197-199.
9. Zhou, Jian & Xu, Zhiguo. (2020). Design of an STM32-based intelligent tracer robot. *Integrated Circuit Applications* (02), 45-47.
10. Jiang Linglong, Hu Jian & Liu Qingyun. (2021). Non-contact object recognition device based on OPENMV. *Science & Technology Innovation* (09), 160-161.

Authors Introduction

Mr. Hongze Liu

He is an undergraduate of Tianjin University of science & technology, his research field is software design.

Dr. Yizhun Peng

He is an Associate Professor in Tianjin University of Science &Technology. He recevived 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.

An Intelligent Guide Hat Based on The Internet of Things

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

With STM32 as the control core, the new guide cap has ranging obstacle avoidance, ranging obstacle avoidance, intelligent recognition, voice interaction, GPS positioning and other functions. The system uses OpenMV4 Cam H7 Plus intelligent camera for continuous real-time monitoring and effective feedback of traffic lights, zebra crossings and other important road information; Assisted by ultrasonic sensor, the electrical signal is converted into ultrasonic output through the transmitter to effectively measure the distance of obstacles; At the same time, the user's location information is sent to the surrounding vehicles and passers-by in real time through the Wi-Fi module and GPS positioning, ensuring the user's safety to the greatest extent.

Keywords: microcontroller, intelligent guid, Internet of things, voice control, target detection

1. Introduction

The number of blind people in the world is very large. According to the World Health Organization, by 2022, there will be between 40 and 45 million blind people and approximately 140 million people with low vision worldwide.

We see that the number of blind people is still on the rise every year: "mobility for the blind" has become a pressing issue. In recent years, with the rapid development of computer software and hardware technology, wearable electronic products have come into our lives, allowing the blind to enjoy the benefits of technological advances. This product is based on solving the problem of blind people's travel, in order to help blind people solve such difficulties, realize voice accessibility control, integrate multi-functional integration, assist blind people to build a mental map, and hope to facilitate the life of blind people.[1]

The rest of this paper is organized as follows. Section II carries out the structural design of the smart guide cap, which contains the overall system scheme design and the introduction of auxiliary sensors. The third section introduces the intelligent guide cap software design and algorithm, including various core technologies such as image processing algorithm, obstacle distance detection algorithm, communication program, positioning algorithm, and voice interaction program. The fourth part summarizes the main contents of this paper.

2. Intelligent Guide Cap Structure Design

2.1. Overall system design

This design uses STM32 as the control core of the new guide cap, which has the functions of distance and obstacle avoidance, intelligent recognition, voice interaction and GPS positioning. The system is continuously monitored in real time by OpenMV4 Cam H7 Plus intelligent camera, which effectively feeds back important road information such as traffic lights and crosswalks, and is supplemented by ultrasonic sensors that convert ultrasonic waves into electrical signals through transmitters to effectively measure the distance of obstacles; at the same time, through Wi-Fi module and GPS positioning, the user's location information is sent to surrounding vehicles and passers-by in real time to ensure maximum safety for the user. To ensure the safety of users to the greatest extent. The design of the 3D model of the product is shown in Fig.1. The design of the physical picture of the product is shown in Fig.2.

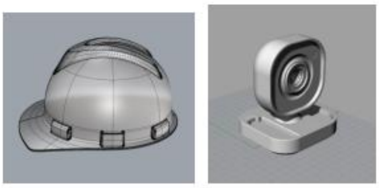


Fig.1 3D model of the product

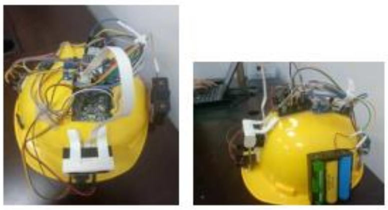


Fig.2 Physical picture of the product

2.2. Auxiliary sensors

(1) OpenMV4 Cam H7 Plus

This product uses a low-cost, scalable, low-power, python-driven machine vision module, OpenMV-H7, based on the 32-bit ARM Cortex-M7 core, equipped with a MicroPython interpreter, which makes programming machine vision algorithms on the embedded easy.[2] At the same time, OpenMV provides communication interfaces such as synchronous asynchronous transceiver, serial peripheral interface and integrated circuit bus, which are capable of processing and programming complex algorithms, providing OpenMV with strong performance that is well capable of shape recognition, line recognition, color recognition, etc. The design of the OpenMV4 Cam H7 Plus picture is shown in Fig.3.

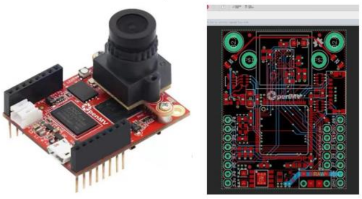


Fig.3 OpenMV4 Cam H7 Plus picture

(2) Ultrasonic Sensor

Ultrasonic distance measurement module uses low power consumption, low price, easy to connect HC-SR04 ultrasonic distance sensor, the core of which is two ultrasonic transducers, one used as a transmitter, the duration of each pulse group emitted is about 0.5ms, the signal is amplified through the transistor, and then after the impedance matching circuit that is transformer, the two ends of the transmitting transducer is coupled with a high voltage, the internal piezoelectric chip began to vibrate and send out 40kHz pulse ultrasonic. When the output ultrasonic wave meets the obstacle return, the TL074 chip to return the sound wave filtering, amplification, demodulation and other processing, and then through the STC11 chip for ultrasonic waveform and level conversion, return ECHO a certain time of high level signal. Then the port enters the low power state, the module stops timing and reads the timer worth to the time required for the sound wave round trip, and then calculates the distance to the obstacle by calculating this time. The control program will make processing of the measured data, if the distance is less than 3 meters voice system will immediately issue an alarm.[3] The design of the ultrasonic sensor is shown in Fig.4.

Fig.4 Ultrasonic sensor

(3) Wi-Fi communication module

This module integrates a 32-bit Tensilica processor, standard digital peripheral interface, antenna switch, RF balun, power amplifier, low-noise amplifier, filter and power management module, etc. It only requires few peripheral circuits and takes up significantly less PCB space. It supports low power consumption mode, the current required to keep Wi-Fi connection in general scenario is only about 50mA, and the power consumption is even reduced to 10uA during deep sleep. The design of the Wi-Fi communication module is shown in Fig.5.

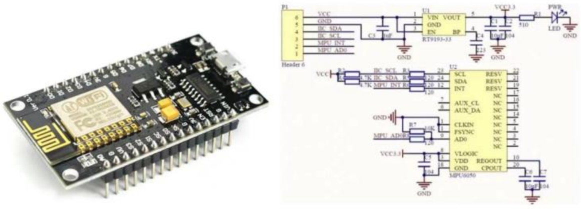


Fig.5 Wi-Fi communication module

(4) GPS Module

The GPS module uses the STM32 microcontroller as the processor, and the data signal output from the GPS signal receiving module is passed to the serial interface of the STM32 microcontroller. The signal receiving module mainly consists of inverter, signal channel, memory, central processor and input/output interface, using the RXD1 and TXD1 pins of the UM220 module, which are connected to the STM32 microcontroller through the serial chip, and the reset circuit realizes low-level reset by jumping from high to low BDRST potential when S11 is pressed; and the peripheral circuit is composed of GPS receiver devices and auxiliary circuits. The design of the GPS module is shown in Fig.6.

Fig.6 GPS module

(5) Voice interaction module

The hardware of voice interaction is mainly composed of AIUI speech recognition module, audio amplifier module and speaker, AIUI is the core component of the voice module, whose CPU uses R16 a cortex-A7 intelligent hardware processor; the running memory uses K4B4G16 with memory size of 1GB; Flas uses NCEMASD9-08GEMMC with storage space.[4] The ADC audio acquisition chip is Cosmos CX20819-11Z, a 4-channel far-field voice capture high performance HD video ADC, and the two are directly connected for data interaction through serial port, and the baud rate of 115200bps is used for serial communication parameters.[5] The design of the voice interaction module is shown in Fig.7.

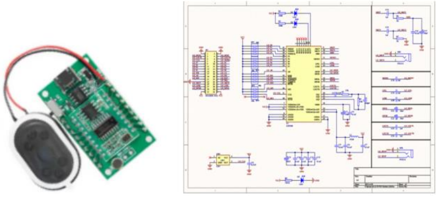


Fig.7 Voice interaction module

(6) Angle sensor module

This product uses the integrated six-axis angle sensor MPU6050 from InvenSense, which integrates a three-axis MEMS gyroscope, a three-axis MEMS accelerometer and an expandable digital motion processor DMP inside, so that it can be combined with a motion processing library to achieve attitude solving. With the self-contained DMP, the data of the nine-axis fusion algorithm can be output through the IIC interface to reduce the load of the motion processing operation on the operating system. The design of the angle sensor module is shown in Fig.8.

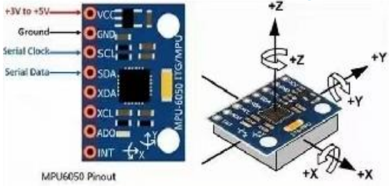

+3V to +5V
Ground
Serial Clock
Serial Data
VCC
GND
SCL
SDA
XDA
XCL
ADO
INT
MPU-6050 ITG/MPU
MPU6050 Pinout
+Z
+Z
+Y
+Y
+X
+X

Fig.8 Angle sensor module

3. Intelligent Guide Cap Software And Algorithm

3.1. Overall algorithm

This product uses STM32 for C language programming, mainly to achieve real-time control and real-time monitoring software two parts, including intelligent image recognition, ultrasonic distance measurement, voice interaction and other functions. In the main program, two interrupts are set, one for the network communication interrupt, which is used to receive instructions and data from the monitoring computer; the other for the timer with a frequency of 1 kHz. When the interrupt is queried, it enters the control interrupt subroutine, which collects and processes the data collected by each module, then judges the current control mode and the type of response, and then further solves and sends the The control signal can be sent according to the set control algorithm, which can complete the work of the guide efficiently and accurately. For example, the OpenMV4 Cam H7 Plus distance measurement module can perform effective and accurate distance measurement work. The design of the overall algorithm communication picture is shown in Fig.9.

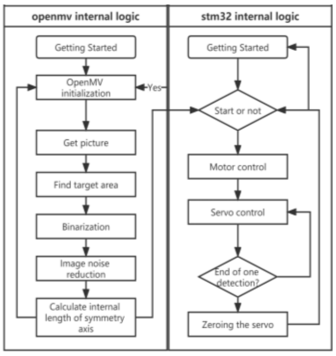

openmv internal logic
stm32 internal logic
Getting Started
OpenMV initialization
Yes
Get picture
Find target area
Binarization
Image noise reduction
Calculate internal length of symmetry axis
Getting Started
Start or not
Motor control
Servo control
End of one detection?
Zeroing the servo

Fig.9 Overall algorithm communication picture

3.2. Communication protocol development

(1) OpenMV uses python to drive machine vision modules, equipped with MicroPython interpreter, which makes it easy to program machine vision algorithms on embedded; meanwhile, OpenMV provides communication interfaces such as synchronous asynchronous transceiver, serial peripheral interface and integrated circuit bus, which can process and program complex algorithms and can be well capable of shape recognition, It is capable of shape recognition, straight line recognition, color recognition and other functions. The design of the OpenMV identification results is shown in Fig.10.

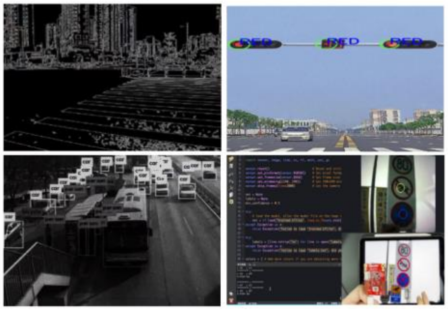


Fig.10 OpenMV identification results

(2) Ultrasonic module first uses DS18B20 digital tempera

ture sensor to measure the ambient temperature, by looking up the pre-established speed of sound - temperature control table to get the current speed of sound; turn on the counter timing, while through the STC16F40K128 PWM module to generate ultrasonic echo, to be echo into the receiving circuit, after shaping and amplification and other hardware processing into the microcontroller. The microcontroller captures the trigger signal of the echo and records the moment of the falling edge; after software filtering and peak time detection algorithm to get the peak moment as the moment of the echo arrival, and finally calculate the distance. The design of the Ultrasonic module test results is shown in Fig.11.

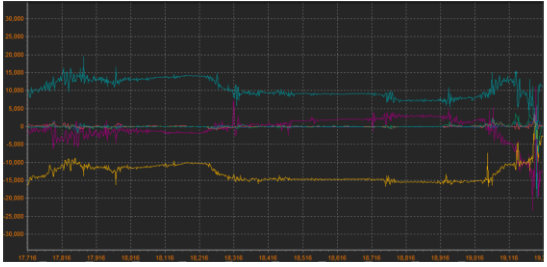
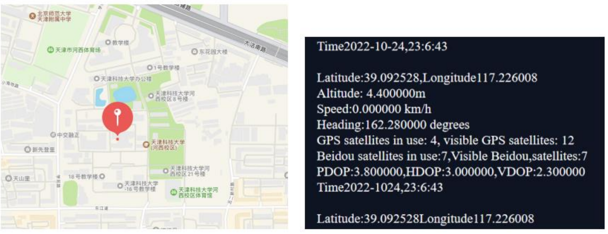


Fig.11 Ultrasonic module test results

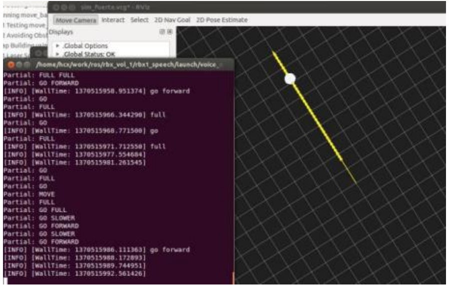
(3) The GPS module consists of four parts: data loading, data pre-processing, calculation and result return, and the GPS receiver interface connects the corresponding positioning module to receive the feedback data in real time. Through the fusion of single-point real-time positioning and post-facto precision single-point positioning settlement solution, the positioning accuracy of the system is greatly improved. In order to enhance the security of the system, RXM/RAW decoding technology is added in the data transmission process to ensure the safety and reliability of the program to the greatest extent.[6] The ability to accurately transmit the location to surrounding vehicles and pedestrians in a timely manner also reduces the security threat to the blind from unexpected situations to a certain extent. The design of the GPS module positioning results is shown in Fig.12.



```
Time2022-10-24,23:6:43

Latitude:39.092528,Longitude117.226008
Altitude: 4.400000m
Speed:0.000000 km/h
Heading:162.280000 degrees
GPS satellites in use: 4, visible GPS satellites: 12
Beidou satellites in use:7,Visible Beidou,satellites:7
PDOP:3.800000,HDOP:3.000000,VDOP:2.300000
Time2022-1024,23:6:43

Latitude:39.092528Longitude117.226008
```

Fig.12 GPS module positioning results

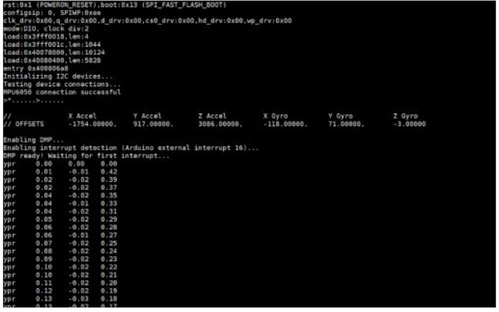
(4) The voice control part is written in C language, the keywords are set during development, and the voice chip extracts the features of the voice and compares them with the keywords. The code is written to send the corresponding data from the serial port to the main control board when the corresponding command is recognized, thus realizing voice control. The voice recognition chip is connected to the MCU and transmits the data in JSON format. After receiving the data, the serial port analyzes it to get the identification code corresponding to the current command and executes the corresponding operation accordingly. The design of the Voice module communication code initialization and operation resultsis shown in Fig.13.



```
Partial: FULL FULL
Partial: GO FORWARD
[INFO] [WallTime: 1370515958.951374] go forward
Partial: GO
Partial: FULL
[INFO] [WallTime: 1370515966.344290] full
Partial: GO
[INFO] [WallTime: 1370515968.771500] go
Partial: FULL
[INFO] [WallTime: 1370515971.712550] full
[INFO] [WallTime: 1370515977.554684]
[INFO] [WallTime: 1370515981.261545]
Partial: GO
Partial: FULL
Partial: GO
Partial: MOVE
Partial: FULL
Partial: GO FULL
Partial: GO SLOWER
Partial: GO FORWARD
Partial: GO SLOWER
Partial: GO FORWARD
[INFO] [WallTime: 1370515986.111363] go forward
[INFO] [WallTime: 1370515988.172893]
[INFO] [WallTime: 1370515989.744951]
[INFO] [WallTime: 1370515992.561426]
```

Fig.13 Voice module communication code initialization and operation results

(5) The gyroscope MPU6050 is used for attitude resolution, it integrates a 3-axis MEMS gyroscope, a 3-axis MEMS accelerometer, and an expandable digital motion processor DMP, which can be connected to a third-party digital sensor, such as a magnetometer, via an I2C interface. After expansion, a 9-axis signal can be output via its I2C or SPI interface.[7] The design of the Gyroscope MPU6050 communication code running results is shown in Fig.14.

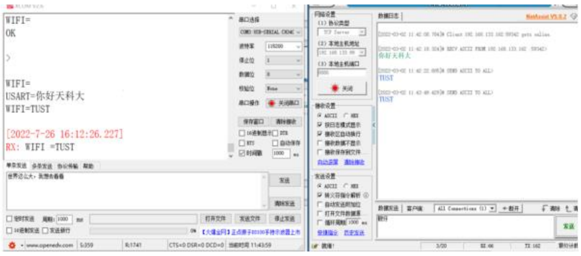


```
rst:0x1 (POWERON_RESET),boot:0x13 (SPI_FAST_FLASH_BOOT)
configsip: 0, SPIWP:0xee
clk_drv:0x00,q_drv:0x00,d_drv:0x00,cs0_drv:0x00,hd_drv:0x00,wp_drv:0x00
mode:DIO, clock div:2
load:0x3fff0018,len:4
load:0x3fff001c,len:1044
load:0x40078000,len:10124
load:0x40080400,len:5828
entry 0x400806a8
Initializing I2C devices...
Testing device connections...
MPU6050 connection successful
>*......>......

//              X Accel         Y Accel         Z Accel         X Gyro          Y Gyro          Z Gyro
// OFFSETS      -1754.00000,    917.00000,      3086.00000,     -118.00000,     71.00000,       -3.00000

Enabling DMP...
Enabling interrupt detection (Arduino external interrupt 16)...
DMP ready! Waiting for first interrupt...
ypr     0.00    0.00    0.00
ypr     0.01    -0.01   0.42
ypr     0.02    -0.02   0.39
ypr     0.02    -0.02   0.37
ypr     0.04    -0.02   0.35
ypr     0.04    -0.01   0.33
ypr     0.04    -0.02   0.31
ypr     0.05    -0.02   0.29
ypr     0.06    -0.02   0.28
ypr     0.06    -0.01   0.27
ypr     0.07    -0.02   0.25
ypr     0.08    -0.02   0.24
ypr     0.09    -0.02   0.23
ypr     0.10    -0.02   0.22
ypr     0.10    -0.02   0.21
ypr     0.11    -0.02   0.20
ypr     0.12    -0.02   0.19
ypr     0.13    -0.03   0.18
```

Fig.14 Gyroscope MPU6050 communication code running results

(6) For the WIFI module, the main processor communicates with the guide cap via the ESP8266, which will print t

he name of the connected wifi and the 8266's own IP address when it starts up, and will always print connected if it is not connected to the server. The design of the WIFI module communication results is shown in Fig.15.



```
WIFI=
OK

>

WIFI=
USART=你好天科大
WIFI=TUST

[2022-7-26 16:12:26.227]
RX: WIFI =TUST
```

```
你好天科大

TUST

TUST
```

Fig.15 WIFI module communication results

4. Conclusion

The intelligent guide cap with integrated multi-sensors and external devices analyzes the demand and experimental conditions and theoretical technology research. This product includes the practical operation through microcontroller, machine vision, Internet of Things, embedded system, voice control, etc. It effectively realizes the information such as the distance status of discerning obstacles and reminding blind people to avoid them in time; prompting blind people the status information of traffic lights at intersections; recognizing crosswalks and guiding blind people along The product is designed to be used in a wide range of applications, such as the identification of zebra crossings and guiding the blind to cross the road safely in the correct direction, voice prompting and other prompting functions, informing the blind of the current status results, and simple human-machine interaction. With the continuous promotion and popularization of this product, more and more blind people are able to go out alone, and the misunderstanding and prejudice of the society towards blind people will be reduced, which makes a contribution to the development of the society.

References

1.Wu Qinqi, Zheng Wenhao, Ye Zhipeng, et al. SMT32-based intelligent guide cane for the blind, *Electronic Fabrication*,2020(13):34-36

2.Shen Zhongkun, Xu Jinjie. Smart wayfinding cart based on OpenMV vision module and MPU6050 angle sensor, *Electronic World*,2022,30(03):8-30+19.

3.Huang Jinjing, He Jinying, Liu Nianghu, et al. Research on intelligent wearable ultrasonic distance measurement and positionable guide belt, *Industrial Control Computer*,2022,35(02):100-101+123.

4.Mi Yuanyuan, Li Guangzong, Qiao Danni. Design of intelligent cart with voice control and automatic obstacle avoidance based on LD3320, *Technology and innovation*,2021(24):36-37+40.

5.Wei Jingting, Chen Liwei, Li Binki, et al. Design of voice-controlled and automatic obstacle avoidance smart car in STM32, *Electronic Testing*, 2019(15):24-25+20.

6.Ding Zulei, Jiang Tianze, Wen Xuping, et al. Research on autonomous motion navigation system based on GPS positioning, *Automation and Instrumentation*,2022,37(08):1-4+29.

7.Zhou Xuhu, Zhang Huawei. Research and implementation of human posture detection and fall determination based on STM32 chip and MPU6050 motion sensor, Compilation of papers from China Medical Equipment Conference and 2021 Medical Equipment Exhibition, 202

Authors Introduction

Ms. Suqing Duan

She is currently pursuing her undergraduate degree at the College 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.

Acquisition of Synergy for Low-dimensional Control of Multi-fingered Hands by Reinforcement Learning

Kazuki Higashi
Graduate School of Engineering Science, Osaka University, 560-8531 Osaka, Japan
E-mail: higashi@hlab.sys.es.osaka-u.ac.jp

Tomohiro Motoda
Graduate School of Engineering Science, Osaka University, 560-8531 Osaka, Japan

Akiyoshi Hara
Graduate School of Information Science and Technology, Osaka University, 565-0871 Osaka, Japan

Kensuke Harada
Graduate School of Engineering Science, Osaka University, 560-8531 Osaka, Japan

Abstract

Synergy is the method that reduces the control inputs of a multi-fingered hand and is utilized for designing underactuated robotic hands and efficient control. Calculating conventional synergies depends on the measured human grasping postures. Therefore, preparing synergies for the not-human-like multi-fingered hands is challenging. We propose a reinforcement learning platform for acquiring synergies of a multi-fingered robotic hand through learning a grasping task. The learning process automatically generates postures for creating synergies so that this system can prepare synergies for any robotic hand. Experiments show that this reinforcement learning platform improves learning tasks and acquires the synergy that is suitable for the learned task.

Keywords: Synergy, Reinforcement Learning, Dimensionality Reduction, Multi-fingered Hand

1. Introduction

Synergy is a practical approach to decrease the control inputs of a robotic system with high degrees of freedom (DoFs) [1]. This unique advantage of synergy makes the multi-fingered hand, which is too complicated for industrial use, easy to control, and more adaptive to various tasks than simple robotic hands. For reducing the control inputs, the synergy is calculated by dimensionality reduction methods such as the principal component analysis and the gaussian process latent variable model. Since these methods are data-driven, many postures of a robotic hand performing tasks must be gathered. Conventional studies generally use human hand postures and transfer the postures to robotic hands [2], [3]. These methods allow for creating a low-dimensional control space for synergy while maintaining the dexterity of human manipulation. Synergies made from human postures are generally applied to humanoid robotic hands. Besides, some methods for making synergies for non-humanoid multi-fingered robotic hands have been proposed. Ficuciello et al. [4] developed a mapping method from human hand posture to robotic hands. However, these methods cannot guarantee synergies to utilize non-humanoid robotic hands' dexterity fully. Because kinematics between the human hand and the non-humanoid hand, such as the number of joints and the dimension, are entirely different. Therefore, more suitable synergies must exist that can fully exploit the non-human robotic hand's kinematics than the synergies made by human postures.

This paper proposes a reinforcement learning (RL) platform to acquire synergies for every kind of robotic hand. First, a multi-fingered hand that acquires synergy through RL is trained on a specific task. As the learning progresses, the postures of successful tasks are accumulated in a database. The principal component analysis is applied to the postures registered in the database to compress the dimension of control inputs. We design a reward for the learning so that the compressed control space can express many postures in which the task can be performed. This allows us to obtain synergy with low dimensional controllability and high task performance. Simulation-based experiments confirm the effectiveness of the method. It is also shown to be effective in improving learning efficiency.

Section 2 describes the details of the method, Section 3 describes the experimental and evaluation methods, Section 4 discusses the experimental results, and Section 5 concludes this paper.

2. Proposed Method

This section explains the details of the proposed RL platform.

2.1 *Reinforcement Learning*

The RL algorithms used in this platform are the Deep Deterministic Policy Gradient (DDPG) method [5], [6], and the Hindsight Experience Replay (HER) [7]. An overview of the learning system is shown in Fig.1. DDPG consists of an actor network that outputs actions and a critic network that computes a Q-function for evaluating actions in a given state. The actor network takes only the state s as input, while the critic network takes the state s and action a as input. These networks consist of an input layer, three all-coupled layers with 256 units in each layer, and an output layer.

Because tasks performed by a multi-fingered hand are represented by continuous values for both actions and states, the number of possible states and actions is significant, and it is often the case that exploration is not rewarded. HER adds experience four times per experience $(s_t \mid\mid g, a_i, r_i, s_{t+1} \mid\mid g)$. The reward r, r' is recalculated for all recorded experiences, along with the calculation of the reward for the successful completion of the task, as described below. Mini-Batch Learning is performed on each network of DDPG using all the experiences thus obtained.

2.2 *Reward system*

Fig. 2 shows the rewards for successful and failed tasks. r_{failure} is a reward in the case of task failure. In the case of task success, we consider a reward r_{success} and a penalty e which is the approximation error of the successful grasping posture projected onto the synergy space. λ means a coefficient of feedback for the approximation error e. *Synergy space* is a principal component space from the dataset of successful grasping postures. We aim to learn a task while increasing the number of postures during task execution that can be accurately represented in the synergy space.

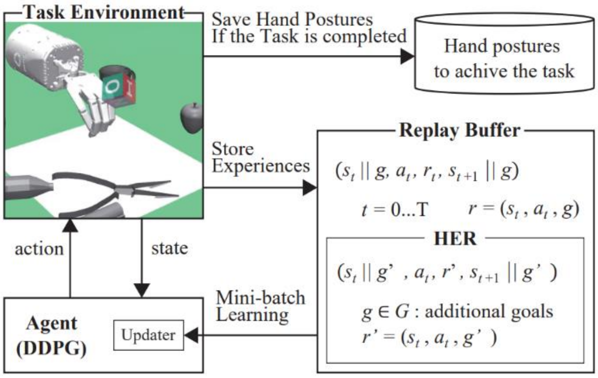

Task Environment
Save Hand Postures
If the Task is completed
Hand postures
to achive the task
Store
Experiences
Replay Buffer
$(s_t \mid\mid g, a_t, r_t, s_{t+1} \mid\mid g)$
$t = 0 \ldots T$
$r = (s_t, a_t, g)$
HER
$(s_t \mid\mid g', a_t, r', s_{t+1} \mid\mid g')$
$g \in G$: additional goals
$r' = (s_t, a_t, g')$
action
state
Mini-batch
Learning
Agent
(DDPG)
Updater

Fig. 1 Overview of the proposed method

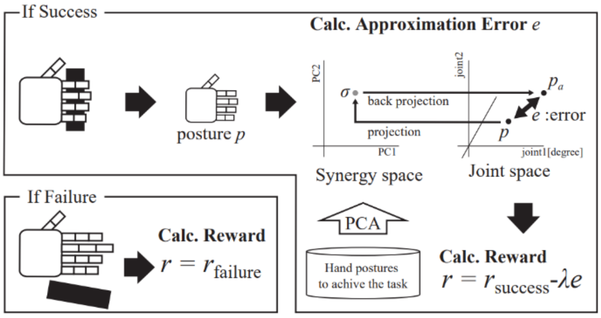

If Success
Calc. Approximation Error e
posture p
PC2
joint2
σ
back projection
p_a
e :error
projection
p
PC1
joint1[degree]
Synergy space
Joint space
If Failure
Calc. Reward
$r = r_{\text{failure}}$
PCA
Hand postures
to achive the task
Calc. Reward
$r = r_{\text{success}} - \lambda e$

Fig. 2 Reward system

3. Experiments

3.1. *Task description*

A grasping task is trained on various objects to compare the results with synergies computed from the postures of humans grasping various objects. The tasks are executed on a physics simulator, Mujoco [8]. The robot hand running in the simulator is a kinematic model like the Shadow Dexterous Hand [9], with 20 actuators driving 24 joints (2 rotational wrist joints and 22 hand joints). In addition, one linear motion joint is added for the vertical motion of the hand, and the robot operated by this system includes 21 actuators and 25 joints. Actions are represented by vectors representing the actuator command values of these 21 joints.

The task is successful if the agent can grasp an object in a random position and orientation on a plane and move the object's center to the goal. Each episode consists of 100 steps, and each step executes a joint command value output from the DDPG and updates the state accordingly. The state is represented in 64 dimensions consisting of joint angles (25 Degrees of Freedoms: DoFs), joint angular velocities (25 DoFs), object position and posture (3 DoFs + 4 DoFs), object velocity and angular velocity (3 DoFs + 3 DoFs) and grasped object ID (1 DoF, seven objects).

As shown in Fig.1, postures that succeed in grasping objects are stored in the dataset. In this experiment, we use the joint angles (25 DoFs) as the type of posture to be preserved. Besides, five principal components will be used for the synergy space.

3.2. *Evaluation*

This experiment evaluates the task success rate during learning and the contribution of each principal component of the synergy. The contribution rate r_i is the ratio of the ith principal component to the total variance and is expressed as follows:

$$r_i = \frac{l_i}{\sum_{i=0}^{N} l_i}$$

l_i means the eigenvalue of the ith principal component. N is the number of the principal components of the synergy.

4. Results

Fig. 3 and Fig. 4 show the transition of the contribution rate (vertical bars) of each principal component of the learning synergy and the task success rate (black line) during the RL with $\lambda = 0, 1$, respectively. The contribution rate is calculated after the number of postures in the dataset exceeds 100. Until then, the penalty term λe in the reward function is zero, and learnings proceed similarly regardless of λ. In both settings, the contribution rate decreases significantly after adding the penalty term and then increases. In Fig. 3, the contribution rate decreases as the episodes pass, whereas in Fig. 4, the contribution rate increases. From these results, the feedback of the approximation error to the learned synergy space makes the synergy able to express grasping postures during the tasks.

An interesting fact is that considering synergy during RL improves both the contribution of synergy and the task success rate. Fig. 3 does not consider synergy, so the task success rate fluctuates around 70%. On the other hand, Fig.4 shows that the task success rate almost converges to 100%. This is because the grasping posture generated by the network follows the synergy, which includes a dataset of successful grasping postures. Thus, the task can be successful with a high probability.

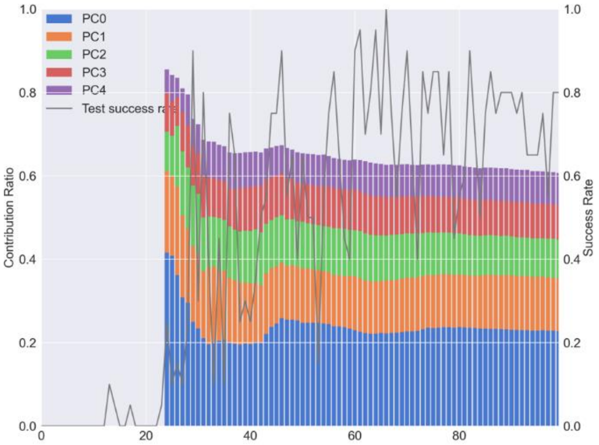


Fig. 3 Transition of contribution rate and task success rate during RL ($\lambda = 0$)

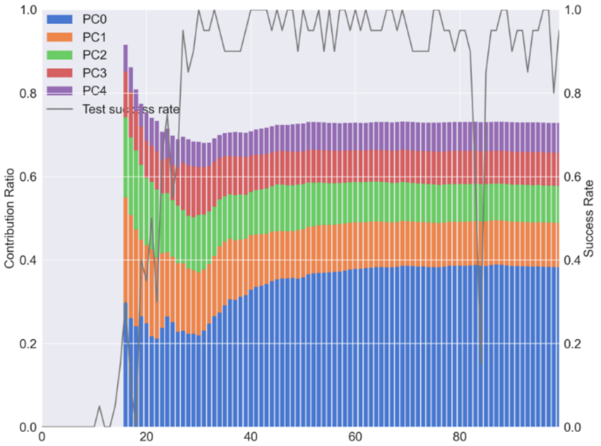


Fig. 4 Transition of contribution rate and task success rate during RL ($\lambda = 1$)

5. Conclusion

In this paper, we propose a platform for learning synergy among various robot models. We conduct experiments in which synergy is acquired in an object-grasping task using a humanoid hand. The results show that this RL platform realizes to make a synergy that can express various grasping postures and improves task success rate during RL.

This proposed system can learn not only grasping tasks but also more complex tasks such as tool manipulation that are appropriate for multi-fingered hands. In the future, we will compare the synergies obtained in this study with those obtained from humans, study the reward function for better synergy acquisition, train and assign synergy to various multi-fingered hand models, and verify the efficiency of RL when the acquired synergies are used as an action space.

References

1. Santello, Marco, Martha Flanders, and John F. Soechting. "Postural hand synergies for tool use." Journal of neuroscience 18.23 (1998): 10105-10115.
2. Ficuciello, Fanny. "Synergy-based control of underactuated anthropomorphic hands." IEEE Transactions on Industrial Informatics 15.2 (2018): 1144-1152.
3. Rodriguez, Diego, et al. "Learning Postural Synergies for Categorical Grasping through Shape Space Registration." 2018 IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids). IEEE, 2018.
4. Ficuciello, Fanny, et al. "A model-based strategy for mapping human grasps to robotic hands using synergies." 2013 IEEE/ASME International Conference on Advanced Intelligent Mechatronics. IEEE, 2013.
5. Silver, David, et al. "Deterministic policy gradient algorithms." International conference on machine learning. PMLR, 2014.
6. Lillicrap, Timothy P., et al. "Continuous control with deep reinforcement learning." arXiv preprint arXiv:1509.02971 (2015).
7. Andrychowicz, Marcin, et al. "Hindsight experience replay." Advances in neural information processing systems 30 (2017).
8. Todorov, Emanuel, Tom Erez, and Yuval Tassa. "Mujoco: A physics engine for model-based control." 2012 IEEE/RSJ international conference on intelligent robots and systems. IEEE, 2012.
9. The Shadow Robot Company. https://github.com/shadow-robot

Authors Introduction

Mr. Kazuki Higashi

He received the M.E. degree from the Graduate School of Engineering Science, Osaka University, Japan in 2020. Currently, he is a Ph.D. candidate in Graduate School of Engineering Science, Osaka University, Japan. His research interests include low-dimensional control (Synergy) for multi-fingered hands and development of robotic hand utilizing synergy.

Mr. Tomohiro Motoda

He received the M.E. degree from the Graduate School of Engineering Science, Osaka University, Japan in 2020. Currently, he is a Ph.D. candidate in Graduate School of Engineering Science, Osaka University, Japan. His research interests deep learning in grasping and manipulation, motion planning, and manufacturing automation.

Mr. Akiyoshi Hara

He received the M.E. degree from the Graduate School of Information Science and Technology, Osaka University, Japan in 2020. Currently, he is a Ph.D. candidate in Graduate School of Information Science and Technology, Osaka University, Japan. His research interests include transcranial current stimulation and reflex response, and tele-existence.

Dr. Kensuke Harada

He received the Ph.D. degree in engineering from the Graduate School of Mechanical Engineering, Kyoto University, Kyoto, Japan, in 1997. He is currently a Professor with the Graduate School of Engineering Science, Osaka University, Osaka, Japan. His current research interests include the mechanics and control of humanoid robots and robotic hands.

Error Recovery Techniques 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 years, working robots have been used in various fields, from manufacturing industries to domestic living spaces. An increasing number of difficult tasks are performed by robots, introducing a necessity for error recovery techniques. Our proposed error recovery technique is based on a method of returning to the process before the step in which the failure occurred and starting over. Of course, in practice, not only such a backward recovery but also a forward recovery that moves forward even after a failure occurs is used. In this study, we consider various paths from failure occurrence to recovery execution.

Keywords: error recovery, task stratification, error classification, robotic manufacturing plant

1. Introduction

In recent years, the speed of automated plants that use robots has increased. However, the possibility of errors occurring in automated operations has increased. Therefore, studying the recovery process after an error occurs remains important.[1]-[5]

We have studied the systematization of error recovery theory for several years and have previously proposed an error recovery method based on the concepts of task stratification and error classifications.[6]-[9] The main part of the method consists of the basic elements of sensing, modeling, planning, and execution sequences (Fig. 1). If an error occurs while working, the process proceeds to the recovery part,inwhich the cause of the error is estimated, the error is classified, the system is corrected, and the process is run with the corrected system and improved reliability.

Our proposed error recovery technique returns to the step before the one in which the failure occurred and starts from there. Of course, in practice, not only this type of backward recovery is used, but also forward recovery, which moves forward after a failure occurs. In this study, we consider various paths from failure occurrence to recovery execution.

The remainder of the paper is organized as follows. The concept of skills, which are motion primitives, is described in Section 2. The fundamental techniques for error recovery are presented in Section 3. Multiple possibilities of various recovery procedures are considered in Section 4, and examples of each are presented in Section 5. Section 6 concludes this paper.

2. Concept of skill

In this section, we briefly describe a skill, which we define as the unit of motion.[10]-[12]

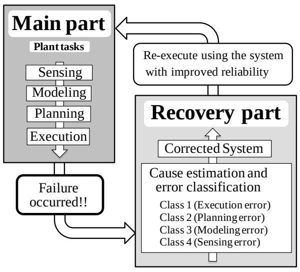
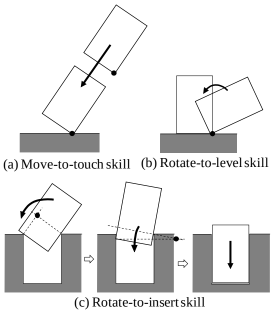

Main part
Plant tasks
Sensing
Modeling
Planning
Execution
Failure occurred!!
Re-execute using the system with improved reliability
Recovery part
Corrected System
Cause estimation and error classification
Class 1 (Execution error)
Class 2 (Planning error)
Class 3 (Modeling error)
Class 4 (Sensing error)

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

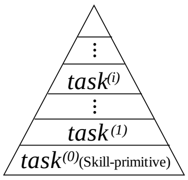

task(i)
task(1)
task(0)(Skill-primitive)

Fig. 3 Hierarchy of tasks

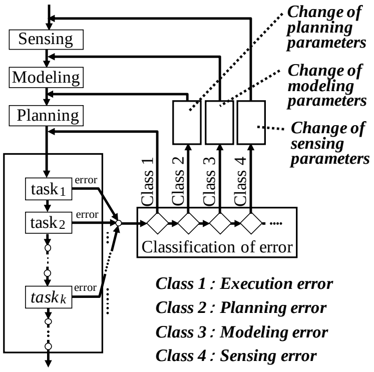

Change of planning parameters
Change of modeling parameters
Change of sensing parameters
Sensing
Modeling
Planning
task1
task2
taskk
error
error
error
Class 1
Class 2
Class 3
Class 4
Classification of error
Class 1 : Execution error
Class 2 : Planning error
Class 3 : Modeling error
Class 4 : Sensing error

Fig. 4 Fundamental process flow with error recovery

2.1. *Skill primitives*

The three skill primitives, shown in Fig. 2 can be considered by analyzing human behavior. A task such as ‘assembly’ consists of these three skill primitives and the resemble skill primitives.

2.2. *Stratification of tasks*

The hierarchy of robotic manipulation tasks is illustrated in Fig. 3. Neglecting the servo layer, the first skill layer consists of movement units, such as the important skill primitives shown in Fig. 2.

3. Error recovery

Errors can occur in equipment for various reasons. In this section, we outline the concept of error classification and the proposed error recovery method.[6]-[9]

3.1. *Error classification*

Errors can be divided into four types, based on their possible causes: execution, planning, modeling, and sensing errors (Fig. 4).[6]-[9]

3.2. *Error recovery based on classification*

When an error occurs, the cause is first estimated, and then a suitable correction based on the estimated cause is implemented. In many methods, the executive process returns to the previous step and the task restarts from that step (Fig. 4).[6]-[9] For small-scale errors, the process returns to the previous step in the lowest layer of the task hierarchy, while for large-scale errors, the process returns to the previous step in the high-ranking hierarchy layers (Figs. 4 and 5). The probability of the same error occurring is reduced because of the correction that has been executed.

4. Various recovery pathways

As described in Section 3, recovery reruns the process by returning to the step previous to that in which an error occurs. In other words, the fundamental recovery process is executed backward. However, backward recovery is not the only conceivable option; other recovery procedures, such as forward recovery, without backward recovery are also possible. The occurrence of an error may affect the surrounding environment, such as the destruction of the arrangement or shape of certain objects.

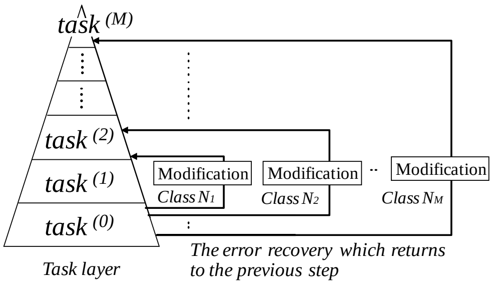

task (M)
task (2)
task (1)
task (0)
Modification
Class N1
Modification
Class N2
Modification
Class NM
Task layer
The error recovery which returns to the previous step

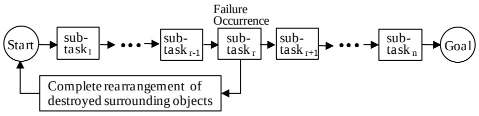
Fig. 5 The expression of task stratification and the process flow of the error recovery

Then, it could happen that the same recovery process as the previous one is not possible without modification when returning to the original process. In such cases, partial restructuring of planning is required. Of course, sensing and modeling are also necessary, even for local planning.

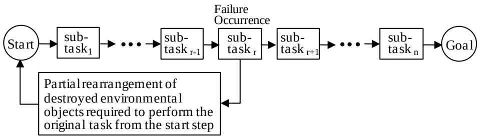
In the remainder of this section, we consider the degree of destruction of the environment surrounding the object and consider multiple possibilities for various recovery procedures.

4.1. *Recovery Formula I (RF-I): Exact original repetition*

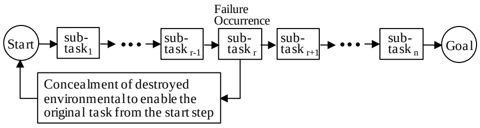
In this method, the environment is completely restored to its original state and work is resumed using the same process as the original starting point (Fig. 6 (a)). The process is performed by replacing the original object, or part of it, with a new object, if necessary. The same


Failure Occurrence
Start
sub-task$_1$
sub-task$_{r-1}$
sub-task$_r$
sub-task$_{r+1}$
sub-task$_n$
Goal
Complete rearrangement of destroyed surrounding objects

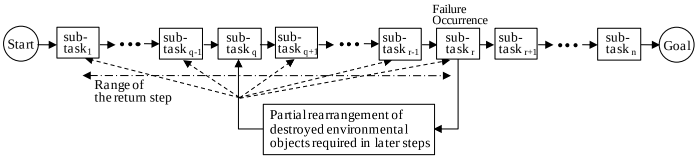
(a) (RF-I) Exact Original Repetition


Failure Occurrence
Start
sub-task$_1$
sub-task$_{r-1}$
sub-task$_r$
sub-task$_{r+1}$
sub-task$_n$
Goal
Partial rearrangement of destroyed environmental objects required to perform the original task from the start step

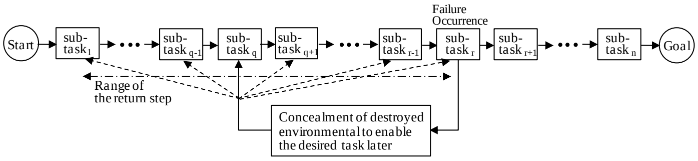
(b) (RF-II) Incomplete Original Reproduction from Start Step


Failure Occurrence
Start
sub-task$_1$
sub-task$_{r-1}$
sub-task$_r$
sub-task$_{r+1}$
sub-task$_n$
Goal
Concealment of destroyed environmental to enable the original task from the start step

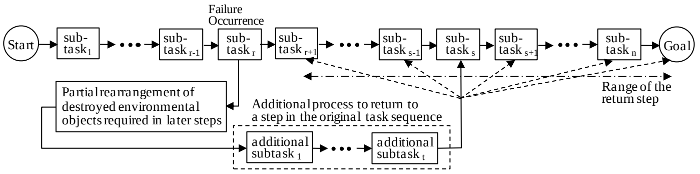
(c) (RF-III) Original Repetition from Start Step with concealment


Failure Occurrence
Start
sub-task$_1$
sub-task$_{q-1}$
sub-task$_q$
sub-task$_{q+1}$
sub-task$_{r-1}$
sub-task$_r$
sub-task$_{r+1}$
sub-task$_n$
Goal
Range of the return step
Partial rearrangement of destroyed environmental objects required in later steps

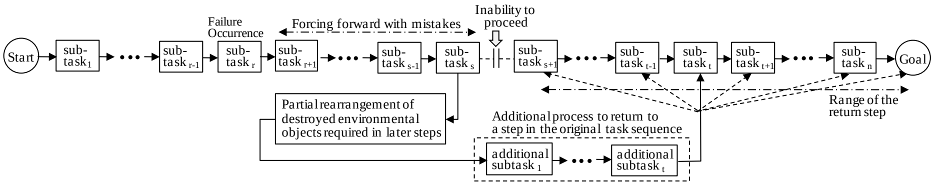
(d) (RF-IV) Incomplete Original Reproduction from Step along the Way


Failure Occurrence
Start
sub-task$_1$
sub-task$_{q-1}$
sub-task$_q$
sub-task$_{q+1}$
sub-task$_{r-1}$
sub-task$_r$
sub-task$_{r+1}$
sub-task$_n$
Goal
Range of the return step
Concealment of destroyed environmental to enable the desired task later

(e) (RF-V) Original Repetition from Step along the Way with concealment


Failure Occurrence
Start
sub-task$_1$
sub-task$_{r-1}$
sub-task$_r$
sub-task$_{r+1}$
sub-task$_{s-1}$
sub-task$_s$
sub-task$_{s+1}$
sub-task$_n$
Goal
Range of the return step
Partial rearrangement of destroyed environmental objects required in later steps
Additional process to return to a step in the original task sequence
additional subtask$_1$
additional subtask$_t$

(f) (RF-VI) Recovery through a different process


Failure Occurrence
Forcing forward with mistakes
Inability to proceed
Start
sub-task$_1$
sub-task$_{r-1}$
sub-task$_r$
sub-task$_{r+1}$
sub-task$_{s-1}$
sub-task$_s$
sub-task$_{s+1}$
sub-task$_{t-1}$
sub-task$_t$
sub-task$_{t+1}$
sub-task$_n$
Goal
Range of the return step
Partial rearrangement of destroyed environmental objects required in later steps
Additional process to return to a step in the original task sequence
additional subtask$_1$
additional subtask$_t$

(g) (RF-VII) Continuation of work, including consideration of any recovery process that may be required later

Fig. 6 Various error recovery paths

approach applies to all recovery formulae outlined in this section (RT-II to RT-VII). Note that any replacement results in additional costs.

4.2. *Recovery Formula II (RF-II): Incomplete original reproduction from start step*

In this method, the environment is not restored to its original state, but is restored to a level that does not cause any problems for subsequent processes; the process is restarted from the original starting point using the same process (Fig. 6 (b)).

4.3. *Recovery Formula III (RF-III): Original repetition from start step with concealment*

In this method, the surrounding environment that has been broken from the occurrence of an error is concealed, and it is superficially restored; work resumes from the original starting point using the same process (Fig. 6 (c)). By concealing the broken environment, no problems occur in subsequent operations. However, if the concealed area reappears for any reason, the appearance of the object may be spoiled. For customer products, for example, this may result in a decrease in value.

4.4. *Recovery Formula IV (RF-IV): Incomplete original reproduction from step along the way*

In this method, the environment is not restored to its original state, but is restored to a level that does not cause any problems for subsequent processes; the process is resumed from a point in the middle of the original process, using the same process (Fig. 6 (d)).

4.5. *Recovery Formula V (RF-V): Original repetition from step along the way with concealment*

In this method, the surrounding environment that has been damaged by an error occurrence is concealed, restoring it superficially; work is resumed from a step in the middle of the original process, using the same process (Fig. 6 (e)). The recovery work procedure and commercial value are subject to the same concerns as those in RF-III.

4.6. *Recovery Formula VI (RF-VI): Recovery through different processes*

In this method, the environment is not returned to its original state, but instead is transformed into a state in which there are no problems with subsequent work. Then, a process that is different from the original process is executed to reach the goal (Fig. 6 (f)). Partial planning becomes necessary for the process to either achieve its goal or to return to a previous step in the middle of the original process; sensing and modeling accompany local planning. The final target state is the same as that of the original. Therefore, if the originally assumed process is restored to the middle, no new planning is required from the middle of the recovery process.

4.7. *Recovery Formula VII (RF-VII): Continuation of work, including consideration of any recovery process that may be required later*

In this method, work is continued for as long as possible, even if an error occurs. Corrections are made to the environment when the work becomes unsustainable, thereby continuing the work (Fig. 6 (g)). Local planning must be performed to execute the modification; sensing and modeling are performed for local planning. Such a forward recovery process is symmetrical to the backward recovery methods outlined in RF-I to RF-VI.

5. Illustrative examples of various recovery pathways

This section provides examples of each method outlined in Section 4.

5.1. *RF-I: Exact original repetition*

An example of RF-1 can be found in “Recovery Type I” in reference [9], where error recovery is considered in an assembly task, in which a hook is stuck to a plate by screws.

5.2. *RF-II: Incomplete original reproduction from start step*

An example of RF-II can be considered in that presented for RF-I, because the outlined example does not consider whether the environment was destroyed. If the environment was destroyed, it is simply restored to such a level that the recovery task can be performed without

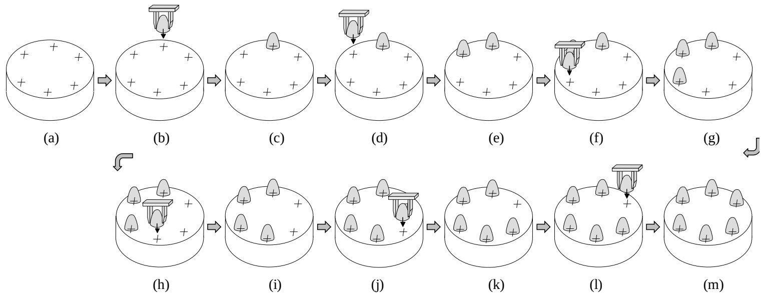


Fig. 7 Cake-making task

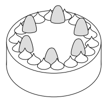


Fig. 9 Whole Cake with Cream Decoration

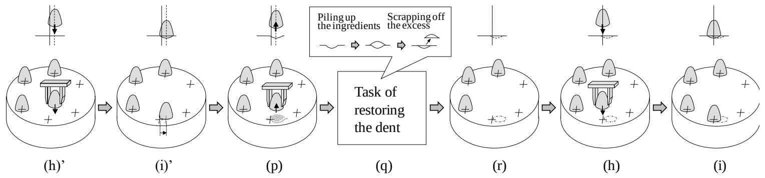


Fig. 8 Error and recovery process in cake-making task

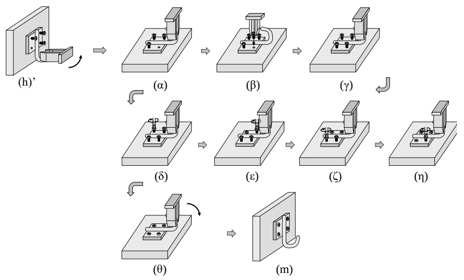


Fig. 10 The recovery task performed by laying the object on its side

problems, and then the original task is rerun from the start. Additionally, reference [7] provides an example of recovery for an error that damages the placement of objects around it.

5.3. *RF-III: Original repetition from start step with concealment*

Examples of RF-III are shown in conjunction with R-V, which are described later.

5.4. *RF-IV: Incomplete original reproduction from step along the way*

An example of RF-IV can be found in "Recovery Type II" in reference [9].

5.5. *RF-V: Original repetition from step along the way with concealment*

Consider the cake-making task shown in Fig. 7. To obtain a cake that is cut with equal divisions from a whole cake, decorations, such as strawberries, must be placed at the correct positions in terms of the angles and position of the cutting device. Let us consider the error in which the position of the decoration is shifted to the side, as shown in Fig. 8. In this case, the following recovery task is performed: the decoration is lifted, the dent in the cream is filled and finished with a scraper, and the decoration is placed back on the cake. Note that the repair task becomes more difficult if the top surface of the cake is decorated with cream before the strawberries are placed on top, as shown in Fig. 9.

The error recovery approach of RF-V is characterized by concealing the failure such that it is not visible. If the process returns to the starting step and the same process can be re-executed, the approach corresponds to RF-III. In the cake-decorating example, implementing RF-III is easy if the placement of the first strawberry fails; however, if the placements of the second or subsequent strawberry fail, returning to the starting state is difficult.

5.6. *RF-VI: Recovery through different processes*

Let us consider an example of RF-VI using reference [9], which considers an assembly task in which a hook is stuck to a plate by screws. The recovery task is performed by laying the object on its side so that the temporarily inserted screw could be stabilized, as shown in Fig. 10. In RF-VI, a process different from the original process is executed to reach the original goal. Note that this method cannot always be used, as problems could arise, such as the inability to lay the object on its side.

5.7. *RF-VII: Continuation of work, including consideration of any recovery process that may be required later*

An example of RF-VII can be found in "Recovery Type III" in reference [9].

6. Conclusion

If an error occurs during work, the process advances to a recovery part. Many types of recovery processes can be selected using our proposed error recovery method, which is based on both task stratification and error classification. In this paper, we considered the various paths from failure occurrence to recovery execution, and not only recovery of backward type but also recovery of forward type were shown.

There are various recovery routes, and various methods are conceivable for which route to choose. In future work, we would like to determine the optimal recovery process. The method of process selection will change depending on the type of robotic system considered; however, deriving a systematized method remains important.

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. 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.
11. 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.
12. 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.

Flexible Assembly System with Stiffness Switching Joint

He Maike, Tokuo Tsuji, Naoki Ichikawa, Takuro Sawada, Tatsuhiro Hiramitsui, Hiroaki Seki
Kanazawa University, Kanazawa, Ishikawa, Japan
E-mail: hemaike@stu.kanazawa-u.ac.jp
www.kanazawa-u.ac.jp

Abstract

In this research, we aim to construct a flexible assembly system and realize precise assembly work by using a flexible-rigidity switchable joint for the wrist part of an industrial robot. One of the basic works in assembly work is Peg-in-Hole task. We install magnets in the joints and change the structure of the joints by changing the state of attraction and repulsion of the magnets. The joint has two states, which are low stiffness and high stiffness. The experimental data shows that switching stiffness improves the success rate.

Keywords: Peg in Hole, Precise assembly, Stiffness switching, Robot joint.

1. Introduction

For many years, the introduction of industrial robots has made progress in automation at production sites. Along with this, the production system is also shifting from line production to cell production, but in many cases, cell production relies on human labor. Especially in the assembly work performed by industrial robots, it is not easy to automate precision fitting, and even today, many assembly tasks rely on human labor. One of the basic works in assembly work is the fitting (Peg-in-Hole) work of the peg to the hole. In order to realize the mating work, it is necessary to suppress relative errors in the position, posture, and insertion direction between the peg and the hole within the clearance. Therefore, in automated assembly work, peripheral devices such as vision sensors, force sensors, and jigs are used to achieve high-precision positioning. In the case of preparing a jig for each part, in addition to the increase in time and money costs, it is not possible to respond flexibly to variable-variety variable-volume production. Amid the problems of labor shortages and rising labor costs due to population decline in Japan, realization of a flexible and rapid assembly system that can handle variable-variety variable-volume production is desired. We decided to achieve this goal by making a joint with variable stiffness. The joint has two states, which are low stiffness and high stiffness. The joint is equipped with magnets and springs, and the structure of the joint is changed by changing the state of attraction and repulsion of the magnets.

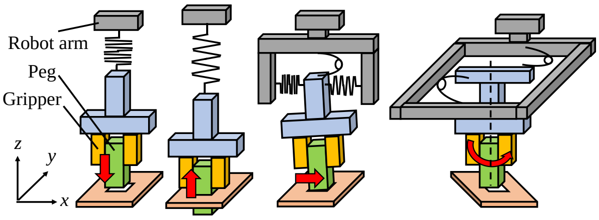
2. Related works

A hardware-based method that elastically supports workpieces and a software-based control method that uses data from sensors are being researched as methods [1],[3],[4] for realizing fitting work. The former is a passive alignment method, and the latter is an active alignment method. A hardware-based method is Remote Center Compliance (RCC) proposed by Whitney[1]. RCC is a function in which the center of elasticity is near the tip of the peg. To realize the function of RCC, the position of the elastic center must be changed according to the size of the workpiece. In addition, many elastic devices such as RCC are designed assuming mating work in which insertion is performed in the direction of gravity. Therefore, an elastic device that can handle various types of workpieces and arbitrary insertion directions is desired. Kim[2] proposed a hybrid variable stiffness actuator (HVSA), which is a variable stiffness unit design. The proposed HVSA is composed of a control module and a drive module. By controlling the gears in the control module, the position and stiffness of a joint can be controlled.

3. Fabrication of joints

3.1. *Purpose*

In fitting work, it is necessary to keep relative errors in the position, posture, and insertion direction between peg holes within the clearance. If an error exceeds the clearance, the peg and hole will come into contact and exert forces and moments on each other. If the error is large, the frictional force increases, and bite may occur. The purpose of making the wrist part of the compliant joint is to passively absorb this error. On the other hand, a weak point of flexible joints is that the posture of the hand is susceptible to external forces. The hand position is not reproducible with the same robot motion. In order to solve these problems, the joint used for the wrist of the robot to be manufactured can be switched between a low-rigidity state and a high-rigidity state with one joint. In this low-stiffness mode of the joint, as shown in Fig.1. The joint is given 6 degrees of freedom to achieve the mating and withdrawal motions of the peg.



(a) z-axis translation (b) x-, y-axes rotation (d) z-axis rotation
(c) x-, y-axes translation

Fig.1 DOF of the compliance at the low stiffness mode

3.2. *Structure*

The developed joint is shown in Fig.2, and the structure of the joint is shown in Fig.3. It becomes possible to displace in the direction when a force or moment acts on the part. The degree of freedom of translational and rotational motion on the x and y axes can be obtained by forming the entire circumference of the two upper and lower shafts with silicone. When a force or moment acts in these directions, the silicone deforms, and the position and posture of the gripper change, making it possible to passively absorb the error between the peg and the hole during mating work. Above the plate connected to the peg is a guide connected to a spring. The space between the plate and the guide is covered with silicone, and when it receives a rotational moment around the z-axis, the silicone deforms, and the gripper rotates. In addition, the N-pole magnet attached to the rotating part around the z-axis faces the N-pole magnet of the flat plate, and the flat plate receives the repulsive force of the magnet. At this time, the plate stops at the position where the magnetic force, the elastic force, and the gravitational force of the gripper and the grasped object are balanced. When an external force in the z-axis direction is applied to the gripper, the joint is displaced in the direction in which the force acts. In the high-rigidity mode, the N-pole magnet of the rotating part faces the S-pole magnet of the flat plate, and the flat plate is attracted to the rotating part by receiving the attraction force of the magnet. Conical pins and holes on the bottom of the flat plate mesh with each other to restrict all movements and suppress the effects of external forces. Switching between both stiffness modes is performed using a servomotor (DynamixelXM430-W210, Robotis).

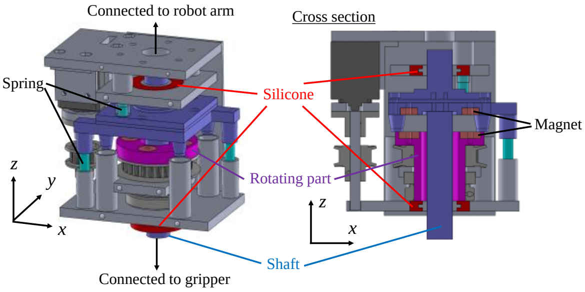


Fig.2 Developed stiffness switching joint

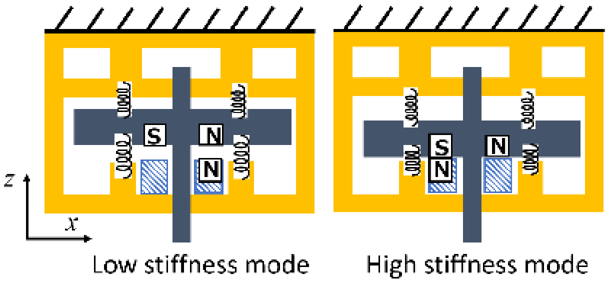


Fig.3 Schematic of both stiffness modes

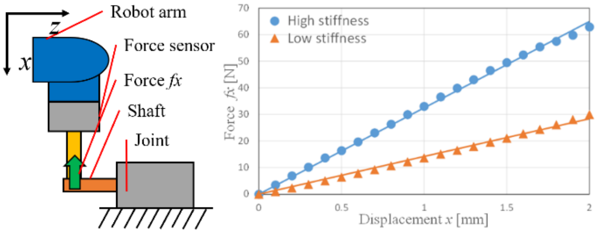


Fig.4 Schematic and result of the stiffness measurement

4. Evaluation experiment

4.1. *Stiffness measurement*

As shown in Fig.4, we measured the stiffness in the x-axis direction in both stiffness modes. Fig.4 shows the measurement results of the force when pushing the joint axis using a robot arm connected to a 6-axis force sensor. The magnitude of the force required for displacement in the x-axis direction in the high-rigidity mode is about 2.3 times that in the low-rigidity mode.

4.2. *Fitting work*

As shown in Fig.5, fitting work was performed using the manufactured joint. The joint is connected to the servo motor and the peg (Φ16.88) in that order, and it is fitted into the hole (Φ17.02) on the board. [5] showed that the success rate is improved by utilizing the rotation of the peg in the fitting work of a robot with a flexible wrist. The motion was 10 N, and the peg was pressed against the hole on the plate, rotated clockwise for 3 seconds, and inserted 50 mm when the fitting was successful. In the low-stiffness mode, the joint has an axial stiffness that displaces about 2mm at 10N. Experiments were performed in both stiffness modes by changing the position of the peg with respect to the hole. Fig.6 shows the position of the peg when the peg insertion was successful and failed. The origin of the peg position is the position where the peg can be inserted into the hole without interference. The frictional force acting in the direction of insertion of the peg is examined. Fig.7 shows the relationship between the distance from the origin of the peg and the magnitude of the friction force in both stiffness modes. The frictional force shown is the maximum value at 50mm insertion. In the high-rigidity mode, the greater the positional error, the greater the frictional force generated, requiring a large insertion force. On the other hand, in the low-rigidity mode, the friction force is suppressed to 1N or less at any positional error, which is lower than in the high-rigidity mode. This is because the joint compliance reduced the drag force between the peg and the hole. Therefore, in the low-rigidity mode, the required insertion force is reduced, and it is possible to cope with cases where the friction coefficient between peg holes is high.

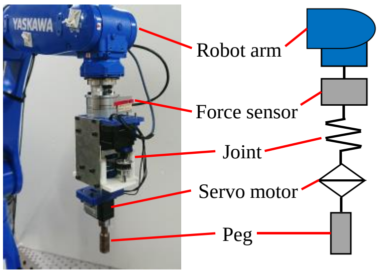
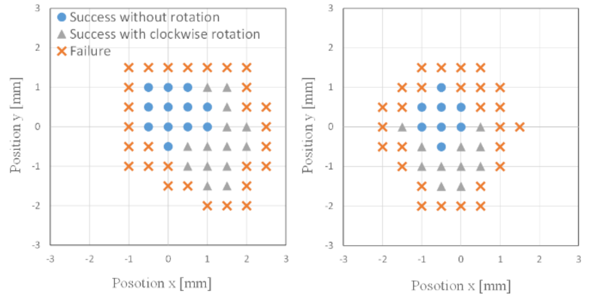


Fig.5 Schematic of the Peg-in-Hole experiment



(a) High stiffness mode (b) Low stiffness mode

Fig.6 Success area of the Peg-in-Hole

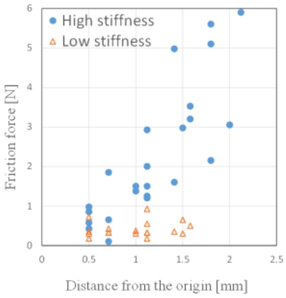
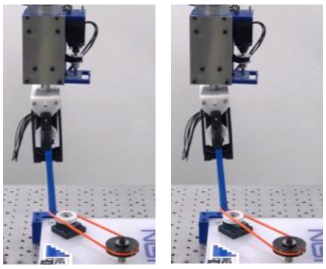


Fig.7 Friction force of the Peg-in-Hole

4.3. *Belt assembly work*

We assemble the belt with the proposed joint. A parallel gripper is connected to the joint, and a belt with a diameter of 4 mm and a circumference of 400 mm is assembled on pulleys with a center distance of 135 mm. Fig.8 shows the moment when the belt is put on the second pulley in both stiffness modes. In the low-rigidity mode, the center position of the joint and the gripper fingertip position deviate greatly due to the displacement of the joint axis due to the belt tension. Therefore, in order to change the position and posture of the gripper fingertip, it is necessary to move the robot considering the deformation of the joints. In the high-rigidity mode, it is possible to work while suppressing the effects of joint deformation due to tension.



(a) High stiffness mode (b) Low stiffness mode

Fig.8 Belt assembly experiment

5. Conclusion

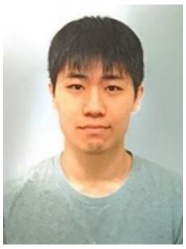
In this research, we developed a flexible/rigid switchable joint for the wrist of a robot to realize precise assembly work. The effectiveness was verified with the actual assembly work.

References

1. Whitney DE. Quasi-static assembly of compliantly supported rigid parts. J Dyn Sys Meas Control-Trans ASME. 1982;104:65–77.
2. Kim B S, Song J B. Design and control of a variable stiffness actuator based on adjustable moment arm[J]. IEEE Transactions on Robotics, 2012, 28(5): 1145-1151.
3. Jasim IF, Plapper PW, Voos H. Position identification in force-guided robotic peg-in-hole assembly tasks. Procedia Cirp. 2014;23:217–222.
4. Park H, Bae J-H, Park J-H, et al. Intuitive peg-in-hole assembly strategy with a compliant manipulator. In: IEEE ISR 2013. 2013 Oct. p. 1–5.
5. M. Tennomi et al, "Development of assembly system for quick and low-cost installation", Advanced Robotics, Vol. 34, No. 7-8, pp.531-545, 2020

Authors Introduction

Mr. Maike He

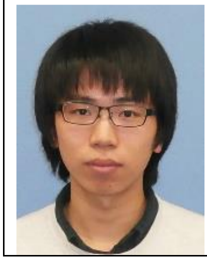


He received his Master of Engineering degree from the Mechanical Science and Engineering, Kanazawa University, Japan in 2021. He is currently a Doctoral course student in Kanazawa University, Japan.

Prof. Tokuo Tsuji

He is an associate professor of Institute of Science and Technology at Kanazawa University in Japan. He received his D. Eng. degree from Kyushu University in 2005. His research interests include manipulation, image processing, and robotic hand.

Mr. Naoki Ichikawa



He received his Master of Engineering degree from the Mechanical Science and Engineering, Kanazawa University, Japan in 2021.

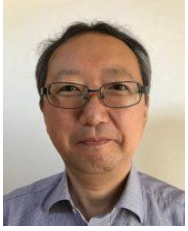
Mr. Takuro Sawada

He received his Master of Engineering degree from the Mechanical Science and Engineering, Kanazawa University, Japan in 2022.

Prof. Tatsuhiro Hiramitsu

He is an assistant professor at the Institute of Science and Engineering, Kanazawa University in Japan. He received Dr E. degrees from school of engineering, Tokyo Institute of Technology, Japan in 2019. His research interest is in the soft structure mechanisms for robotic systems.

Prof. Hiroaki Seki



He is a Professor of Institute of Science and Technology at Kanazawa University in Japan. He received his D. Eng. degree in Precision Machinery Engineering from the University of Tokyo in 1996. His research interest includes novel mechanism and sensing in robotics and mechatronics.

Pose Detection for Flexible-Indefinite Objects using Pseudo-Bone Data

Yuma Yoshimoto
Department of Creative Engineering, National Institute of Technology (KOSEN), Kitakyushu College,
5-20-1, Shii, Kokura-Minami-ku, Kitakyushu-shi, Fukuoka, Japan

Hakaru Tamukoh
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu-ku, Kitakyushu 808-0196, Japan

E-mail: yoshimoto@kct.ac.jp
https://researchmap.jp/yoshimoto116

Abstract

This paper proposes a method for recognizing the poses of a flexible-indefinite object. Some flexible-indefinite objects, such as fried shrimp, differ between individuals. Therefore, it is difficult to estimate the pose of these foods by point cloud fitting or other methods. We propose "pseudo-skeleton" data for these objects. Pseudo-skeleton data consists of "key-points," which are joints, and "bones," which connect between key-points. For example, fried shrimp are given 3 key-points; "head," "belly," and "tail." In addition, the bones that connect them are given. In the experiment, objects pose recognition model based on the human pose recognition model trains pseudo-skeleton data of fried shrimp. We confirmed that the model estimates the poses in the images.

Keywords: Pseudo-skeleton, Flexible-indefinite objects, Pose estimation model, Key-points and bones

1. Introduction

The population decline in recent years has introduced robots into factories. The industrial field is especially true. However, the food industry, such as making lunch boxes, is handled because the robots need to grasp some flexible-indefinite objects, such as fried shrimp. It is not easy to estimate the posture of flexible-indefinite objects.

In a lunch factory, food ingredients are stacked in bulk. To recognize industrial objects stacked in bulk, the robots generally fit the design data, such as computer aided design (CAD) data, to objects in bulk. However, we cannot prepare the design data of food ingredients. The method is not suitable for the situation. As a method for such cases, camera data is input into a neural network, and the grasping point is estimated directly. However, this method is unsuitable for cases such as foods that are required to dish up. The reason is that robots cannot estimate the posture of foods using the method.

In this study, we propose a method for estimating the posture of flexible-indefinite objects such as food ingredients using pose estimation neural network models for humans. In the methods, the structures of flexible-indefinite objects are described, such as "pseudo-skeleton," using a combination of key-points and bones.

2. Related Works

2.1 *Skeleton Data*

Skeleton data consists of "key-points," which are joints, and "bones," which connect between key-points. Figure 1 shows an example of human skeleton data. In the data, head, hand, arm, and etcetera, are key-points. In addition, bones connect the key-points.

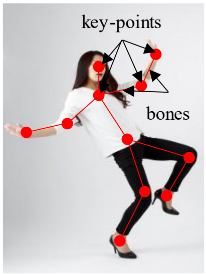


Fig. 1. Human skeleton data

2.2 *Pose Estimation Neural Network Models*

Pose estimation models are generally used to estimate the human pose. These methods are two types: bottom-up type [1] and top-down type [2]. Bottom-up type first search all key points in an image. Then, there are matched and connected for each person. The top-down type first detects people using an object detection algorithm. Then, the models estimate key-points and bones for each person. The latter is more robust to occlusion and motion blur than the former. The latter is used in this paper.

3. Proposed Methods

In this study, we apply the pose estimation method, which is generally used for humans, to flexible-indefinite objects, such as food ingredients. In the method, the "pseudo-skeleton" of these objects are defined for estimation poses. Fig. 2. Fried shrimp's pseudo-skeleton shows an example of this method. In the figure, the proposed method is applied to fried shrimp. In this example, "head," "belly," and "tail" are defined as key-points. The connections between the head and belly and between the belly and tail are defined as bones.

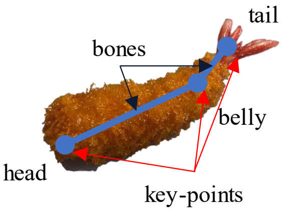


Fig. 2. Fried shrimp's pseudo-skeleton

4. Experiments

4.1 *Setup*

The experiment was conducted to evaluate the proposal. We prepared 325 pairs of fried shrimp's images, pseudo-skeleton data, and bounding boxes. Figure 3 shows an example of datasets. The 300 pairs were used as training data, and 25 pairs as test data. In addition, a combination of ResNet50 [3] and Top-down Heatmap Simple Head was used as a pose estimate model [2].

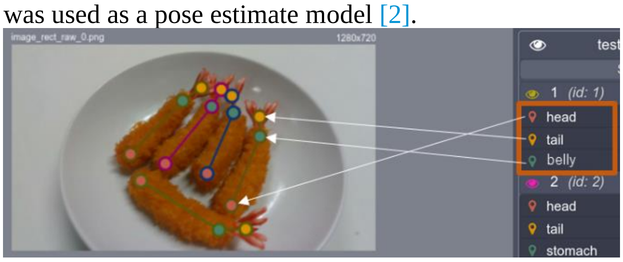


Fig. 3. Datasets

4.2 *Results and Discussion*

Table 1 shows the experimental results. We calculated the mean average precision (mAP), which is used to confirm dataset effectiveness. The same model for estimating the human poses, is included as a comparison [2]. The results show that the proposed method can estimate the pseudo-skeleton of flexible-indefinite objects at the same level as the human skeleton. Figure 4 shows an image of the estimate results. The result shows that the pseudo-skeleton of a fried shrimp can be inferred.

Table 1. Experimental results

Training Data	mAP
Fried Shrimp's Pseudo-Skeleton (Ours)	**0.795**
Human Skeleton [2]	0.724

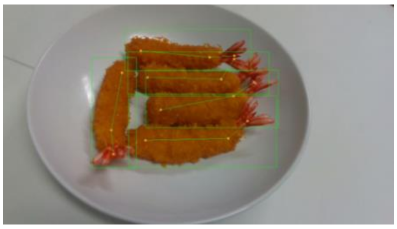


Fig. 3. Estimate results

5. Conclusion

We propose a pseudo-skeleton method for the pose estimate of flexible-indefinite objects in the paper. In the proposed method, the "pseudo-skeleton" of these objects are defined for estimation poses. Experimental results show that the model, which was trained pseudo-skeleton, can estimate the posture of flexible-indefinite objects at the same level as a human skeleton. In the future, we will incorporate this system into a robot to verify grasping.

References

1. J. Li, S. Bian, A. Zeng, C. Wang, B. Pang, W. Liu, C. Lu, "Human Pose Regression with Residual Log-likelihood Estimation," 2021 International Conforence on Computer Vision (ICCV 2021), pages 11025-11034, 2021.
2. B. Xiao, H. Wu, Y. Wei, "Simple Baselines for Human Pose Estimation and Tracking," 2018 European Conference on Computer Vision (ECCV 2018), pages 466-481, 2018.
3. K. He, X. Zhang, S. Ren, J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016), pages 770-778, 2016.

Authors Introduction

Asst. Prof. 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.

Prof. Hakaru Tamukoh

He received the B.Eng. degree from Miyazaki University, Japan, in 2001. He received the M.Eng. and the PhD 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 2006 to 2007. He was an Assistant Professor of Tokyo University of Agriculture and Technology, from 2007 to 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.

Object Search and Empty Space Detection System for Home Service Robot

Tomoya Shiba
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka 808-0196, Japan

Ono Tomohiro
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka 808-0196, Japan

Hakaru Tamukoh
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka 808-0196, Japan

E-mail : Tomoya.shiba627@mail.kyutech.jp
https://www.brain.kyutech.ac.jp/~tamukoh/

Abstract

Home service robots have pick-and-place tasks to grasp and carry objects. When performing pick-and-place tasks, robots need to process to search for objects in the case of missing a target object. We propose approaches for the robot to search for a target object from a shelf and a method for selecting empty spaces to move off-target objects. The proposed method is based on two recognition models to select a space and plan motions. The object is to be moved and the space is selected based on the size and position of empty spaces. In the experiments, the robot planned actions to search for a target object from a shelf and to place an object near a group of similar objects. We used these in the RoboCup@Home competition to evaluate the effectiveness of the proposed method.

Keywords: Home service robot, mobile manipulator, Object detection, RoboCup@Home

1. Introduction

Home service robots have begun to attract increasing attention due to the needs of an aging society with a declining birth rate [1],[2],[3]. The essential functions of home service robots are object recognition, picking and placing, recognition of people and environment, and interaction with people. We have conducted various studies using TOYOTA’s home service robot HSR [4] and our robot Exi@ [5]. So far, Various studies on the social implementation of service robots for home use have been presented through competitions [6],[7],[8],[9]. These studies have mainly proposed tidying environment recognition, object recognition at edge devices, efficient motion planning, and pick-and-place strategies.

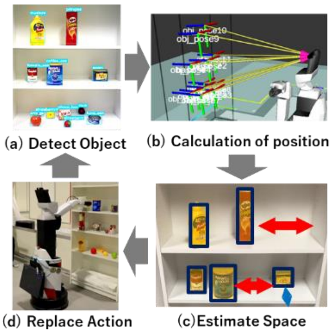


Fig. 1. System Overview

Home service robots use these proposed techniques to perform various tasks, such as shelf placement, room cleaning, and restaurant clerking. Common to all these tasks is searching for the target object.

If the person cannot find the target object, they perform actions such as moving an obstacle object. The person moves an object with predictability, and we wondered if it would be possible for the robot.

This paper proposes a method for the robot to search the shelf. This method defines space at the back end of object recognition systems. And the robot moves obstructive objects into empty spaces.

We implement these methods on TOYOTA's home service robot HSR. As a validation, we used them in RoboCup@Home [10], a competition for home service robots, to verify their effectiveness.

2. Previous research

2.1. *Object recognition & estimate position*

There are various methods for mobile manipulators to detect the location of objects and pick them. In our previous research, we created one that utilizes YolactEdge [11] and point cloud information. This recognition technology can estimate the position and posture of an object [12].

2.2. *Problem*

There are many objects in the home environment, including furniture such as desks and shelves. Because of this, objects overlap, creating blind spots where the desired object cannot be detected, and the desired object may not be located.

3. Proposed Method

3.1. Motion to move obstacle objects

First, we propose a simple method for locating objects on the shelf. In this model, the size of the shelf is assumed to be known. The theme of this model is how to grab an object in the way. Then, it is how to move it to an available location. (Fig. 1.)

(a). Object recognition by YolactEdge
(b). Calculation of position from depth information
(c). Estimating empty and unseen areas based on the size and position of an object
(d). Move larger-sized objects to a space.
(e). Object detection again (to (a))

3.2. Estimate space

The second method is estimating the number of available placement locations using semantic segmentation to detect planes. We use Estimating Available Placement Locations [13] (Fig.2.).

(a). Planar detection by semantic segmentation
(b). Clop and convert planar point cloud information
(c). Morphology transformations
(d). Blurred to remove information on the edges

The robot must decide where to place the object from the candidate regions. The robot scores each point in the candidate region and places the object at the highest rated point. In the experiment, the robot selects the point in the candidate region that is the nearest to it (Fig.3.). In the case of (Fig. 3.), the point closest to the robot was the bottom left of the output image in (Fig. 2.). We defined a distance map such that the bottom left is the highest point, and the score decreases with each Euclidean distance away from the robot.

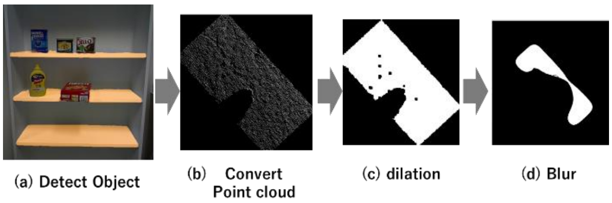


Fig. 2. Estimate empty space

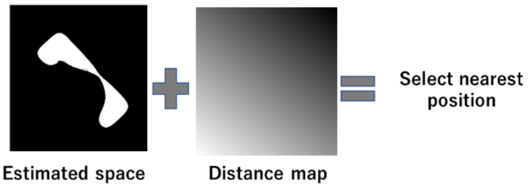


Fig. 3. Select the available position

4. Experimental

4.1. *RoboCup@Home*

The system's validity was verified and experimented with in this section through a RoboCup@Home competition. The conditions and results are summarized.

4.2. *Go and get it task*

We have tested the effectiveness of an object search task by running it on a shelf at the World Robot Challenge 2020 (WRC2020) [14], the RoboCup Asia-Pacific 2021

(RCAP2021) [15], and RoboCup Japan Open 2021(RCJ2021) [16].

The task performed was to search for the target object on a shelf. The object used is the YCB object [17]. It searches for the target object on a shelf with 15~20 objects. An example of object placement is shown in (Fig. 4).

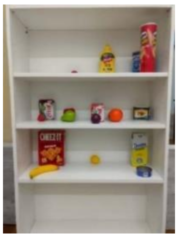


Fig. 4. YCB object placed on the shelf [14]

4.3. *Storing Groceries task*

In RoboCup@Home 2022 [18], the task of placing an object on a shelf was performed. The method of estimating space described in 3.2 has been verified.

An example of object placement is shown in (Fig. 5). From the initial state, four to six objects are placed on the shelf. The robot must place objects on the shelf near objects of the same category. The robot can place up to 5 objects on the shelf.

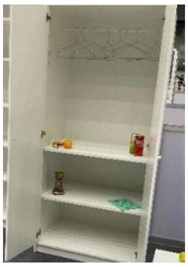


Fig. 5. Storing Groceries task shelf

5. Experimental result

5.1. *Go and get it task*

We tested the effectiveness of this task six times in WRC2020 and four times in RCAP2021 and RCJ2021 in each competition. The results are summarized in Table 1. As a result, the target object was found 12 out of 14 times with a success rate of more than 85%.

Table 1. Success rate in grasping the target

Competition	Challenge [times]	Success [times]	Success rate [-]
WRC 2020	6	6	1.00
RCAP 2021	4	3	0.75
RCJ 2021	4	3	0.75
total	14	12	0.86

5.2. *Storing Groceries task*

We tested the operation of the estimation of empty space in four trials at RoboCup 2022. Ten empty space estimations were performed in the four trials. The results are summarized in Table 2.

The task has two evaluations: the number of successful attempts to place an empty space and the number of times an object could be placed near the same category. The number of successful attempts to place the object in an empty space was 10 out of 10, but the overall correct classification rate was about 60%.

There are cases of classification errors. This is because the robot tries to place objects in larger spaces. The robot tends to select open spaces between object categories.

Table 2. Storing Groceries grasping task

Trials	Challenge [times]	Success [times]	successful classification [times]	successful classification rate
First	0	-	-	0
Second	0	-	-	0
Third	5	5	2	0.4
Fourth	5	5	4	0.8
total	10	10	6	0.6

Notice : The first and second attempts did not proceed to object grasping due to robot errors.

6. Discussion

In this study, we proposed search method from shelf for home service robots and conducted experiments in RoboCup tasks. The experimental results show that our method is effective in that.

However, when classification is required, as in the storing groceries task, it makes an incorrect select. This is because it selects an area that is easier to place a larger area without even an object of the same category.

To solve this problem, it is necessary to define a new region for each category and select a space from among them.

7. Conclusion

In this study, we proposed and applied a search method that fully utilizes object rearrangement and spatial detection to search for objects on the shelf of a home service robot.

The proposed method is simple and can be implemented at a later stage of existing recognition methods, and its effectiveness was evaluated in competitions such as RoboCup@Home. In the future, we aim to incorporate overlap recognition methods into the recognition itself, such as UOAIS [19], and to define candidate regions for each category for accurate placement selection.

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. “RoboCup Federation website” (2022 12/15 accessed)
11. 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.
12. Ono, T., Kanaoka, D., Shiba, T., Tokuno, S., Yano, Y., Mizutani, A., ... & Tamukoh, H. (2022). Solution of World Robot Challenge 2020 Partner Robot Challenge (Real Space). Advanced Robotics, 36(17-18), 870-889.
13. Ono T, Tamukoh H. A fast pick-and-place method for home service robots using 3d point clouds. Proceedings of International Conference on Artificial Life and Robotics. 2020;25:195-196.
14. “World Robot Summit website” (2022 12/15 accessed)
15. “RoboCup Asia-Pacific website” (2022 12/15 accessed)
16. “RoboCup JapanOpen website” (2022 12/15 accessed)
17. Calli, B., Walsman, A., Singh, A., Srinivasa, S., Abbeel, P., & Dollar, A. M. (2015). Benchmarking in manipulation research: The YCB object and model set and benchmarking protocols. arXiv preprint arXiv:1502.03143.
18. “RoboCup 2022 official website” (2022 12/15 accessed)
19. Back, S., Lee, J., Kim, T., Noh, S., Kang, R., Bak, S., & Lee, K. (2022, May). Unseen object amodal instance segmentation via hierarchical occlusion modeling. In 2022 International Conference on Robotics and Automation (ICRA) (pp. 5085-5092). IEEE.

Authors Introduction

Mr. Tomoya Shiba

He received the B.Eng. degree from National Institute of Technology, Kagoshima College, Japan, in 2021.
He is currently in a Master’s 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. Tomohiro Ono

He received the B.Eng. degree from National Institute of Technology, Ube College, Japan, in 2018. He received the M.Eng. from Kyushu Institute of Technology, Japan, in 2020. He is currently in a Ph.D. student in the graduate school of Life Science and Systems Engineering, Kyushu Institute of Technology. Since 2020, he has also been a research fellow of the Japan Society for the Promotion of Science (JSPS). His research interest includes image processing, motion planning and domestic service robots. He is a student member of RSJ.

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 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.

Robust Classification Model with Multimodal Learning for Home Service Robots

Ikuya Matsumoto
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

Hakaru Tamukoh
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology, 2-4 Hibikino,
Wakamatsu, Kitakyushu, 808-0196, Japan
E-mail:matsumoto.ikuya585@mail.kyutech.jp, kanaoka.daiju327@mail.kyutech.jp, tamukoh@brain.kyutech.ac.jp
https://www.lsse.kyutech.ac.jp/english/

Abstract

We propose an auxiliary data stream structure as a robust classification model. The model treats one modal as the main input and the other modals as supports. The model chooses how much of the sub-modal is used for classification. We experimented with two and three modal inputs. Moreover, we added pseudo-shadows to the visual information for the experiment with three modal inputs. In all experiments, our proposed model improves the accuracy and robustness to environmental disturbances by using multiple modals.

Keywords: Multimodal learning, Classification, Home service robot, Gate structure

1. Introduction

In recent years, service robots have attracted attention owing to the shortage of workers caused by the aging population[1]. Among them such robots, home service robots work in human living environments to assist people. We expect home service robots to perform domestic tasks in the home, such as cleaning up rooms[2][3][4][5][6]. Object classification is one of the functions required for home service robot operation. Currently, object classification uses visual information for home service robots. However, recognition accuracy decreases if the conditions for capturing images are poor (e.g., shadows, blurring). Multimodal learning is one of the solutions to this problem. The method combines several different types of information, such as image and tactile[7]. The application of multimodal learning to object classification will solve the abovementioned problem (e.g., operating in a poor visual environment).

We propose a novel multimodal object classification model in which friction and pushing modals to assist with the visual modal.

2. Related Work

2.1. *Multimodal learning*

Multimodal learning is a learning method that combines multiple types of modals, such as images, text, and voice. This learning method produces more accurate output than unimodal learning and has been applied in various fields, such as speech recognition and disease prediction[8][9].

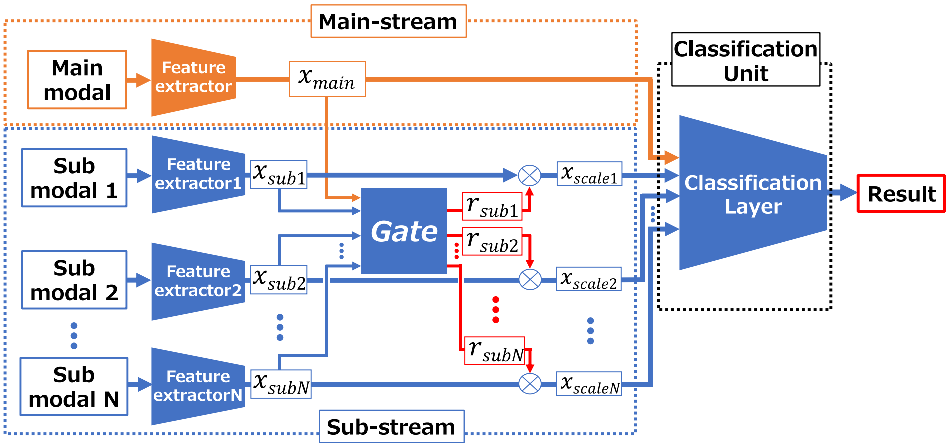


Fig. 1. Multimodal learning with auxiliary data stream model

2.2. *Multimodal learning without gate structure (Gao's model)*

Gao et al. conducted object classification experiments by using images and tactile information[10]. Gao's model was generally more accurate than the unimodal model. However, for some objects, multimodal was degraded accuracy than unimodal. We need to devise a structure for input rather than simply inputting multimodal.

2.3. *Deep gated multimodal learning (DGML)*

Anzai et al. proposed a gate structure for multimodal learning to estimate the pose of a grasping object using images and tactile inputs[11]. The gate structure calculates reliability values that adjust the effect of each modal on the overall recognition result. The sum of the reliability values for the image and haptic is one, and a modal with a higher reliability value is more effective for the estimation. Anzai et al. showed that gate structures are effective for multimodal learning. Therefore, gate structures can be applied to multimodal object classification to improve accuracy.

3. Proposed Model

We propose multimodal learning with an auxiliary data stream model shown in Figure 1. The proposed model is treated as one modal as the main input, whereas the other modals are used to support it. The proposed model can input multiple sub-modals, while maintaining the dominance of the main-modal.

3.1. *Auxiliary data stream model*

The proposed model consists of a mainstream, a substream, and a classification unit. The mainstream and a substream extract features using a feature extractor that is appropriate for the input modal. The extracted main-modal features (x_{main}) and sub-modal features ($x_{sub1}, x_{sub2}, \dots, x_{subN}$) are input to the gate, which calculates the sub-modal reliability value ($r_{sub1}, r_{sub2}, \dots, r_{subN}$). The substream scales the sub-modal features by multiplying the calculated reliability values by the sub-modal features. Finally, the main-modal features and the scaled sub-modal feature (x_{scale}) input to the classification unit output the results.

3.2. *Gate structure*

The gate structure used in the proposed model is shown in Figure 2. The fully connected (FC) layers ($FC0, FC1, \dots, FCN$) reduce the number of feature dimensions extracted from each modal to one ($x_{\mathrm{FC1}}, x_{\mathrm{FC2}}, \dots, x_{\mathrm{FC}N}$). The features were concatenated and input into FC. The FC layer and the sigmoid function calculated the reliability values of the sub-modals. The reliability value was obtained through learning.

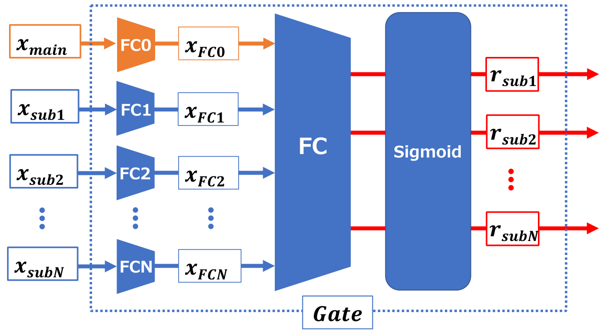


Fig. 2. Gate structure. The reliability value can be in the range of [0, 1]. The larger value, the more effective the modal is for classification.

4. Experiments

4.1. *Dataset*

We used the LMT haptic texture database for this experiment[12]. The dataset contains 10 training and 10 test data per class. We used visual, friction, and pushing modals in the experiment. The details of each modal are presented in Table 1. The visual modal is RGB images taken without flash. The friction modal is time series data of three axes (x, y, z) acquired by sliding the tactile sensor, and the pushing modal is time series data of one axis (z) acquired by pushing the tactile sensor.

Table 1. Data details

Modal	Shape
Visual	480 × 320 [px]
Friction	3 axes (x, y, z) × 48000 [steps]
Pushing	1 axis (z) × 1601 [steps]

4.1.1. *Preprocess*

We preprocessed each modal using the following procedure. The visual modal compressed the image size. The friction and pushing modals were downsampled to remove unstable data.

- Visual
  (i) Resized images to 50 × 50 [px].
  (ii) Normalized the images to [0 , 1].

- Friction
  (i) Skipped 3000 [steps] at the beginning as the data are unstable.
  (ii) Normalized data to [0, 1] for each axis.
  (iii) Down sampled 45000 [steps] to 300 [steps] using an anti-aliasing filter.

- Pushing
  (i) Skipped 101 [steps] at the beginning as the data are unstable.
  (ii) Normalized data to [0, 1] for each axis.
  (iii) Down sampled 1500 [steps] to 300 [steps] using an anti-aliasing filter.

4.2. *Evaluation experiment with two modals*

4.2.1. *Overview*

We conducted 108 class classification experiments using visual and friction as modal inputs. In this experiment, visual is the main-modal, and friction is the sub-modal. Figure 3 shows the model used in this experiment. The visual feature extractor was a CNN, and the friction feature extractor was a 3D CNN.

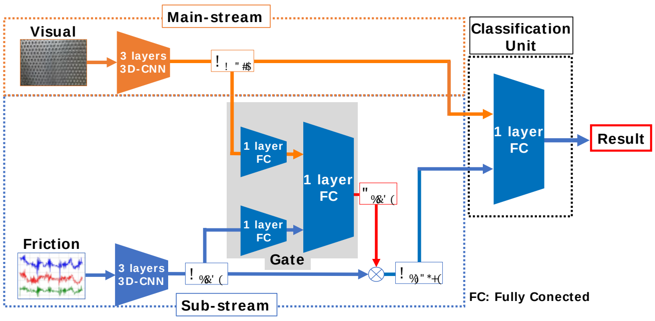


Fig. 3. Proposed model for two modals

4.2.2. *Result*

The experimental results are shown in Table 2. The proposed model has a gated structure and main and sub-stream, which allowed it to achieve an accuracy higher than that of other models. The proposed model is more accurate than DGML, and the structure that gives the main and sub-functions to the modal is effective.

Table 2. Results of the with two modals

Model	Accuracy [%]
Gao's	93.0
DGML	94.0
Ours	94.9

Table 3. Network design for 3 modals experiments

	Layer	In	Out	Filter Size	Activation Function
Visual Feature Extractor	1st 2D Conv.	3	10	(5, 5)	ReLU
	2nd 2D Conv.	10	30	(5, 5)	ReLU
	3rd 2D Conv.	30	108	(3, 3)	ReLU
	Adaptive Avg.Pool2D	108	108	-	-
Friction Feature Extractor	1st 3D Conv.	3	10	(1, 1, 6)	ReLU
	2nd 3D Conv.	10	30	(1, 1, 3)	ReLU
	3rd 3D Conv.	30	108	(1, 1, 3)	ReLU
	Adaptive Avg.Pool3D	108	108	-	-
Pushing Feature Extractor	1st 3D Conv.	1	10	(1, 1, 6)	ReLU
	2nd 3D Conv.	10	30	(1, 1, 3)	ReLU
	3rd 3D Conv.	30	108	(1, 1, 3)	ReLU
	Adaptive Avg.Pool3D	108	108	-	-
Gate	1st $\text{FC}_{\text{visual}}$	108	1	-	-
	1st $\text{FC}_{\text{friction}}$	108	1	-	-
	1st $\text{FC}_{\text{pushing}}$	108	1	-	-
	2nd FC	3	2	-	Sigmoid
Classifier	FC	324	108	-	-

4.3. *Evaluation experiment with three modals*

4.3.1. Overview

We conducted 108 class classification experiments using visual, friction, and pushing as the modal inputs. In this experiment, visual was the main-modal, and the others were sub-modal. The model used in this experiment is shown in Figure 4, and the network design is presented in Table 3. The feature extractor for pushing was a 3D CNN. A classification experiment was also conducted by adding pseudo-shadows to the visuals to simulate a real environment.

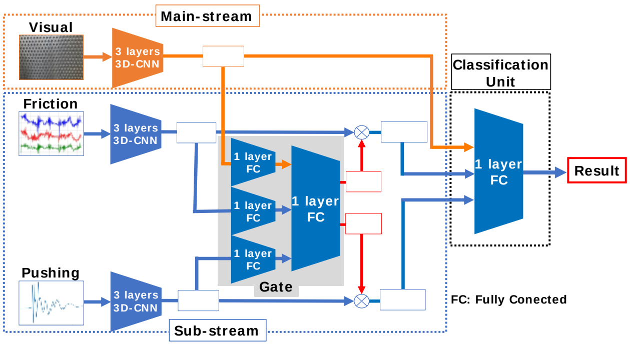


Fig. 4. Proposed model for three modals

4.3.2. *Result*

Table 4 summarizes the experimental results. The proposed model achieved higher recognition accuracy than Gao's model. Our model has a smaller loss of accuracy than Gao's model because the useful modals can be combined in pseudo-shadowed visual experiments.

Table 4. Results of the with three modals

Model	Accuracy [%] (w/o shadows)	Accuracy [%] (w/ shadows)	Amount of decrease
Gao's	97.7	97.0	−0.7
Ours	98.1	97.6	−0.5

5. Conclusion

We proposed an auxiliary data stream model as a robust object classification model and conducted validation experiments in this study. The proposed model achieved higher accuracy than the conventional model, and the experiments simulated real-world conditions also showed high accuracy and robustness of the model.
However, to apply the proposed model to a home service robot, it is necessary to validate the model using actual

data acquired by the robot. In addition, although we focused on shadows in the real environment to verify the robustness of the proposed model, it is necessary to verify the model against other disturbances, such as blurring and backlighting. In future works, we will attempt to implement and evaluate the proposed model on a home service robot.

References

1. New Energy and Industrial Technology Development Organization (NEDO), Future Robot Market Forecasts Released, (Accessed 2022-01-23).
2. Tomohiro Ono, Daiju Kanaoka, Tomoya Shiba, Shoshi Tokuno, Yuga Yano, Akinobu Mizutani, Ikuya Matsumoto, Hayato Amano, and Hakaru Tamukoh, "Solution of World Robot Challenge 2020 Partner Robot Challenge (Real Space)," Advanced Robotics, Vol. 36, Issue 17-18, pp. 870-889, 2022
3. Yuma Yoshimoto, Hakaru Tamukoh, "FPGA Implementation of a Binarized Dual Stream Convolutional Neural Network for Service Robots," Journal of Robotics and Mechatronics, Vol. 33, No. 2, pp. 386-399, 2021.
4. Yuichiro Tanaka, Takashi Morie, Hakaru Tamukoh, "An amygdala-inspired classical conditioning model on FPGA for home service robots," IEEE Access, Vol. 8, pp. 212066-212078, November 2020.
5. Yutaro Ishida, Takashi Morie, Hakaru Tamukoh, "A hardware intelligent processing accelerator for domestic service robots," Advanced Robotics, Vol. 34, Issue 14, pp. 947-957, June 2020.
6. Yutaro Ishida, Hakaru Tamukoh, "Semi-Automatic Dataset Generation for Object Detection and Recognition and its Evaluation on Domestic Service Robots," Journal of Robotics and Mechatronics, Vol. 32, No. 1, pp. 245-253, 2020.
7. Denis Ivanko, Alexey Karpov, Dmitrii Fedotov, Irina Kipyatkova, Dmitry Ryumin, Dmitriy Ivanko, Wolfgang Minker, and Milos Zelezny, "Multimodal Speech Recognition: Increasing Accuracy Using High Speed Video Data." Journal on Multimodal User Interfaces, Vol. 12, No. 4, pp. 319-328, 2018.
8. Hiroshi Kawasaki, "Clinical Multimodal Data Analysis (in Japanese), " Japanese Journal of Allergology, Vol. 69 No.8. pp. 708-709, September 2020.
9. Youssef Mroueh, Etienne Marcheret, and Vaibhava Goel. Deep Multimodal Learning for Audio-Visual Speech Recognition. In 2015 IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 2130–2134, 2015.
10. Yang Gao, Lisa Anne Hendricks, Katherine J. Kuchenbecker, and Trevor Darrell. Deep Learning for Tactile Understanding From Visual and Haptic Data. IEEE International Conference on Robotics and Automation, Vol. 2016-June, pp. 536–543, 2016.
11. Tomoki Anzai and Kuniyuki Takahashi. Deep Gated Multi-modal Learning: In-hand Object Pose Changes Estimation using Tactile and Image Data. IEEE International Conference on Intelligent Robots and Systems, pp. 9361–9368, 2020.
12. Chair of Media Technology Technical University of Munich. LMT Haptic Texture Database. (Accessed 2021-10-26).

Authors Introduction

Mr. Ikuya Matsumoto

He received the B.Eng. degree from Kyushu Institute of Technology, Japan, in 2022. He is currently in a Master's degree student the graduate school of Life Science and Systems Engineering, Kyushu Institute of Technology. His research interest includes multimodal learning and domestic 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 in a Ph.D. student in the graduate school of Life Science and Systems Engineering, Kyushu Institute of Technology. His research interest includes object recognition, multimodal learning, and domestic service robots. He is a student member of IEEE.

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 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..

Flexible Human-Robot Interaction in Domestic Environment Using Semantic Map

Yuga Yano
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu，Fukuoka 808-0196, Japan.

Yukiya Fukuda
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu，Fukuoka 808-0196, Japan.

Tomohiro Ono
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu，Fukuoka 808-0196, Japan.

Hakaru Tamukoh
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu，Fukuoka 808-0196, Japan.

E-mail: yano.yuuga158@mail.kyutech.jp
https://www.brain.kyutech.ac.jp/~tamukoh/

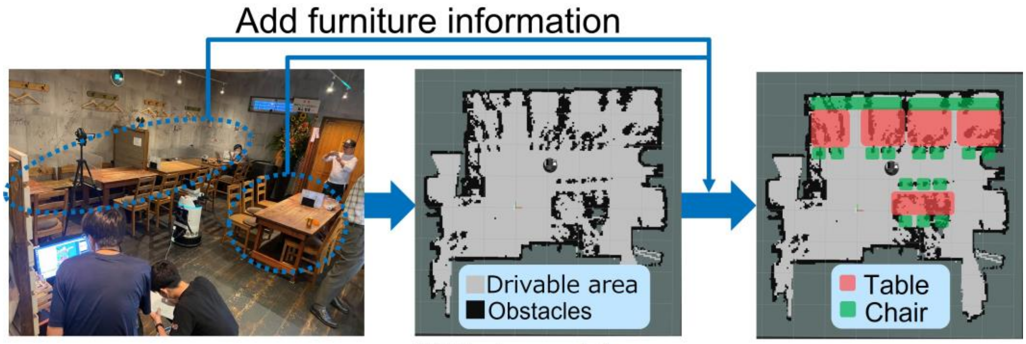
Abstract

We propose an efficient semantic map to realize flexible human-robot interaction (HRI) in domestic environments. Our proposed map is created from an output of Simultaneous Localization and Mapping and already-known environmental information such as furniture and room. In this study, we evaluated the effectiveness of our proposed method on two benchmark tests for HRI in RoboCup@Home held in Bangkok in 2022. In the RoboCup@Home, we employ 3D human recognition to apply our proposed map to HRI, such as "find and offer an empty seat." Our proposed method had the best score of all teams on both tests. The results of our experiments are available at https://youtube.com/playlist?list=PLfbN50Mwh2DG3OPDeCHo4TNuyrU4qYCrJ.

Keywords: Human-robot interaction, semantic mapping, Speech recognition, Human recognition

1. Introduction

In recent years, service robots have been widely used in domestic environments. Typical examples are nursing care robots, cleaning robots, and food delivery robots in restaurants [1], [2], [3], [4], [5] . However, these robots are still limited in their functions and cannot perform flexible actions. Therefore, home service robots that can realize flexible actions with Human-Robot Interaction (HRI) are needed. Estimating the surrounding environment and self-position is necessary to achieve autonomous mobility for home service robots. simultaneous localization and mapping (SLAM) is commonly used for creating an environmental map and self-localization simultaneously. Figure 1(b) shows the example of created environmental map by SLAM. However, SLAM only estimates the surrounding environment in terms of binary values, obstacle or not. Therefore, the robot cannot perform flexible actions such as guiding a person to an empty seat. To address this issue,


Add furniture information
Drivable area
Obstacles
Table
Chair

(a) Actual environment.

(b) Environmental map created by SLAM.

(c) Semantic map.

Fig. 1 An example of semantic map.

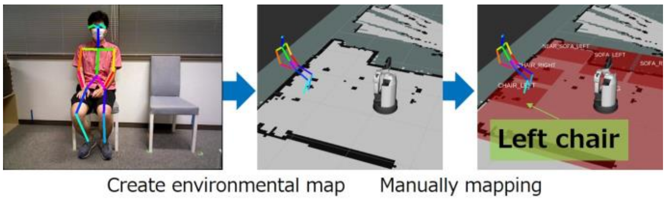

Left chair
Create environmental map
Manually mapping

Fig. 2 Proposed semantic map

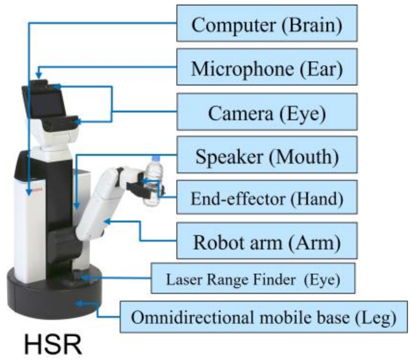

Computer (Brain)
Microphone (Ear)
Camera (Eye)
Speaker (Mouth)
End-effector (Hand)
Robot arm (Arm)
Laser Range Finder (Eye)
Omnidirectional mobile base (Leg)
HSR

Fig. 3 overview of HSR.

a semantic map, which assigns attributes to the surrounding environment, is effective. Figure 1(c) shows an example of a semantic map. In this figure, the robot can assign the attribute to an area previously judged to be an obstacle.

In this study, we propose an efficient and effortless semantic map for a more flexible HRI. We evaluate the effectiveness of our proposed semantic map in RoboCup@Home [6]. RoboCup@Home conducts multiple benchmark tests in a room that imitates an actual home environment. Therefore, we can evaluate the performance for social implementation.

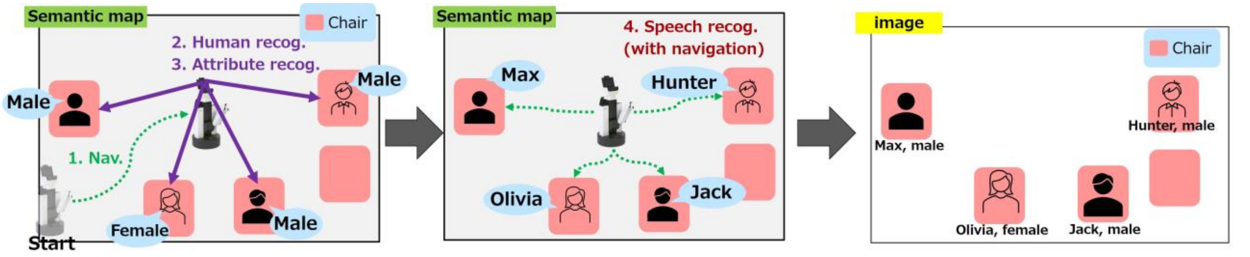
2. Proposed Method

In this study, we propose an efficient semantic map. Figure 2 shows an overview of our method. First, our semantic map is created based on an environmental map using SLAM. Next, the semantic map is updated by manually mapping the furniture location to the coordinates in the environment map. Manual mapping is effortless and fast because it requires only two diagonal coordinates for each object. For example, a semantic map for a standard room can be created in less than an hour. In addition, we use the human and action recognition method proposed by Ono et al. [7] Figure 2 shows the semantic map. In this case, the robot estimates that the waving customer is sitting in the left chair.

3. Experiments

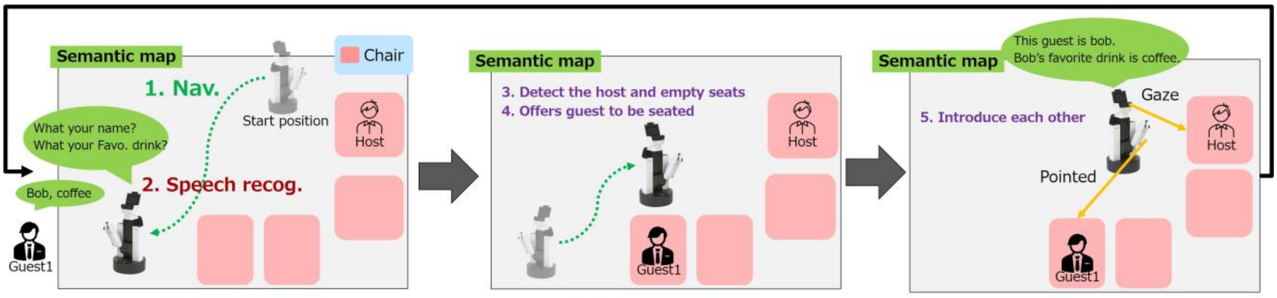
We participated in RoboCup@Home in Bangkok in July 2022 to evaluate the effectiveness of our proposed semantic map. We focus on two tasks which can evaluate flexible HRI in the domestic environment. In RoboCup@Home, we use Human Support Robot (HSR) [8] as a domestic standard robot. Figure 3 shows the HSR appearance and mainly sensors. In addition, the PC specifications to control HSR are as follows: CPU: Intel core i7-7820HK, GPU: Geforce RTX 1080, Memory: 32 GB, OS: Ubuntu18.04.

In this experiment, we use Vosk [9] as a speech recognition method and class-specific residual attention [10] as a human attribute recognition system. In addition, we use real-time appearance-based mapping [11] as a SLAM.


Semantic map
Chair
2. Human recog.
3. Attribute recog.
Male
Male
1. Nav.
Female
Male
Start
Semantic map
4. Speech recog.
(with navigation)
Max
Hunter
Olivia
Jack
image
Chair
Max, male
Hunter, male
Olivia, female
Jack, male

(a) Human and attribute recognition (b) Speech recognition with navigation (c) Visualize guests' location and characteristics

Fig. 4 Solution for Find My Mates.


Semantic map
Chair
1. Nav.
Start position
Host
What your name?
What your Favo. drink?
2. Speech recog.
Bob, coffee
Guest1
Semantic map
3. Detect the host and empty seats
4. Offers guest to be seated
Host
Guest1
Semantic map
This guest is bob.
Bob's favorite drink is coffee.
Gaze
5. Introduce each other
Host
Pointed
Guest1

(a) Invite a new guest (b) Detect and offer an empty seat (c) Introduce each other

Fig. 5 Solution for Receptionist.

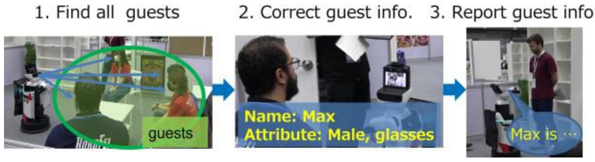

1. Find all guests
2. Correct guest info.
3. Report guest info.
guests
Name: Max
Attribute: Male, glasses
Max is …

Fig. 6 Task flow of Find My Mates.

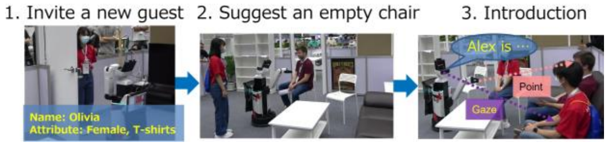

1. Invite a new guest
2. Suggest an empty chair
3. Introduction
Name: Olivia
Attribute: Female, T-shirts
Alex is …
Point
Gaze

Fig. 7 Task flow of Receptionist.

3.1. *Find My Mates*

Find My Mates (FMM) is a benchmark test that collects information about guests visiting the house and reports it to the host. Figure 6 shows the task flow of FMM. Figure 4 shows our solution using our semantic map for FMM. First, HSR navigates to the center of the room and detects where guests are sitting using our proposed semantic map. Next, HSR navigates to the front of each guest based on the location of our semantic map and asks the guest's name using speech recognition. Furthermore, HSR detects the gender of the guest using attribute recognition. Finally, HSR visualizes collected information to inform the host using the head display. Figure 6 (c) shows the visualized image by our solution for FMM. Therefore, this task can evaluate a basic HRI that considers the locations of people and furniture.

3.2. *Receptionist*

Receptionist (RP) is a benchmark test in which guests visiting a house are escorted to their rooms and introduced to their hosts. Figure 7 shows the task flow of RP. Figure 5 shows our solution using our semantic map for RP. First, HSR corrects a guest's name and favorite drink using speech recognition. Next, HSR detects which chair the host is sitting in using our proposed semantic map. In addition, HSR detects empty seats and offers guests to be seated. Finally, HSR realizes pointing to guests based on the seat locations registered in our semantic map and self-location. RP can evaluate more flexible HRI than FMM because RP needs to realize multiple approaches using furniture and human location.

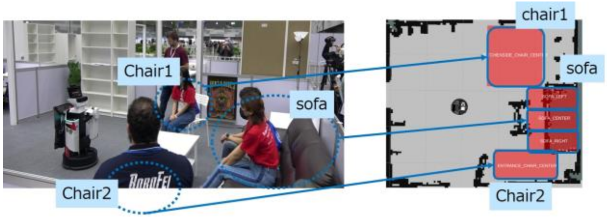
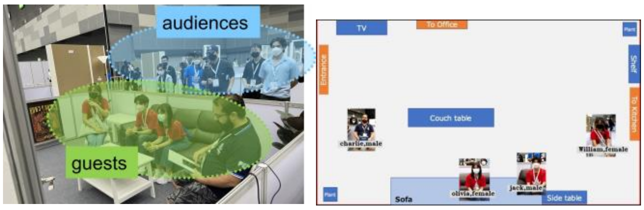


Fig. 8 Created semantic map in RoboCup@Home.

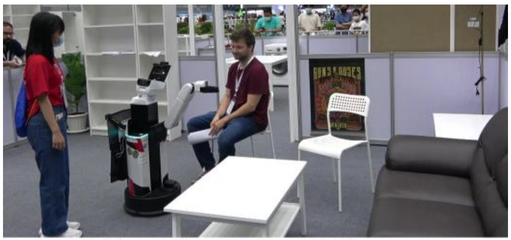


(a) Actual environment (b) visualized image

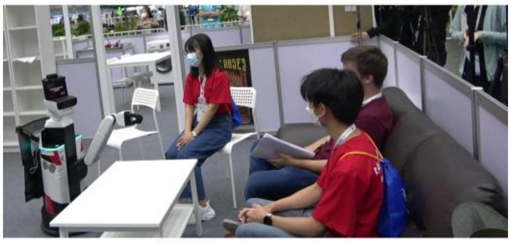
Fig. 9 Visualized image to inform guests location and characteristics.

Table 1 Result of Find My Mates.

	First try	Second try
1. Navigation	-	✓
2. Human recognition	✓	✓
3. Attribute recognition	✓	✓
4. Speech recognition	-	-
5. Report to the host	Due to Navigation failed	✓



(a) Suggest an empty seat to the guest



(b) Introduce each other

Fig. 10 Behavior of HSR in Receptionist.

Table 2 Result of Receptionist.

	First guest	Second guest
1. Navigation	✓	✓
2. Speech recognition	-	-
3. Detect empty seats	✓	✓
4. Offer a guest to be seated	✓	✓
5. Introduce a new guest	✓	✓
5-A. Look at a guest	-	-
5-B. Point at a guest	-	-
6. Describe the first guest to the second guest		✓

4. Experimental Results

4.1. *Semantic map*

Figure 8 shows created semantic map in RoboCup@Home. In the experiment, we mapped only the chairs and sofas necessary to perform the task, omitting the desks and the like. We mapped the sofas onto the environmental map in three separate sections because the sofas were designed for three people. The creation of the semantic map, together with the acquisition of the environmental map, took about 30 minutes. It took about 30 minutes to create the semantic map.

4.2. *Find My Mates*

We tried FMM twice at RoboCup@Home. Table 1 shows our result of FMM. In the first trial, HSR could detect human location correctly using our semantic map. In the first trial, HSR failed in the initial navigation, but HSR could detect the person's location using a semantic map correctly. However, due to the low resolution of the guest images, HSR could not perform attribute recognition correctly. As a result, location reports were not accepted because the person images shown on the head display were unclear. In the second trial, we address the navigation issue of the first trial. HSR could get high-resolution guest images and recognize guest attributes correctly. In addition, HSR could navigate in front of the guest but could not get the guest's name by speech recognition. Figure 9(b) shows the image that visualizes guest locations and characteristics in the second trial. In this trial, guests were seated, as shown in Fig. 9(a), and the image correctly reports where guests were seated. Furthermore, the gender and name were also correct, resulting in a perfect score.

4.3. *Receptionist*

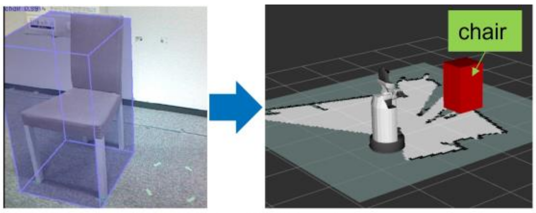
We tried RP only one time. Table 2 shows our results of RP. HSR could navigate in front of the door and ask the guests. However, HSR could not recognize their names and drinks by speech recognition. Furthermore, the chairs were placed in different locations from the predefined location beforehand. As a result, HSR misidentified one of the spectators as the host. Despite the misidentification, HSR could invite all guests and introduce each host and all guests. Figure 10 (a) shows that the HSR suggests an empty seat for the guest. HSR points toward the empty seat and looks at the guest to suggest guest be seated. Figure 10 (b) shows how HSR introduces each guest and host. In this case, the HSR points toward the host and introduces the second guest while looking at the second guest.

5. Discussion

In this study, we proposed a semantic map and conducted experiments in RoboCup@Home. Since HSR could report guest location using our proposed semantic map in FMM, our proposed semantic map can realize a basic HRI that considers the person's location and furniture information. However, HSR could not operate correctly in RP because of the chair position change. Because our proposed map is created manually and we cannot update the map in real-time. In future works, we will enhance the proposed semantic map to operate in dynamic environments using furniture detection. Figure 11 shows our new semantic map with 3D object detection. We use omni3d [12] as a furniture detection method capable of 3D detection and tracking detected furniture.

6. Conclusion

In this study, we proposed efficient semantic mapping and its application to realize flexible HRI using the domestic standard service robot. We also evaluated the effectiveness of our proposed method in RoboCup@Home. In RoboCup@Home, we achieved first place in each task. In future works, we will study semantic mapping using 3D object recognition.



(a) 3D furniture detection (b) Mapping to environmental map

Fig. 11 Semantic mapping using 3D object detection.

References

1. Y. Yano, I. Matsumoto, Y. Fukuda, T. Ono, H. Tamukoh, "Proposal for Solution of Human Interaction Task in RoboCup@Home," JSAI SIG on AI Challenge, SIG-Challenge-060-02, 2022.
2. Y. Tanaka, T. Morie, H. Tamukoh, "An amygdala-inspired classical conditioning model on FPGA for home service robots," IEEE Access, Vol. 8, pp. 212066-212078, 2020.
3. T. Matsushima, Y. Noguchi, J. Arima, T. Aoki, Y. Okita, Y. Ikeda, K. Ishimoto, S. Taniguchi, Y. Yamashita, S. Seto, S. Shane Gu, Y. Iwasawa, Y. Matsuo, "World robot challenge 2020 – partner robot: a data-driven approach for room tidying with mobile manipulator," Advanced Robotics, Vol. 36, No. 17-18, pp. 850-869, 2022.
4. KEENON website, catering robot T-8, https://www.keenonrobot.com/jp/index/Page/index/catid/32.html, (accessed 2022-12-22).
5. PUDO Robotics, Bella Bot, https://www.pudurobotics.com/products/bellabot, (accessed 2022-12-22).
6. " RoboCup@Home," https://athome.robocup.org/, (Accessed 2022-09-03)
7. T. Ono, D. Kanaoka, T. Shiba, S. Tokuno, Y. Yano, A. Mizutani, I. Matsumoto, H. Amano, H. Tamukoh, "Solution of World Robot Challenge 2020 Partner Robot Challenge (Real Space)," Advanced Robotics, Vol. 36 No. 17-18, pp. 870-889, 2022.
8. 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, Art. no. 4, 2019.
9. alpha cephei Vosk Offline speech recognition. https://alphacephei.com/vosk/, (Accessed 2022-0904).
10. Ke Zhu, Jianxin Wu, "Residual Attention: A Simple But Effective Method for Multi-Label Recoginition," arXiv, 10.48550/ARXIV.2108.02456, 2021.
11. Mathieu Labbé and François Michaud, "RTAB-Map as an open-source lidar and visual simultaneous localization and mapping library for large-scale and long-term online operation," Journal of Field Robotics, Vol. 36, No. 2, pp. 416–446, 2019.
12. G. Brazil, J. Straub, N. Ravi, J. Johnson, G. Gkioxari, "Omni3D: A Large Benchmark and Model for 3D Object Detection in the Wild," arXiv, 2022.

Authors Introduction

Mr. Yuga Yano

He received the B.Eng. degree from Kyushu Institute of Technology, Japan, in 2022. He is currently in a master's course student in the graduate school of Life Science and Systems Engineering, Kyushu Institute of Technology. His research interest includes image processing, autonomous driving, and domestic service robots.

Mr. Yukiya Fukuda

He received the B.Eng. degree from National Institute of Technology, Ube College, Japan, in 2020. He received the M.Eng. from Kyushu Institute of Technology, Japan, in 2022. 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.

Mr. Tomohiro Ono

He received the B.Eng. degree from National Institute of Technology, Ube College, Japan, in 2018. He received the M.Eng. from Kyushu Institute of Technology, Japan, in 2020. He is currently in a Ph.D. student in the graduate school of Life Science and Systems Engineering, Kyushu Institute of Technology. Since 2020, he has also been a research fellow of the Japan Society for the Promotion of Science (JSPS). His research interest includes image processing, motion planning and domestic service robots. He is a student member of RSJ.

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.

Impact of PS Load on FPGA Object Detection System Performance

Yusuke Watanabe
CRAFT WORK Co., Ltd,
5F OS Bldg., 3-5-15 Shibasaki-cho, Tachikawa, Tokyo, 190-0023, Japan
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka, 808-0196, Japan

Hakaru Tamukoh
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka, 808-0196, Japan
E-mail: watanabe.yusuke898@mail.kyutech.jp, tamukoh@brain.kyutech.jp
http://www.lsse.kyutech.ac.jp/english/

Abstract

A field-programmable gate array (FPGA) device which has a Zynq architecture becomes popular these days. It is featured by integration of processing system (PS) and programmable logic (PL) into a single chip. While we tend to focus on the performance of PL, we can not ignore PS load completely. In this paper, using a Zynq FPGA board, we explore how our object detection system performance changes with PS load and report our experiment results.

Keywords: FPGA, Zynq, PL, PS, Parallel Processing, System Performance.

1. Introduction

In robotics, demands for running neural networks such as object detection on a low energy consumption device are extremely high. We can soon find papers about relationship between neural networks and energy consumption [1], [2] rogrammable gate arrays (FPGAs) meet the requirement as they are well known for their low energy consumption [3], [4]. Thanks to the recent advancement of FPGAs, architectures such as Zynq which are featured by integration of the software programmability of a processing system (PS) and the hardware programmability of a programmable logic (PL) into a single device have become popular.

When we employ FPGA devices, though we tend to focus on the PL performance with regard to the system performance [5], [6], we can not ignore the PS performance completely. The larger a system becomes, the larger its PS load tends to be. Therefore, we research an impact of PS load on the system performance by measuring the execution time. Considering both the PS and the PL performance together, we intend to improve the whole system performance.

2. Method

We use our object detection system shown in Fig. 1 to research influence of the PS load on PL and system performances. Our object detection in PL is implemented based on Ref. 7 and Ref. 8 and performs convolution and max pooling operations which are widely used in a neural network. Our System works on an FPGA board having a Zynq architecture to which a USB camera and a monitor are directly connected. An image file and the USB camera image are used as input data to our system. We change the PS load in our system by switching these two input methods. After receiving input data, our system

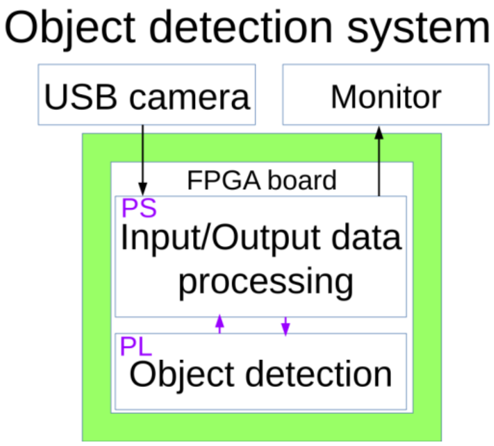

Object detection system
USB camera
Monitor
FPGA board
PS
Input/Output data processing
PL
Object detection

Fig. 1: Our object detection system on a FPGA board

then detects objects and pass an output image on which bounding boxes are drawn to the monitor. We measure execution time of both object detection in PL and inference in the application to evaluate our system performance. The inference time represents a period of time from getting input image data to displaying output image data.

We conduct several experiments, collect the execution time and compare the time to evaluate the impact of the PS load on our object detection system performances.

3. Experiment

We employ a Zynq UltraScale+ MPSoC ZCU102 evaluation board, Logicool C270 HD WEBCAM and BenQ GW2480T as our experiment environment and conduct three experiments executing the following applications.

(i) FPGA application using multiple threads.
(ii) C++ and FPGA applications.
(iii) FPGA application using only the main thread.

Our FPGA applications in all experiments are generated by AMD Xilinx tool and perform completely the same object detection in PL and the other operations in PS as shown in Fig. 1. While there is no difference in the PL operations, the PS operations vary in each application. As execution time, we measure CPU time by C library clock() function and measure wall time by C library clock_gettime() function. CPU time refers to the time CPU is busy in processing the program's instructions. Wall time refers to the elapsed time during measurement. OpenCV imshow() function is called to display an output image from our system.

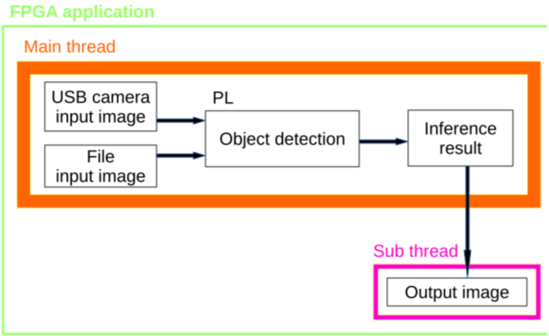

FPGA application
Main thread
USB camera input image
File input image
PL
Object detection
Inference result
Sub thread
Output image

Fig. 2: An overview of the first experiment application

Fig. 2 represents an overview of processing in the FPGA application executed in the first experiment. We switch two input methods by changing arguments passed to the application. We call pthread_create() function in the POSIX thread libraries to display an output image in a sub thread.

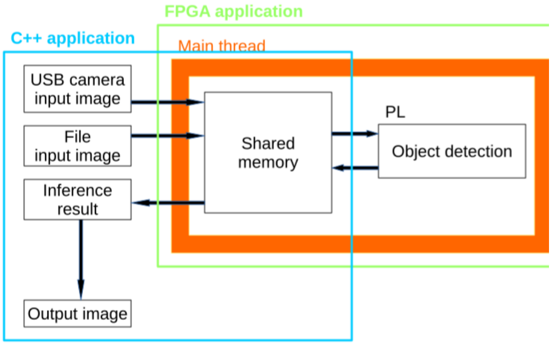

FPGA application
C++ application
Main thread
USB camera input image
File input image
Inference result
Shared memory
PL
Object detection
Output image

Fig. 3: An overview of the second experiment

Fig. 3 represents an overview of the second experiment application. We execute the C++ application which is generated by cmake and automatically executes the FPGA application. Shared memory is used to pass data between the C++ and the FPGA applications. Whatever the input methods are, input data to the FPGA application are passed through the shared memory and therefore we only employ the image file to measure the execution time. In the second experiment, we also measure the execution time of the FPGA application which is the time from right after the beginning of the FPGA application to right before the end of the FPGA application.

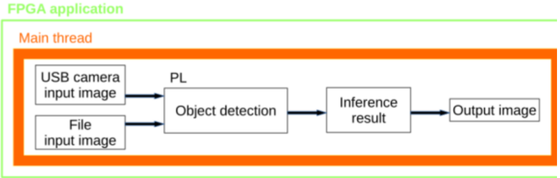


Fig. 4: An overview of the third experiment application

Fig. 4 shows an overview of the third experiment application which is almost the same as the first experiment application. The only difference is that the function to display an output image is called in the main thread instead of called in a sub thread.

We execute each application in the three experiments for ten times and record its execution time. When we use the USB camera, we regard the time when any object is detected as the execution time. We prepare an image file so that any object is detected by the system and we use it in all experiments.

4. Results

Table 1, Table 2, Table 3, Table 4 and Table 5 show the results of our experiments. Min. is the minimum execution time and Max. is the maximum time. Ave. means the mean time of the ten measurement. The time is showed in the second.

Table 1. CPU time in the first experiment

			Image input	
			Camera	File
CPU time (sec)	Inference	Min.	1.344	5.978 x10^{-1}
		Ave.	1.348	5.988 x10^{-1}
		Max.	1.350	5.997 x10^{-1}
	PL	Min.	1.042 x10^{-1}	1.016 x10^{-1}
		Ave.	1.046 x10^{-1}	1.028 x10^{-1}
		Max.	1.080 x10^{-1}	1.048 x10^{-1}

Table 2. Wall time in the first experiment

		Image input	
		Camera	File
Inference	Min.	1.215	5.136 x10^{-1}
	Ave.	1.219	5.145 x10^{-1}
	Max.	1.221	5.161 x10^{-1}
PL	Min.	5.231 x10^{-2}	5.271 x10^{-2}
	Ave.	5.232 x10^{-2}	5.273 x10^{-2}
	Max.	5.233 x10^{-2}	5.281 x10^{-2}

Table 1 is the CPU time of the first experiment. Table 2 is the wall time of the first experiment. From Table 1 and Table 2, the PS load of the USB camera is higher than the image file and the execution time in PL is almost the same in both input methods. These results indicate that there is no impact of the PS load on the PL execution.

Table 3. CPU time in the second experiment

			Image input
			File
CPU time (sec)	Inference	Min.	5.797
		Ave.	5.808
		Max.	5.817
	FPGA application	Min.	4.071
		Ave.	4.073
		Max.	4.074
	PL	Min.	5.285 x10^{-2}
		Ave.	5.294 x10^{-2}
		Max.	5.301 x10^{-2}

Table 3 is the CPU time of the second experiment. We do not show the wall time because both the CPU and wall time are almost the same. From Table 3, the CPU time in PL is shorter than the first experiment application though the inference CPU time is longer, and the wall time in PL is almost the same.

Table 4. CPU time in the third experiment

			Image input	
			Camera	File
CPU time (sec)	Inference	Min.	1.258 x10^{-1}	5.067 x10^{-1}
		Ave.	1.279 x10^{-1}	5.076 x10^{-1}
		Max.	1.285 x10^{-1}	5.082 x10^{-1}

PL	Min.	5.218 x10^{-2}	5.262 x10^{-2}
	Ave.	5.227 x10^{-2}	5.272 x10^{-2}
	Max.	5.231 x10^{-2}	5.281 x10^{-2}

Table 5. Wall time in the third experiment

			Image input
			Camera
Wall time (sec)	Inference	Min.	1.209
		Ave.	1.213
		Max.	1.236
	PL	Min.	5.231 x10^{-2}
		Ave.	5.232 x10^{-2}
		Max.	5.232 x10^{-2}

Table 4 is the CPU time of the third experiment. Table 5 is the wall time of the third experiment. We do not show the wall time of the image file input because both the CPU and wall time are almost the same. From Table 4 and Table 5, the CPU time of both inference and PL is shorter than the first experiment application and the wall time is almost the same.

From the results of all experiments, we paid attention to the relationship between the number of threads in PS and the execution time in PL and show it as Table 6. The execution time in Table 6 is the mean time of the image file input.

Table 6. Relationship between the number of threads and the execution time in PL

		The number of threads	Mean CPU time in PL (sec)	Mean wall time in PL (sec)
Experiment	One	2	1.028 x10^{-1}	5.273 x10^{-2}
	Two	1	5.294 x10^{-2}	
	Three	1	5.272 x10^{-2}	5.232 x10^{-2}

5. Conclusion

There is no impact of the PS load on PL execution as for the wall time. Although the CPU time in PL increases when we increase the number of threads in PS, the wall time in PL does not change. We can conclude the system performance highly depends on the PS load from our experiments. To improve the system performance, we can consider the PS and PL separately because the PS load does not affect the wall time in PL.

As a future work, we need to further research whether the increase of the CPU time in PL affects the wall time in PL.

References

1. C. Profentzas, M. Almgren, O. Landsiedel, "Performance of deep neural networks onlow-power IoT devices", Proceedings of the Workshop on BenchmarkingCyber-Physical Systems and Internet of Things, 2021.
2. L. Caballero, Á. Perafan, M. Rinaldy, W. Percybrooks, "Predicting the Energy Consumption of a Robot in an Exploration Task Using Optimized Neural Networks", Electronics, 10, 920, 2021.
3. M. Qasaimeh, K. Denolf, J. Lo, K. Vissers, J. Zambreno and P. H. Jones "Comparing Energy Efficiency of CPU, GPU and FPGA Implementations for Vision Kernels", *IEEE International Conference on Embedded Software and Systems (ICESS)*, pp. 1-8, 2019.
4. H. Nakahara and T. Sasao, "A High-speed Low-power Deep Neural Network on an FPGA based on the Nested RNS: Applied to an Object Detector", IEEE International Symposium on Circuits and Systems (ISCAS), pp. 1-5, 2018.
5. L. Mo, C. Wu, L. He and G. Chen, "Layout driven FPGA packing algorithm for performance optimization", IEICE Electronics Express, 2017.
6. T. Nguyen, C. MacLean, M. Siracusa, D. Doerfler, N. J. Wright and S. Williams, "FPGA-based HPC accelerators: An evaluation on performance and energy efficiency", Concurrency Computat Pract Exper, 2022.
7. M. Courbariaux, I. Hubara, D. Soudry, R. El-Yaniv and Y. Bengio, "Binarized Neural Networks: Training Deep Neural Networks with Weights and Activations Constrained to +1 or -1", 2016.
8. H. Nakahara, H. Yonekawa, H. Iwamoto and M. Motomura, "A Batch Normalization Free Binarized Convolutional Deep Neural Network on an FPGA (Abstract Only)", pp. 290-290, 2017.

Authors Introduction

Mr. Yusuke Watanabe

He received his B.Eng. and M.Eng. degrees from Waseda University, Japan, in 2007 and 2009. He is pursuing his Ph.D. degree in Kyushu Institute of Technology. He is working for CRAFT WORK Co., Ltd.

Mr. Hakaru Tamukoh

He received his B.Eng. degree from Miyazaki University, Japan, in 2001. He received his M.Eng. and Ph.D. degrees from Kyushu Institute of Technology, Japan, in 2003 and 2006, respectively. He was a postdoctoral research fellow at Kyushu Institute of Technology, from 2006 to 2007. He was an assistant professor at Tokyo University of Agriculture and Technology, from 2007 to 2013. He is currently a professor in the graduate school of Life Science and Systems Engineering, Kyushu Institute of Technology, Japan. His research interest includes digital hardware design, soft-computing and home service robots. He was the author of works that won the Best Paper Award at IJCNN 2019, the Best Live Demonstration Award at ISCAS 2019, the Best Paper Award at ICONIP 2013. He is a member of IEEE, IEICE, JNNS.

An Effective Method for Minimizing Domain Gap in Sim2Real Object Recognition Using Domain Randomization

Tomohiro Ono[*]
Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka, 808-0196, Japan [†]

Akihiro Suzuki
Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka, 808-0196, Japan

Hakaru Tamukoh
Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka, 808-0196, Japan
E-mail: ono.tomohiro342@mail.kyutech.jp, suzuki@brain.kyutech.ac.jp, tamukoh@brain.kyutech.ac.jp
http://www.lsse.kyutech.ac.jp/english/

Abstract

Manual annotation is common, but problems occur, such as oversight and mislabeling via human error. These problems are known to affect the quality of datasets significantly. To resolve these problems, we propose a method to automatically generate high-quality and large datasets in a short time using a simulator. Our proposed method uses domain randomization to minimize domain gaps without faithfully reproducing real scenes. The generated dataset achieved more than 80% recognition accuracy against the real image dataset.

Keywords: Data-centric Deep Learning, Object Recognition, Sim2Real, Dataset Generation.

1. Introduction

With the advent of deep learning, various technologies have been realized. In particular, the impact of deep learning on object recognition has been significant, and various models and large datasets have been released for validation. Deep learning performance is determined by $Code(model, algorithm) * Data$. There are two approaches to improving this performance: model-centric, which aims to improve performance based on the model, and data-centric, which seeks to improve performance based on data. Many studies have been conducted on improving the performance of these models. However, the reality is that the performance of models has reached a plateau and has yet to further improve dramatically. Therefore, recently, the data-centric approach, which improves the quality of datasets, has attracted research attention[1].

Figure 1 shows the four essential elements that define the quality of the dataset. The first is consistent labeling (Fig. 1(a)). For example, consider the annotation of the image of two elephants shown in Fig. 2 Suppose the annotator is instructed to "enclose the elephants with bounding boxes (BBs)." At this time, there are three ways to annotate the elephants: annotate them with a single BB, as in Fig. 2(a), annotate the visible part of each elephant, as in Fig. 2(b), or annotate only the non-overlapping parts of each elephant, as in Fig. 2(c). Since the optimal solution for these ways depends on the situation, it is necessary to establish clear rules in advance and annotate according to the rules. However, such rules are often very complex and vary widely. The second is the accuracy (Fig. 1(b)). Accuracy refers to the ability to annotate appropriately. Annotation is simple and may be affected by the presence or absence of a concentration or time margin, even for the same annotator. The third and fourth phases (Fig. 1(c) and (d)) contain missing labels and images that are mislabeled, respectively. As with accuracy, these are also caused by human error and can significantly degrade the quality of the dataset. In general, double-checking, or other approaches is required to ensure quality. The dataset quantity is also essential; generally, the larger the training dataset, the higher the accuracy. However, annotation costs increase proportionally to the volume of data, and the more people involved, the greater the risk of information leakage.

---

[*].

These problems are unavoidable if people are involved. In recent years, considerable research has been conducted on automatic dataset generation using 2D image synthesis[2][3] and 3D simulation[4][5][6]. These methods do not require manual annotation, and the simulation renderer can quickly obtain synthetic images and high-quality ground-truth annotations. Thus, high-quality datasets can be generated without human error. However, the distribution of the generated synthetic images differs from that of real images. This is the so-called domain gap. Domain randomization[7] and domain adaptation[8] were proposed to fill this domain gap in Sim2Real studies. In recent years, the appearance of neural radiance fields (NeRF)[9] and others have enabled the easy acquisition of high-quality 3D models. In this study, we have focused on regions other than objects for recognition, E.g., the background. We propose an effective method to minimize the domain gap using domain randomization without reproducing real scenes. Particularly, we propose a method to generate synthetic data that encompasses the distribution of real images by complexly changing various elements in the simulator (e.g., camera positions, angles, and background textures). We conclude by effectively using complex and diverse fractal images as background textures in Sim2Real. To improve the quality of the dataset, we propose a method to remove the annotations of objects that are missing in the image and have significant feature deficits. We trained the synthetic data generated by the proposed method on You Only Look Once v7 (YOLOv7)[10] and experiments for real images.

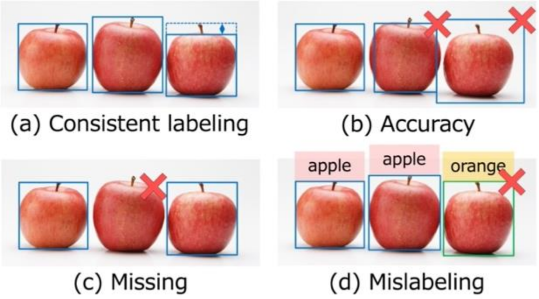


Fig. 1. Quality of dataset.

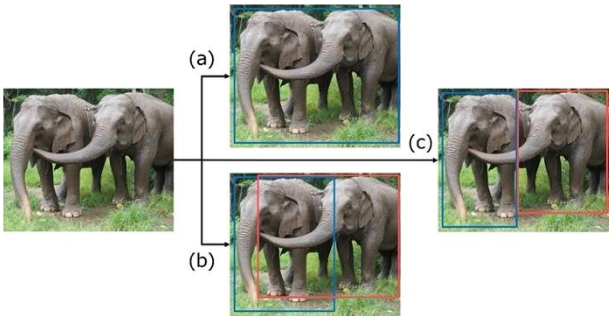


Fig. 2. Example of inconsistent labeling: when instructed to enclose the elephants in bounding boxes.

2. Proposed Method

We propose a dataset generation method using the real-time physics simulator PyBullet[11].

Figure 3 presents an overview of the proposed method. The procedure is as follows.

1. Acquire 3D models

   First, a 3D model of the object to be recognized is acquired. A high-quality 3D model can be easily obtained using NeRF, a 3D scanner, or a smartphone application. 3D modeling software can also be used to create the 3D model.

2. Initialize the simulator

   Set up the simulator environment for the assumed scene. For example, we set up simple furniture such as a desk, chair, and shelves, because we assumed that the simulator would be used in a home environment. Changing the scene according to the assumed environment enabled generation of a high-quality dataset. Each piece of furniture was

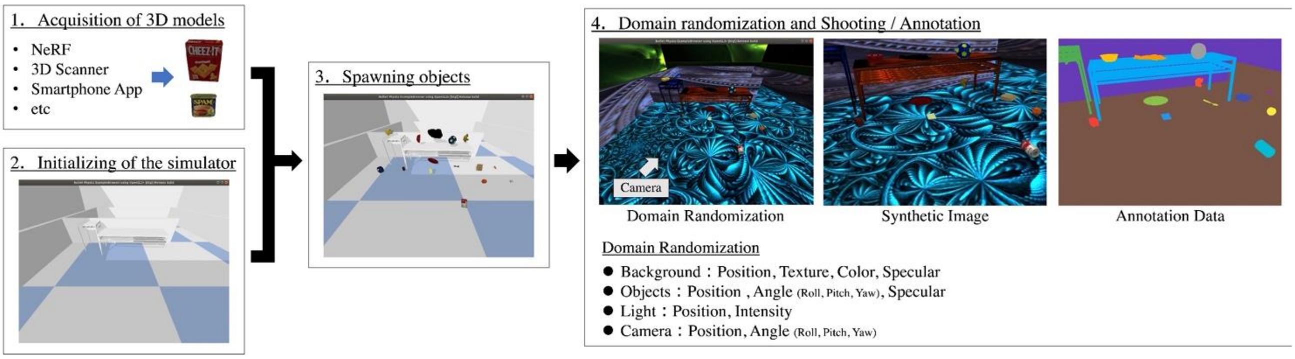

1. Acquisition of 3D models
NeRF
3D Scanner
Smartphone App
etc
2. Initializing of the simulator
3. Spawning objects
4. Domain randomization and Shooting / Annotation
Camera
Domain Randomization
Synthetic Image
Annotation Data
Domain Randomization
● Background : Position, Texture, Color, Specular
● Objects : Position , Angle (Roll, Pitch, Yaw), Specular
● Light : Position, Intensity
● Camera : Position, Angle (Roll, Pitch, Yaw)

Fig. 3. Overview of the proposed method.

randomly rearranged after a certain number of data acquisitions.

3. Spawn objects

   The 3D model created in Step 1 is generated in the simulator environment set up in Step 2. At this time, the position and orientation of the objects are randomly determined. In addition, the objects are generated in advance at a higher position than the ground and placed on the ground using physical operations to make the object placement look real. After spawning objects, the physics operations are temporarily disabled to speed up the subsequent dataset generation process.

4. Randomize the domain, shoot, and annotate

   By simulating the environment in a complex way, we improved the dataset's quality. Here, the background conditions (including furniture), objects, light source, and camera were randomly changed, and synthetic image and annotation data were generated. These data were generated by using the rendering function of the simulator. The annotation data was outputted according to the Object Detection in COCO format [12], and used for object detection, semantic segmentation, panoptic segmentation, and instance segmentation. In this case, we used monochromatic, Perlin noise and fractal images for the background textures. Our method was characterized by the fact that it did not use realistic textures. We applied this to domain randomization, and from ten to one hundred different synthetic images and annotations were generated per scene. Then, we returned to Step 2 to generate a predetermined number of synthetic images and annotations.

Furthermore, we can quickly generate data by running the simulator in parallel. Using an Intel i9-12900 K CPU (16 core, 24 threads, 3.20 GHz) and a GeForce RTX 3090 GPU, approximately 100,000 images could be processed in one hour when run ten in parallel.

3. Experiments

The synthetic dataset generated by the proposed method was trained on YOLOv7, a real-time object detection model, and verification was conducted to see if real images could be recognized.

We use 56 classes from the YCB object and model sets[13][14] published as benchmark objects as training classes. We manually annotated 50 real images as validation data. We generated three datasets, 10,000, 100,000, and 500,000 images. The human-annotated dataset was prepared for comparison with the proposed synthetic and real data in terms of dataset performance. The dataset contained 7,093 images clipped from 18 videos at 15 frames per second (fps), and the videos were filmed at 30 fps.

These datasets were used to train YOLOv7 in 32 batch sizes and 20 epochs, and COCO metrics calculate the mean average precision (mAP), which is used to compare dataset performance.

3.1. *Result*

The experimental results in Table 1 show that the mAP score improves with the number of images. Furthermore, the proposed method achieves higher accuracy than the real-world environment dataset, indicating its effectiveness.

Because of these considerations, a domain randomization technique was required to realize Sim2Real object detection. In particular, complex, and

Table. 1. Results of recognition accuracy.

Dataset	10,000	100,000	500,000	Real
mAP_{50}	0.586	0.823	**0.841**	0.667
mAP	0.471	**0.675**	**0.675**	0.528

diverse textures, such as fractal images, assisted in the Sim2Real effect without preparing a texture of a real scene.

Compared with the lead time to prepare both images and annotations, it took five annotators almost two weeks to obtain approximately 7,000 real images. In contrast, it took only one hour to get 100,000 synthetic images and annotations data using the proposed method.

This comparison and the accuracy above comparisons indicate that the proposed method excels in accuracy and reduces the time cost.

4. Conclusions

This study proposes dataset generation using a 3D simulator and domain randomization for Sim2Real. The experimental results showed that accuracy was higher when trained by a synthetic dataset than when trained by a real image dataset with human annotation. The proposed method of applying a fractal image as a texture in the simulator realized Sim2Real object detection. However, the performance in Sim2Real is still about 84%, and this is open to further study.

Future work will consider dataset generation and model updating as one step in a deep learning training procedure and will be designed to improve the deep learning in a data-centric manner.

Acknowledgements

JSPS KAKENHI (grant number:20J23242) supported this work. This study is based on results obtained from a project, JPNP16007, commissioned by the New Energy and Industrial Technology Development Organization (NEDO).

References

1. “Data-centric approach vs model-centric approach”, https://www.linkedin.com/pulse/data-centric-approach-vs-model-centric-steve-nouri/?trk=public_post
2. Y. Ishida, H. Tamukoh, “Semi-Automatic Dataset Generation for Object Detection and Recognition and its Evaluation on Domestic Service Robots,” Journal of Robotics and Mechatronics, Vol. 32, No. 1, pp. 245-253, 2020.
3. Y. Abe, Y. Ishida, T. Ono, H. Tamukoh, “Acceleration of training dataset generation by 3D scanning of objects,” The 2020 International Conference on Artificial Life and Robotics (ICAROB2020), OS20-4, Oita, Japan, January 13-16 (14), 2020.
4. M. Denninger, M. Sundermeyer, D. Winkelbauer, Y. Zidan, D. Oler, M. Elbadrawy, A. Lodhi, H. Katam, ″Blenderproc: Reducing the reality gap with photorealistic rendering″, In International Conference on Robotics: Sciene and Systems (RSS), 2020.
5. S. Max, B. Sven, ″Stillleben: Realistic scene synthesis for deep learning in robotics″, In IEEE International Conference on Robotics and Automation (ICRA), 2020.
6. G. Klaus, B. Francois, B. Lucas, D. Carl, D. Yilun, D. Daniel, F. David, J. G. Dan, G. Florian, H. Charles, K. Thomas, K. Abhijit, L. Dmitry, L. Issam, L. Hsueh-Ti, M. Henning, M. Yishu, N. Derek, O. Cengiz, P. Etienne, R. Noha, R. Daniel, S. Sara, S. M. S. Mehdi, S. Matan, S. Vincent, S. Austin, S. Deqing, V. Suhani, W. Ziyu, W. Tianhao, M. Yi Kwang, Z. Fangcheng, T. Andrea, “Kubric: a scalable dataset generator”, In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2022.
7. J. Tobin, R. Fong, A. Ray, J. Schneider, W. Zaremba, P. Abbeel, ″Domain randomization for transferring deep neural networks from simulation to the real world″, In 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pages 23-30, 2017.
8. B. Imbusch, M. Schwarz, S. Behnke, “Synthetic-to-real domain adaptation using contrastive unpaired translation”, In Proceedings of 18th IEEE International Conference on Automation Science and Engineering (CASE), 2022.
9. B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, R. Ng, “Nerf: Representing scenes as neural radiance fields for view synthesis”, In European Conference on Computer Vision (ECCV), 2020.
10. C. Wang, A. Bochkovskiy, H. M. Liao, “YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors”, arXiv preprint arXiv:2207.02696, 2022.
11. “Pybullet, a python module for physics simulation for games, robotics and machine learning”, http://pybullet.org/
12. T. Lin, M. Maire, S. J. Belongie, L. D. Bourdev, R. B. Girshick, J. Hays, P. Perona, D. Ramanan, P. Doll'ar, C. L. Zitnick, “Microsoft COCO: common objects in context”, Computing Research Repository (CoRR), abs/1405.0312, 2014.
13. B. Calli, A. Walsman, A. Singh, S. Srinivasa, P. Abbeel and A. M. Dollar, ″Benchmarking in Manipulation Research: Using the Yale-CMU-Berkeley Object and Model Set″, in IEEE Robotics & Automation Magazine, vol. 22, no. 3, pp. 36-52, Sept. 2015.
14. B. Calli, A. Singh, A. Walsman, S. Srinivasa, P. Abbeel and A. M. Dollar, ″The YCB object and Model set: Towards common benchmarks for manipulation research″, International Conference on Advanced Robotics (ICAR), Istanbul, 2015, pp. 510-517.

Authors Introduction

Mr. Tomohiro Ono

He received the B.Eng. degree from National Institute of Technology, Ube College, Japan, in 2018. He received the M.Eng. from Kyushu Institute of Technology, Japan, in 2020. He is currently in a Ph.D. student in the graduate school of Life Science and Systems Engineering, Kyushu Institute of Technology. Since 2020, he has also been a research fellow of the Japan Society for the Promotion of Science (JSPS). His research interest includes image processing, motion planning and domestic service robots. He is a student member of RSJ.

Mr. Akihiro Suzuki

He obtained B.E. degree in electrical and electronic engineering from Tokyo University of Agriculture and Technology, Tokyo, Japan, in 2014.

He has begun working under Tamukoh laboratory in Kyushu Institute of Technology, Fukuoka, Japan. He obtained a M.E. and a Ph.D degree in the Kyushu Institute of Technology, in 2016 and in 2019, respectively. He has been a postdoctoral research fellow in Tamukoh laboratory. His research interests include deep neural networks in computer vision domain.

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 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.

A Fundamental Study on Car Sickness Using Data Science

Tsutomu Ito, Matsuno Seigo
Department of Business and Administration, Ube National College of Technology,
2-14-1 Tokiwadai, Ube, Yamaguchi, 755-0096, Japan

Makoto Sakamoto, Satoshi Ikeda
Faculty of Engineering, University of Miyazaki,
1-1 Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan

Takao Ito[*]
Graduate School of advanced Science and Engineering, Hiroshima University,
1-4-1 Kagamiyama, Higashi-Hiroshima, 739-8527, Japan
E-mail: itotakao@hiroshima-u.ac.jp
** Corresponding author*

Abstract

Car sickness often occurs with dizziness and discomfort accompanied by vomiting and headaches lasting several days in severely affected person. Car sickness has been studied from various standpoints on countermeasures and onset mechanisms, but a general measure has not been established yet. In this study, all dataset of the motion that occurs in the car and the head of the person sitting and/or driving in the car are collected based on data science. By validating the data, the characteristics of person who has experience of car sickness are tested. This study proposes a new measure aiming at development of motion sickness countermeasures that do not depend on car performance and find that the quick rotational motion of the head generated in the car could be considered as one of the factors that cause car sickness.

Keywords: Data-Science, Car sickness, QoL, Statistical Analysis.

1. Introduction

Automobiles contribute to our society not only as private cars, but also as public transportation modes, such as buses and trucks that offer convenience and economical travel methods. Auto sickness is known is a disadvantage that occurs when using automobiles. Car sickness often manifests as dizziness and discomfort, which can be accompanied by vomiting and headaches lasting several days in severely affected individuals. Car sickness has been studied from various viewpoints on countermeasures and onset mechanisms, but a general coping method has not been established. This research endeavor employs a data science perspective to draw metrics on motion sickness that occurs in persons riding in the vehicle. By validating the data, the characteristics of people who experience car sickness are extracted. This study aims to provide insights for proposing motion sickness countermeasures that do not stem from vehicle performance.

2. Related research

With regard to car sickness (motion sickness), a plethora of studies are still being conducted to find countermeasures and preventive measures. According to a study by Schmidt et al. [1], car sickness is a symptom experienced by approximately 2 out of 3 person, and has the following characteristics.

- Those who are completely deaf do not get car sickness. Interestingly, patients with inner ear disorders can also develop visual motion while not physically moving.

- Blind people get car sickness, and sighted people get car sickness even with their eyes closed.
- When you read a book or look at a PC screen in a moving car, it is likely to cause car sickness.
- It is reduced when looking at an open car window, such looking at the road ahead. Car sickness is more likely when sitting backwards in a car.
- It is more likely to occur in the back seat of a car and less likely to be a driver. 46.3% in the rear seat, 36.7% in the front seat, and 17.2% in the driver.
- 62.1% of people experience motion sickness as seasickness when traveling by ships.
- 54.5% of people who are afflicted by this condition develop it even if they are looking outside the auto.
- More than 70% of passengers develop symptoms within 25 minutes after alighting the car.

For car sickness characteristics, research on the onset mechanism [2] and research on a judgment method based on eye movement [3] have been conducted. Furthermore, in recent years, research [4] aimed at countermeasures for "visual sickness" accompanying the development of VR technology has been promoted, but no clear countermeasures against motion sickness have been proposed. In considering car sickness, research [5] focused on the driving state of a vehicle has been conducted, but it is considered necessary to pay attention to the movement of the passenger that is afflicted by car sickness. Accordingly, in this study, we focus on riding posture and vehicle movement, and verify the characteristics of car sickness from the difference between passenger's motion and motion of the vehicle.

3. Objective of this research

The purpose of this study is to propose a riding posture or driving route that is less likely to cause car sickness. To achieve this goal, an experiment was performed and the moving data of a car and passenger’s head were collected as time series data of under conditions of acceleration. The method is to develop a comparative study using these data and find common characteristics to of a person who get car sickness. While several factors have been reported to cause motion sickness, such as car odors and vision, this study focuses on car sickness caused by vehicle movement.

4. Proposed method

4.1. *Experiment*

An acceleration sensor (Witmotion BET901CL-E) is attached to the hat of a passenger in order to acquire data on the movement of the head while the car is in motion. The sensor specifications are shown in Table 1.

Table 1. Specification of Accelerometer.

Manufacturer	Witmotion
Model	BET901CL-E
Communication method	Bluetooth
Baud rate	115200
Output frequency	0.1～200Hz
Acceleration range	±16g
Angular velocity range	±2000 °/s

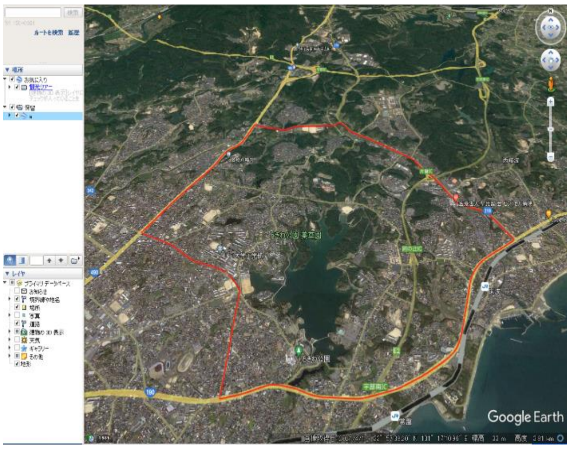


Figure 1

Similarly, the same sensor is fixed on the floor of the vehicle. In addition, a GPS data logger (Digspaice ⅲ) is attached to the vehicle to record the operation information of the vehicle. For the operation route of the experiment, we used an open road near NIT-Ube.

Table 2. List of Average Values of Acceleration and Rotational Speed Generated at the Passenger's Head Obtained.

No.	Ride Time (m)	Symptom	ax(m/s^2)	ay(m/s^2)	az(m/s^2)	a(x,y,z)(m/s^2)	wx(deg/s)	wy(deg/s)	wz(deg/s)
No.3	24.6	-	-0.095	-0.189	0.054	0.003	-4.503	-4.561	-3.096
No.6	30.0	-	-0.006	0.160	-0.042	-0.003	7.731	8.093	7.788
No.11	28.0	Mild	0.070	0.152	-0.056	-0.002	10.755	7.623	9.585
No.19	25.7	-	0.007	0.215	-0.049	-0.003	4.150	4.032	4.851
No.20	28.7	-	-0.039	0.197	-0.031	-0.002	4.167	3.960	3.720
No.21	27.1	-	0.134	0.649	-0.412	-0.006	5.492	4.205	3.854
No.22	29.2	-	0.031	0.184	-0.064	0.001	5.151	4.936	6.254
No.23	28.1	Severe	0.133	0.157	-0.056	-0.001	8.476	8.337	15.637
No.24	27.4	-	0.062	0.320	-0.110	-0.002	6.410	6.059	5.215
Mean	27.6	-	0.033	0.205	-0.085	-0.002	5.314	4.743	5.979

Figure 1 shows an example of an operating route. Since this research aims to prevent car sickness independent on the performance of the vehicle, the official car of NIT-Ube was mainly used because a general car used for the experiment is considered sufficient. The instructor drives the test vehicle, and the subject sits in the back seat. Subjects were 25 male and female students aged 17 to 23 at NIT-UBE.

4.2. *Experimental result*

Figure 2 shows the time-series data of acceleration of the vehicle and the passenger's head movement obtained by the experiment. As shown in Figure 2, it can be seen that the passenger’s head and the vehicle floor recorded different movement patterns. Only two cases, No. 11 (19-year-old female) who developed mild dizziness, and No. 23 (23-year-old male) developed car sickness during the test.

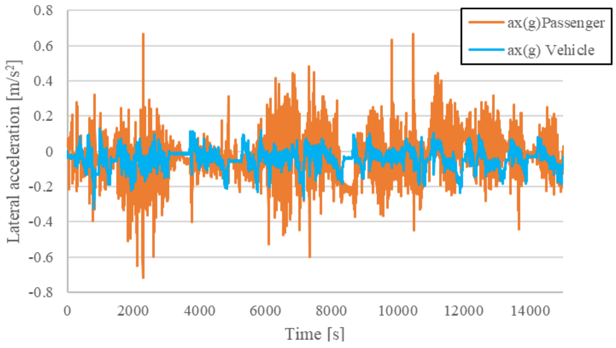


Figure 2. Time Series Results for Lateral Acceleration on Passenger and Vehicle.

5. Analysis and discussion

Table 2 shows the acceleration and rotation speed in the XYZ direction obtained from the ride experiment. Table 2 summarizes the data for ride durations of 20 to 30 minutes. It is difficult to find the tendency leading to car sickness from the movement of the vehicle and the passenger's head. According to the study by Okuyama et al. [3], considering that motion sickness-like symptoms occur even with visual information that does not involve movement, the difference of motion of the car and passenger’s movement should be considered. Here, let the motion of the vehicle be (V), and the motion of the passenger's head be (H). By taking the difference, the index (I) considering the difference between the movement predicted by the passengers from visual information and the actual movement can be calculated as follows.

$$(I) = (H) - (V) \qquad (1)$$

Table 3 shows the difference from the average value of index (I) for each direction of rotation. From Table 3, it is conceivable that in the experiments that caused motion sickness, the rotational speed of the passenger's head tended to be faster in one direction than the average. Next, the quantification of motion sickness symptoms is performed in order to perform correlation analysis. For quantification, the 5-level evaluation items established by NASA were classified into 6-level numerical values by adding 0 to completely asymptomatic. Table 4 shows the results of correlation analysis performed based on classification. From Table 4, a high correlation coefficient of 0.934 was obtained for the difference in rotational speed around the Z axis. This means that the occupant rotates his/her head in the horizontal direction quickly with respect to the yaw generated in the vehicle.

Table 3. The Difference of the Average Rotational Speed of the Passenger's Head and Symptoms.

No.	Symptom	wx(deg/s)	wy(deg/s)	wz(deg/s)
No.3	-	-0.81	-0.18	-2.88
No.6	-	2.42	3.35	1.81
No.11	Mild	5.44	2.88	3.61
No.19	-	-1.16	-0.71	-1.13
No.20	-	-1.15	-0.78	-2.26
No.21	-	0.18	-0.54	-2.12
No.22	-	-0.16	0.19	0.28
No.23	Severe	3.16	3.59	9.66
No.24	-	1.10	1.32	-0.76

Table 4. Correlation between the differences from the average value of the evaluation index related to rotation speed and symptoms.

	Symptom	wx(deg/s)	wy(deg/s)	wz(deg/s)
Symptom	1.000			
wx(deg/s)	0.633	1.000		
wy(deg/s)	0.667	0.887	1.000	
wz(deg/s)	0.934	0.744	0.838	1.000

This seems to support the analysis results of other studies that motion sickness is caused by the discrepancy between sensory motion and actual motion.

6. Conclusion

To shed light on the issue of riding posture that does not lead to motion sickness, an experiment to identify the motion during riding that is common to people who suffer from car sickness was conducted. Based on the findings, it suggested that the quick rotational motion of the head generated in the vehicle is one factor that causes car sickness. On the other hand, for the purpose of identifying the characteristics of movements that cause car sickness from the data science perspective, experimental data on subjects that suffer from car sickness requires further study. Thus more longitudinal data is required for future experiments. In addition, analyzing the motion elements that cause car sickness should be conducted by utilizing frequency estimation using FFT.

References

1. Eike A. Schmidt, Ouren X. Kuiper, Stefan Wolter, Cyriel Diels, Jelte E. Bos, An international survey on the incidence and modulating factors of carsickness, Transportation Research Part F, Traffic Psychology and Behavior, Vol. 71, 76–87, May 2020.
2. Kaname Hirayanagi, A present state and perspective of studies on motion sickness, Ergonomics, Vol. 42, No.3, 200-211, 15 June 2006.
3. Shohta Okuyama, Jun Toyotani, Nae Urata, and Yuto Omae, Automatic Recognition Model of Motion Sickness and Hierarchical Classification by Random Forest of Line of Sight, Japan Society of Directories, Journal of Japan Information Directory Society Vol. 19, No.1, 2-9, 31 march 2021.
4. Hiroyasu Ujike, Developing an evaluation system of visually induced motion sickness for safe usage of moving images, Synthesiology, English edition, Vol.5, No.3, 139-149, December 2021.
5. Kouhei Matsumoto, Norihiro Fujii and Kousuke Ohnishi, Method of motion Sickness Evaluation for Vehicle taking account of Running Condition, The Japan Society of Mechanical Engineers, Proceedings of the Japan Society of Mechanical Engineers Kansai Branch Annual General Meeting Vol. 80, 14.1-14.2, 18 March 2005.

Authors Introduction

Dr. Tsutomu Ito

Dr. Tsutomu Ito is Assistant 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

Dr. Seigo Matsuno

Dr. Seigo Matsuno is Professor of the Department of Business Administration at National Institute of Technology, Ube College, Japan. He received his Ph.D. degree in Economics from Kyushu University, Japan in 2004. His current research interests are in the areas of IT management and strategy, information systems outsourcing, and interfirm relationship management.

Dr. Makoto Sakamoto

Makoto Sakamoto received the Ph.D. degree in computer science and systems engineering from Yamaguchi University. He is presently Professor in the Faculty of Engineering, University of Miyazaki. 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, augmented reality, entertainment computing, complex systems and so on.

Dr. 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. Takao Ito

Dr. Takao Ito is Professor of Management of Technology (MOT) in Graduate School of Engineering at Hiroshima University. His current research interests include automata theory, artificial intelligence, and systems control, quantitative analysis of interfirm relationships using graph theory, and engineering approach of organizational structures using complex systems theory.

Analysis of Quoridor by reusing the results of reduced version

Takuro Iwanaga
Graduate School of Engineering, Miyazaki University, Japan
E-mail: hm17008@student.miyazaki-u.ac.jp

Makoto Sakamoto
Faculty of Engineering, University of Miyazaki,
Miyazaki-City, Miyazaki, Japan

Takao Ito
Graduate School of Engineering, Hiroshima University
Higashi-Hiroshima, Hiroshima, Japan

Satoshi Ikeda＊
Faculty of Engineering, University of Miyazaki, Miyazaki-City, Miyazaki, Japan,
E-mail: bisu@cs.miyazaki-u.ac.jp

＊Corresponding Author

Abstract

Retrograde analysis is a representative game analysis method that first lists all legal positions and searches for the best move by analyzing from the final positions to the initial position. This method is effective for analyzing the type of game in which the same position appears many times when the game is analyzed sequentially from the initial position. However, it also has the drawback of requiring a huge amount of memory to enumerate the legal solutions. In this paper, we take up the board game "Quoridor" released by Gigamic.

Keywords: Perfect play, Retrograde Analysis, Quoridor, combinatorial theory

1. Introduction

Game tree is a graph structure that represents a game as a directed graph with game positions as nodes and players' moves as edges [1]. The standard method for game analysis is to expand this game tree. Fig1 is a game tree of tic tac toe. A complete game tree will always find the best move. However, there are some problems with this method. One of the problem is the possibility of repeating the same positions. For Tic-Tac-Toe, the number of squares that can be placed on the board decreases with each move to an arbitrary position, and positions that have appeared before the current position will never appear again. However, in chess and shogi, once a move is made, it can be moved back to the previous position, so the same position can appear many times at different nodes of game tree.

Retrograde analysis was devised to solve this problem. Retrograde analysis enumerates all regal positions and propagates the win/loss information from the final position where the winner is decided to the initial position. When the win-loss information is no longer updated, the initial position is classified as either a must-win game, a must-lose game, or a tie for the first player.

Retrograde analysis avoids the loop in the game tree that occurs when the game continues to move through several positions. However, Retrograde analysis requires. that all regal positions be enumerated and used in the analysis, which creates a space-computing problem for keeping track of the game. Therefore, it is necessary to reduce the number of enumerated positions as much as possible.

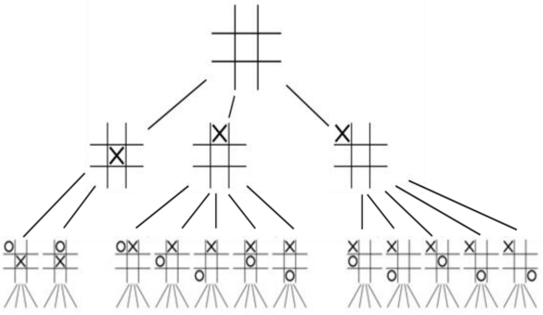


Fig 1The game tree of Tic-tac-toe. This figure shows all moves up to the second move. With a complete game tree, the best move can always be found.

The purpose of this study is to reduce the number of positions in the retrograde analysis by reusing the results of the reduced version. In this study, the phase set is partitioned as shown in Fig. 2. Here, the reduced version of the original game is the problem in which the number of items or board size has been reduced.

2. Quoridor

In this paper, we deal with a miniature board Quoridor[2]. This section describes the rules of a 5x5 board for two-player with one fence each.

2.1. *Object of the Game*

Object of the game is the same as the standard version [3], to be the first to reach the line opposite to one’s base line.

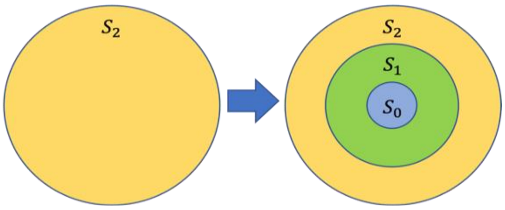


Fig 2 Before and after reuse. When the set of phases $S_k(k=0,1,...)$ is $S_0 \subset S_1 \subset S_2$... as shown in the figure, and the only phase transitions are between phases in S_k and from S_k to S_{k+1}, the results analyzed in $S_k(k=0,1,...,n-1)$ can be reused. This reduces the number of phases that need to be enumerated in S_n. Note that S_{k-1} is a reduced version of S_k in which the number of items is reduced by one step.

2.2. *Game Play (2 players)*

Each player in turn, chooses to move his pawn or to put up one of his fence. When he has run out of fences, the player must move his pawn.

At the beginning the board is empty. Choose and place your pawn in the center of the first line of your side of the board, your opponent takes another pawn and places it in the center of the first line of his side of the board (the one facing yours). Then take one fence each.

2.3. *Pawn moves*

As shown in Fig 3, the pawns are moved one square at a time, horizontally or vertically, forwards or backwards, never diagonally. The pawns must bypass the fences. If, while you move, you face your opponent's pawn you can jump over.

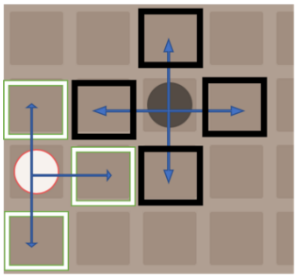


Fig 3 How to move pawn. The white square is where the white pawn can move and the black square is where the black pawn can move.

2.4. *Positioning of the fences*

The fences must be placed between 2 sets of 2 squares. By placing fences, you force your opponent to move around it and increase the number of moves they need to make. However, you are not allowed to lock up to lock up your opponents pawn, it must always be able to reach it's goal by at least one square.

2.5. *Face to face*

As shown in Fig 4, when two pawns face each other on neighboring squares which are not separated by a fence, the player whose turn it is can jump the opponent's pawn (and place himself behind him), thus advancing an extra square.

If there is a fence behind the said pawn, the player can place his pawn to the left or the right of the other pawn.

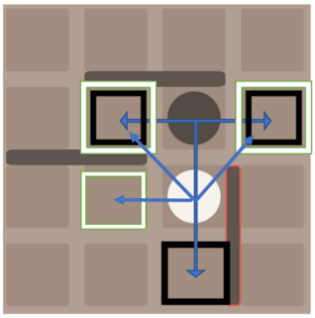


Fig 4 When two pawns are next to each other or when the path is blocked by a fence.

2.6. *End of the game*

The first player who reaches one of the 5 squares opposite his base line is the winner.

3. Retrograde Analysis

In this study, we conducted an experiment using retrograde analysis[4][5]. This method goes back one step at a time from the final stage where the victory or defeat is decided toward the initial board. In the process, if the previous move is connected to the victory phase, the victory information is received, and if all are connected to the defeat phase, the defeat information is received and the flow is repeated, so that the victory or defeat of the first phase can be known. The update of win/loss information is performed as shown in Fig 5. The advantage of this method is that you can also consider the case of a tie, which involves repeating the same move with each other, which is called “Sennichite”.

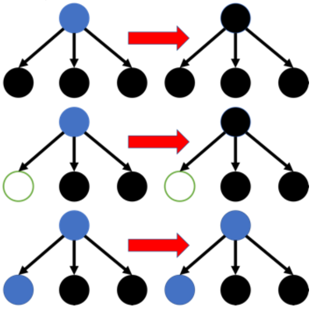


Fig 5 Blue represents an undecided game, white represents a game won by the white player, and black represents a game won by the black player.

4. Research Methods

In order to reduce the number of games to enumerate, we reuse the results of the reduced version of the game. The reuse of the results is done by using the results of the inclusion groups when performing backtracking analysis on each of the games that produce irreversible moves. In this paper, the set relation is defined as $S'_k=Sk\setminus S_{k-1}$ $(k=1,2,3,...)$ for $S_0\subset S_1\subset S_2$.... S_0 is the part that includes the initial phase.

When the value of S_k is unchanged, the entire analysis in S_k can be shortened because it does not affect the final result. When the value of S_k changes, it is updated until the phase information of the changed part is no longer propagated.

In addition, the search for all possible fronts is necessary for the regression analysis, so the regular 9x9 board size with a total of 20 fences is too large to be handled. Therefore, in this experiment, a reduced board with a smaller board size and fewer fences is used.

The size of the reduced board size is increased in each experiment, with 5x5 as the standard size. However, only odd numbers of horizontal lengths are used. This is to fix the rule that the initial placement starts from the center in front of the player. The set of games with zero fences is denoted by S_0, and the set of games with one fence is denoted by S_1. The fence is denoted by S_{10} if it was placed by the first player and by S_{01} if it was placed by the second player. The number of fences is based on each player having one fence. When two fences are placed, S_{20}, S_{11}, and S_{02} exist, but only S_{11} is used when each player has one fence available.

5. Results

The results of the experiment are shown in the Table 1.

Table 1 Number of games won or lost when the number of fences set up in the game decreased.

Number of fences installed	Number of all phases	Number of phases in which win/loss information changed
0 fence	960	764
1 fence	61440	49273
2 fence	436364	

When the game with two fences was not solved and the player moved from a game with one fence to a game with zero fences, the win/loss information for 764 games was swapped. When moving from a phase with two fences to a phase with one fence, the win/loss information for 49273 stations was swapped. When the number of fences held by both players is zero or one each, both players must win the game [6]. When there are zero fences(S_0), the late player has the advantage, but when the one fence is placed(S_1), the game changes to the first player's advantage in almost all situations, regardless of which player places the fence. Next, when the game transitions from a game with one fence(S_1) to a game with two fences(S_2), the game changes from first move advantage to second move advantage. This means that when there are an odd number of fences, the game is biased in favor of the first player, and when there are an even number of fences, the game is biased in favor of the second player. This made it almost impossible to shorten the analysis. However, by using the results of the reduced version in enumerating the phases, we were able to analyze only the added phases. In addition, we were able to reduce the number of phases to be handled at one time by separating the phases with the same number of fences that did not interfere with each other.

6. Conclusion

The way the set of phases was divided in this case, the win/loss information was almost completely changed because of the large change in advantage/disadvantage. Therefore, in addition to irreversible transitions to the next phase, it is necessary to make the division of the phase set into segments where the advantage/disadvantage does not change significantly. In the Quoridor used in this study, the results for the combined set of S_2 and S_1 can be compared with S_0 to further reduce the portion that must be reanalyzed.

References

1. Hu, Te Chiang; Shing, Man-tak, *Combinatorial Algorithms*. Courier Dover Publications. 2002.
2. https://quoridor.jp/#
3. "Quoridor Game Rules" https://www.ultraboardgames.com/quoridor/game-rules.php
4. Teturo Tanaka An Analysis of a Board Game "Doubutsu Shogi" Journal of Information Processing, Vol. 48, No. 11, pp. 3470–3476(2007).
5. J. Romein and H. Bal: Solving the Game of Awari using Parallel Retrograde Analysis, IEEE Computer, Vol. 36, No. 10, pp. 26 – 33(2003).
6. Takuro IWANAGA, Makoto SAKAMOTO, Takeo ITO, Satoshi IKEDA, Analysis of Quoridor, 2022 Information Processing Society of Japan.

Authors Introduction

Mr.Takuro Iwanaga

He received his B.S. degree in Information Systems Engineering from the University of Miyazaki in 2021. He is currently enrolled in the master's program at the University of Miyazaki.

Prof. Makoto Sakamoto

Makoto Sakamoto received the Ph.D. degree in computer science and systems engineering from Yamaguchi University. He is presently an associate professor in the Faculty of Engineering, University of Miyazaki. 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, augmented reality, entertainment computing, complex systems and so on.

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.

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.

Prototype Software for Designing Hula Accessories

Takumi Nakahara
Graduate School of Engineering, University of Miyazaki, Japan

Satoshi Ikeda, Amane Takei, Kenji Aoki, Makoto Sakamoto＊
Faculty of Engineering, University of Miyazaki, Japan

Tsutomu Ito

National Institute of Technology, Ube College, Japan

Takao Ito

Graduate School of Engineering, Hiroshima University, Japan

＊Corresponding Author

E-mail: hm18030@student.miyazaki-u.ac.jp, fruits2000jp@yahoo.co.jp

Abstract

Today, there are two main types of hula: the classical hula "kahiko" and the modern hula "auana". Music, costumes, and accessories are essential to these hula. Auana, in particular, is danced to music played on Western instruments and expresses the mythology, history, and various aspects of Hawaiian culture. The costumes and accessories worn by the dancers are designed to match the music. Therefore, we thought that it would be possible to design costumes and accessories more in line with the image of the dancers if we could freely design them using a personal computer and simulate how they would look when finished. In this study, we focused on accessories and conducted basic research to develop a CAD system for the design of "leis," one of the accessories.

Keywords: hula,Accessories,3Dmodel ,Software

1. Introduction

Today, hula (hula in Hawaiian) can be broadly divided into two types: classical hula "kahiko" and modern hula "auana," both of which have been passed down through generations as traditional Hawaiian dances. The music, costumes, and accessories are essential to these hula.

Therefore, we thought that if we could freely design costumes and accessories to match the music using a computer and simulate the finished product, we would be able to create a design that better matches the image of the hula. In this study, we focused on accessories, aiming to develop a CAD system for designing "leis," one of the accessories, and conducted a basic study of the system.[1]

1.1. *About Leis*

Leis are accessories worn around the head, neck, and shoulders, and are thought to have been introduced by Polynesians around the 12th century. Leis have been used since ancient times to ward off evil, as offerings, and as symbols of social status, and evolved greatly around the

19th century through the use of plants brought by travelers and immigrants. The materials used to make leis include nuts, leaves, flowers, shells, bird feathers, and animal bones, with many variations[Figure 1] [2].

Figure 1 Accessory Materials

1.2. *Research Background*

Leis are made of colorful materials such as flowers, leaves, nuts, shells, bird feathers, and animal bones. In order to create a lei that meets the wishes of a hula dancer with colorful flowers and colored nuts and variations of examples, it is necessary to make the image of the finished product easier to understand. Therefore, we decided to create a design program using 3DCG so that the lei designed by the user would be closer to the actual finished lei.

1.3. *Research Objectives*

This time, several functions have been added, taking advantage of our previous research. These functions are "length adjustment," "display of number of pieces," and "real-time display of designed ray. In the 3D simulation, the size of each part is randomly generated to achieve a more realistic design, and the size of each part is displayed as an ellipse to help the user have an image of attaching the part.

Development Environment

The development environment for this program is as follows

PC: Windows 10 Pro
Intel(R) Core(TM) i7-7700 CPU @ 3.60GHz
16.0 GB
Software: Unity 2020.3.18f1
Blender 2.92.0
Microsoft Visual Studio
Language: C#

2. 3DCG Design Program

This time, we created a program that allows users to select the "parts to be used" and "length" of the ray they wish to create (in the case of kukui, they select the color and the number of types to be used), display the results of the design in 3DCG, and view the designed ray in real time from various directions by changing the viewpoint by moving the mouse cursor on the 3D screen while clicking.

2.1. *Ray Length*

There are three types of leis: lei ai (necklace), lei po'o (headwear), and kuupe'e (wristwear). In past studies, lei eye was set to be 80 cm long for kukui, 100 cm for flowers and foliage, and 76 cm for shells; lei po'o was set to be 80 cm long for kukui, 63 cm for flowers and foliage, and 20 cm for all kuupe'e lengths.

In this study, however, the lengths were extended to allow users to design more freely by keyboard input.

2.2. *Selecting a Part Type*

Six types of lei parts were prepared: "Kukui (white, red, black)," "Denfale (red)," "Maile," and "Shell.

The respondents choose which parts to select from the four types: "Kukui," "Denfale," "Maile," and "Shell."

If kukui is selected, you can choose from 1 to 3 how many types of kukui to use[Figure 2].

2.3. *Design Display*

By splitting the part selection screen and the display screen 1:1, the results can be viewed simultaneously with the input[Figure 3]. The number of parts used is also displayed.

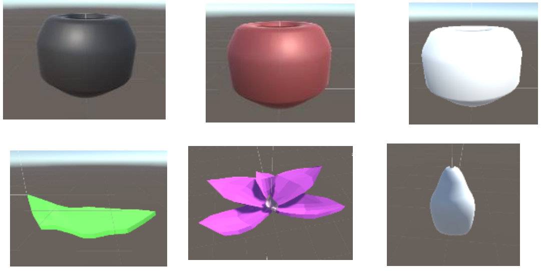


Figure 2 Types of Parts

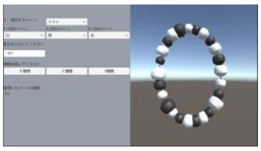


Figure 3 Design Indication

2.4. *Randomness of Parts*

The actual kukui, leaves, and flowers are unique in size. Since kukui are between 4 cm and 7 cm in diameter, they were randomly generated at 4 cm~7 cm when designed.

2.5. *Lay Parts in an Ellipse*

A: Ratio of ellipses = b/a
a (short diameter):b (long diameter) = 2:3
L: Length of string entered by user
V: Size of the part = 4
D:Length of short diameter
It:Number of parts
R:Number of cycles = 1
Rat:Location of cycles
C:Period = 2\pi

1. Find the length D of the short radius (1)

$$\mathrm{D} = \sqrt{\frac{L^2}{4} \times \frac{1}{4 - 8A + 4A^2 + \pi^2 A}} \tag{1}$$

2. E is determined by the following equation (2)

$$\mathrm{E} = \frac{\mathrm{D}}{2 * \mathrm{V}} \tag{2}$$

3. Find the number of parts It (3) The shell is small and has twice as many parts.

$$It = \frac{L}{V} \tag{3}$$

4. Set i=0 and repeat the following from i to It
5. Generate randomly from 4 cm to 7 cm
6. Find the position Rat of the cycle (4)

$$Rat = \frac{i}{It} \tag{4}$$

7. Determine the position (X,Y) to place the part (5)(6)

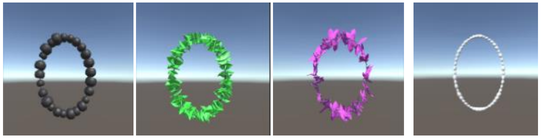
$$\mathrm{X} = a \times Cos\left(Rat \times C \times \frac{R}{2\pi}\right) \times E \tag{5}$$

$$Y = b \times Sin\left(Rat \times C \times \frac{R}{2\pi}\right) \times E \tag{6}$$

8. Rotate the part along the center of the circle.

3. Execution Result

Figures 7 show some of the results of this program[Figure 4]. The size, orientation, etc. can be adjusted in each case.



.

Figure 4 Execution results

4. Consideration

This study was extended to a 3DCG design program using Unity to create an interface design that can be operated more easily and to improve scalability by supporting multiple multi-platforms. In addition, the following functions were added: "adjust length," "display number of pieces," "display designed ray in real time," and "randomize material size.

As for the type of parts, there are still many materials used as material for the ray, so more part types and colors need to be enhanced. When displaying the design, the parts are arranged in an ellipse to make it easier to visualize the lei. However, because the size of the parts changed randomly, there were sometimes gaps between the parts. In order to prevent such gaps, it was necessary to be able to simulate the appropriate number of parts and their positions in response to the random size changes.

Through the creation of the 3DCG design program, we were able to identify various issues for the practical application of a 3DCG design program for hula necklaces (leis) in the future.

5. Conclusion

In this study, a prototype CAD system for ray design was developed through the creation of parts using Blender and 3DCG simulation using Unity. We were able to find various issues for practical use. In the future, we would like to continue to study various aspects, such as having a 3DCG fladancer wear the lei to visualize how it will look when finished.

References

1. Aloha Program [Online].
   https://www.aloha-program.com/
2. Hawaii Tourism Authority [Online].
   https://www.gohawaii.jp/

Authors Introduction

Mr. Takumi Nakahara

He is a master student at Department of Computer Science and System Engineering, University of Miyazaki. His current research interests include computer graphics and VR technology.

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 The Institute of Electronics, Information and Communication Engineers (IEICE), a delegate of the Kyushu branch of Institute of Electrical Engineers of Japan (IEEJ), a director of Japan Society for Simulation Technology (JSST).

Prof. Kenji Aoki

He received Ph.D. of Engineering from Kagoshima University in 2010. He is currently working in Information Technology Center at University of Miyazaki as Associate Professor, since 2010. His research interests include bio-informatics, evolutionary computation, information system and Intelligent systems. He is a member of IPSJ and JSET.

Prof. Makoto Sakamoto

Makoto Sakamoto received the Ph.D. degree in computer science and systems engineering from Yamaguchi University. He is presently an associate professor in the Faculty of Engineering, University of Miyazaki. 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, augmented reality, entertainment computing, complex systems and so on.

Prof. Tsutomu Ito

He is an assistant professor in department of business Administration, Ube National college of technology. He has published many papers in refereed journals and proceedings, particularly in the area of industrial management, and computer science. His current research interests include internet of things (IoT), mechanical engineering, artificial intelligence (AI), automata theory, quantitative analysis of Japanese Keiretsu. He was one of the winners of the Best Paper Award in the International Conference on Artificial Life and Robotics (ICAROB) in 2015and 2016. Dr. Ito earned his Doctor degree of Engineering from Hiroshima University, Japan in 2018.

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.

Basic Study on Museum Exhibition Support Using AR Technology

Kakeru Takemura
Graduate School of Engineering, University of Miyazaki, Japan

Satoshi Ikeda, Amane Takei, Masahiro Yokomichi, Makoto Sakamoto*
Faculty of Engineering, University of Miyazaki, Japan

Shuichi Kurogi
Miyazaki Prefectural Museum of Nature and History, Japan

Tsutomu Ito
National Institute of Technology, Ube College, Japan

Takao Ito
Graduate School of Engineering, Hiroshima University, Japan

E-mail: hm18022@student.miyazaki-u.ac.jp, fruits2000jp@yahoo.co.jp

**Corresponding Author*

Abstract

There are many bioluminescent fungi in Miyazaki Prefecture. All of these bioluminescent fungi glow in the dark and are attracting attention as a new tourist resource in Miyazaki. However, they are small and not easily observed in museums. The purpose of this study was to facilitate observation of such small mushrooms exhibited in museums by using AR technology to display a 3DCG model of the mushroom body in AR to aid observation.

Keywords: AR, Exhibition Support, Tourism Support.

1. Introduction

Augmented reality is a technology that superimposes digital information on the real world[1]. 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.

2. Research Background

A mushroom exhibition was held at the Miyazaki Prefectural Museum in 2021. Many glowing mushrooms were also on display and attracted attention. Dried specimens of the glowing mushrooms were about 5mm in diameter and difficult to observe. *Cruentomycena orientalis* was displayed with a magnifying glass. Therefore, this study examined the use of AR technology to assist in the observation of these small mushrooms to make them easier to observe.

3. AR Application Development

3.1. Development Environment

An AR application was created for use with smartphones. The development environment is shown Table 1. Unity was used to develop applications for smartphones[2]. Blender is a comprehensive 3DCG software[3]. It was used to create a 3DCG model of a mushrooms[4].

Table 1. Development Environment

OS	Windows 10 Pro
software	Unity 2020.3.22f1 Blender 2.93 Visual Studio 2019
Smart phone	Galaxy S8

3.2. Target mushrooms

The *Favolaschia peziziformis*[Figure 1] and *Cruentomycena orientalis*[Figure 2], which were particularly difficult to observe among the bioluminescent fugi on display at the museum, were the subjects of this study.

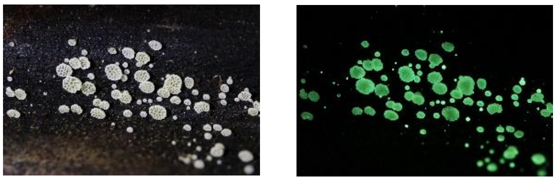


Figure 1. *Favolaschia peziziformis*

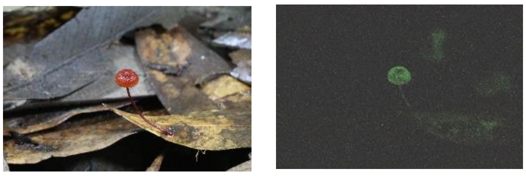


Figure 2. *Cruentomycena orientalis*

3.3. 3DCG of mushrooms

The CG model was created based on the images provided[Figure 3] [Figure 4].

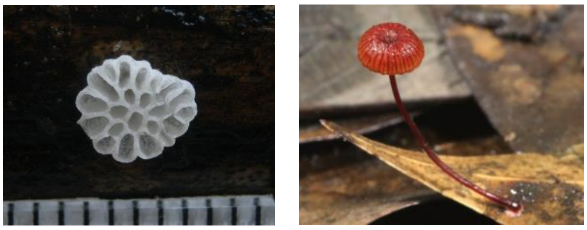


Figure 3. Courtesy image

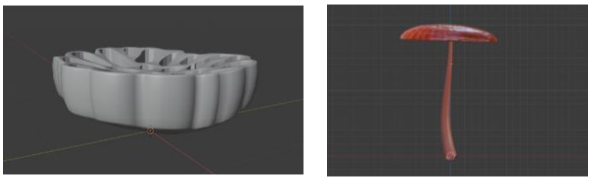


Figure 4. 3DCG models

3.4. AR application

Load an image of a mushroom to be used as a marker, and display a 3DCG model of the mushroom on top of the image in AR. When a marker disappears from the screen, the displayed 3DCG model also disappears, and when the same or another marker is scanned, the 3DCG model appears again. The use of the application is shown below[Figure 5].

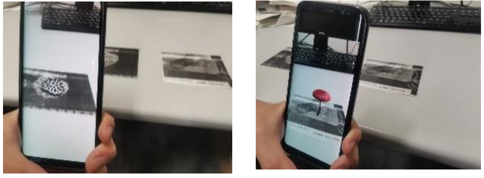


Figure 5. Execution Result

4. evaluation experiment

4.1. experimental procedure

The experiment was conducted with the cooperation of 15 students from the University of Miyazaki. The procedure is as follows.

(i) Explain the use of photographs instead of actual specimens and the placement of markers for AR.
(ii) Observe the photo for 1 minute.
(iii) Observe 3DCG models of mushrooms displayed in AR with the developed application.
(iv) After the observation is completed, a questionnaire is administered.

4.2. Questionnaire Results

The questionnaire consisted of 5-point rating questions, yes-or-no questions, and open-ended questions. The closer to 5, the better the result. Fifteen participants in the evaluation experiment answered the questionnaire. The results of the questionnaire are shown Table 2.

Table 2. Questionnaire results

item	1	2	3	4	5
Ease of use of the application	0	1	1	8	5
Ease of observing *Favolaschia peziziformis*	1	1	3	5	5
Ease of observation of *Cruentomycena orientalis*	0	2	2	6	5
Legibility of text above 3DCG mushrooms	0	2	11	1	1
	Yes			No	
discover anything new by using the app	15			0	
use the application in a museum	15			0	

Descriptive questions are as follows:

- Matters to be improved
    - When observing from above, the letters overlap the mushrooms and get in the way.
    - Difficult to observe mushrooms from below.
- What features do you want?
    - Rotation, enlargement, and reduction functions.
    - observe how it glows.
- Other Notices
    - *Cruentomycena orientalis*'s color differs between the exhibit and when observed using the app.
    - Disturbing others in the museum because I have to move when making observations from various angles.

From the survey results:

We consider it an easy-to-use application. Some users commented that loading markers was not smooth. Since the subjects were all in their 20s, smartphone operation and observation went smoothly. There was a comment that the *Favolaschia peziziformis* were difficult to observe. It was found that this mushroom has a characteristic that makes it more difficult to observe from the side than other mushrooms because it is flat. The most common request was to improve the legibility of the displayed text. Many users requested the ability to rotate and magnify the mushrooms displayed in AR. We believe this is an important feature for observing mushrooms in detail.

5. Conclusion

This study aimed to use AR technology to assist in the observation of mushrooms in museums.

The results of the evaluation experiment confirmed the usefulness of the developed application, as all 15 respondents indicated that they had made new discoveries by using the application and that they would use the application if it were introduced in a museum. On the other hand, we found various points to be improved, such as the lack of functions to rotate, zoom in, and zoom out the 3DCG model, which made it difficult to observe. This point needs to be improved because it is directly related to the most important aspect of this study, which is the observation of mushrooms. However, there are many opinions that the labeling of the mushrooms' names is an obstacle to mushroom observation, and since mushroom names are actually displayed and explained in museums, it is necessary to consider how to label the mushrooms. As a future issue, the application will be improved based on the opinions obtained from this evaluation experiment. There are some inadequacies in the evaluation experiment, such as the fact that real mushrooms could not be used, that the age range of the participants was biased, and that the experiment could not be conducted in a museum.

The above results suggest that further progress may be possible in this research. Therefore, we will continue to conduct research and experiments for further progress of this study.

Acknowledgements

We would like to express our sincere gratitude to the professors who guided us in this research and to those who cooperated in the questionnaire survey. Finally, we would like to express our deepest condolences to our collaborator, Dr. Etsuko Harada (Faculty of Agriculture, University of Miyazaki), who passed away in October of the year before last.

References

1. T. Hashimoto, "AR Programming - A Recipe for Augmented Reality Created with Processing", Ohmsha Corporation, 2012.

2. N. Tada, "Introduction to AR App Development with Unity AR Foundation", R&D Corporation, 2021.
3. S. Itami, "Introduction to Blender 2.9 - 3D production starting from scratch", Shuwa System Co, 2020.
4. Benjamin, "Blender 2.9 3DCG Modeling Master", Sewtec Corporation, 2021.

Authors Introduction

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.

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.

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.

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 The Institute of Electronics, Information and Communication Engineers (IEICE), a delegate of the Kyushu branch of Institute of Electrical Engineers of Japan (IEEJ), a director of Japan Society for Simulation Technology (JSST).

Prof. Masahiro Yokomichi

He is an Associate Professor of Faculty of Engineering at University of Miyazaki, Japan. He received his D. Eng. degree in Precision Enngineering from Hokkaido University in 1995. His research interest is Robotics and Computer Vision.

Prof. Makoto Sakamoto

He received the Ph.D. degree in computer science and systems engineering from Yamaguchi University. He is presently an associate professor in the Faculty of Engineering, University of Miyazaki. 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, augmented reality, entertainment computing, complex systems and so on.

Prof. Tsutomu Ito

He is an assistant professor in department of business Administration, Ube National college of technology. He has published many papers in refereed journals and proceedings, particularly in the area of industrial management, and computer science. His current research interests include internet of things (IoT), mechanical engineering, artificial intelligence (AI), automata theory, quantitative analysis of Japanese Keiretsu. He was one of the winners of the Best Paper Award in the International Conference on Artificial Life and Robotics (ICAROB) in 2015and 2016. Dr. Ito earned his Doctor degree of Engineering from Hiroshima University, Japan in 2018.

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.

Tourism Support for *Bioluminescent Fungi* Using Video Technology

Bidesh Biswas Biki, Kodai Hasebe, Fumito Hamakawa
Graduate School of Engineering, University of Miyazaki, Japan

Satoshi Ikeda, Amane Takei, Makoto Sakamoto*
Faculty of Engineering, University of Miyazaki, Japan

Shuichi Kurogi
Miyazaki Prefectural Museum of Nature and History, Japan

**Corresponding Author*

E-mail: ti22069@student.miyazaki-u.ac.jp, fruits2000jp@yahoo.co.jp

Abstract

Amongst 70 types of bioluminescent mushrooms, more than a dozen is found in Japan. Only in Miyazaki prefecture, there are 11 types of glowing mushrooms. But most of people are not aware of this scenic beauty. So the primary goal of this research is to make a promotional video to attract more people to come to Miyazaki and see the beauty of the bioluminescent mushrooms. After the PV, the research will be extended to VR, using Unity Engine, the project will be capable to give a real time experience for the users.

Keywords: Bioluminescent, glowing mushrooms, Virtual Reality, Unity Engine, Promotional Video (PV), Tourism.

1. Introduction

About from 70 to 100 species of glowing mushrooms have been observed around the world, and more than a dozen of them have been observed in Japan. Miyazaki Prefecture is one of the few places in Japan where 11 different types of glowing mushrooms can be seen, making it one of the few places in Japan that is blessed with such an environment. The most notable of these mushrooms is *Favolaschia peziziformis*. Although *Favolaschia peziziformis* is a very small mushroom, measuring only 6 mm in diameter, it grows on dead billow leaves and fallen leaves, and plays a part in maintaining the ecosystem by breaking down plant fibers and returning them to the soil. If you go to Aoshima Island, a tourist spot in Miyazaki, Japan, at night, you can see their bluish-white glow in the *Livistona chinensis* forest.

This study is part of a collaborative research project with the Pilz Lab of Mushroom Science, Faculty of Agriculture, University of Miyazaki and the Miyazaki Prefectural Museum, Pilz Lab is the only one in Kyushu., to create three-dimensional computer graphics (3DCG) models of 11 species of glowing mushrooms found in Miyazaki Prefecture, and The project also produced an educational promotional video (PV) on *Favolaschia peziziformis* for the general public, which was used in VR to enhance the entertainment aspect of the project. A questionnaire survey was conducted to confirm the usefulness of the educational PV.

2. Research Background

In general, when tourism [1] develops, economic activities in various fields such as lodging, transportation, food and beverage, and travel services become more active, and the economic ripple effects are greater. While there is no prospect of the new coronavirus being contained, the travel industry is becoming active in new attempts such as online tours and virtual tours. In many cases, PR activities, etc., are being promoted by skillfully using easy-to-understand 3DCG and other visual technologies.

VR is an abbreviation of Virtual Reality and is called "virtual reality" in Japan. It is a technology that creates a virtual environment on a computer and creates the illusion of actual being in that space by projecting a 360° image that covers the view of the user wearing special goggles.

Unity is a game engine developed by Unity Technologies in 2004. It is the most popular game engine in the world and is used by one million developers worldwide. It has a built-in "Integrated Development Environment" and supports multiple platforms. It adds physics to objects, which can be edited in real time as the game is played. C#, JavaScript, and Boo can be used as programming languages for writing scripts [2].

3. Development Environment

The development environment is shown in Table 1 with some use of Unity and the corresponding programming language C#.

Table 1. Development environment

Calculator	Precision Tower 3620
OS	Windows10 pro
Software	Blender
	Unity 2018.1.0f2
	SAI 2
	Power Directer1.8

4. Details of implementation

4.1. *Mycena manipularis*

The umbels are 1-3 cm in diameter and the stems are 1.5-4 cm long. It grows in clusters in summer on dead trees of various broad-leaved species and is almost white in color. The back of the umbrella is characterized by a tube hole. The glowing power is weak, and the stem emits a greenish light (Figures 1 and 2)

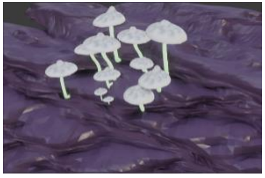


Figure 1: *Mycena manipularis* (day).

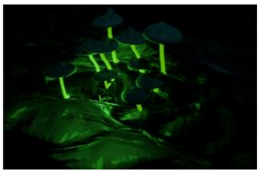


Figure 2: *Mycena manipularis* (night)

4.2. Ayahikaritake *(Roridomyces* sp.*)*

The umbrella is about 1.5 cm in diameter and 4~5 cm tall and white in color. Only the spores are originally glowing, and the whole plant appears to be glowing because the spores are attached to the mucus secreted from the stem and issued by the mucus. In addition, the spores emit light where there is moisture, but not in dry places (Figures 3 and 4).

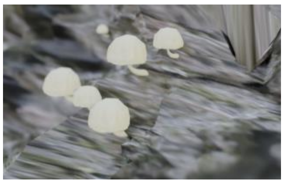


Figure 3: *Ayahikaritake* (day).

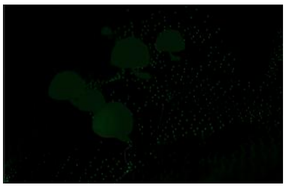


Figure 4: *Ayahikaritake* (night).

4.3. *Cruentomycena orientalis*

The umbel is deep red to blood-red in color, usually umbilicated in the center, and viscid or dry when wet. Folds sparse, drooping, fringed; scapes viscid or dry when wet (Figures. 5 and 6).

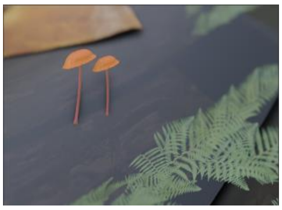


Figure 5: *Cruentomycena orientalis (day).*

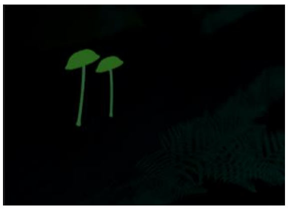


Figure 6: *Cruentomycena orientalis (night).*

4.4. *Resinomycena fulgens*

Mushrooms with umbrellas 2-5 mm in diameter that occur on the rotting bark of Daggi trees during the rainy season and autumn rains. The folds are slightly drooping. The edges of the folds of the umbrella are luminous. The Japanese name "Ginga take" was given to this mushroom because it looks like a glowing "Milky Way" when you look up at a group of *Resinomycena fulgens* (Figures 7 and 8).

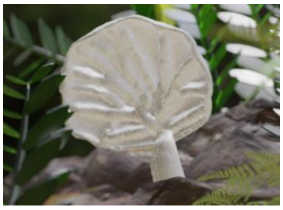


Figure 7: *Resinomycena fulgens* (day).

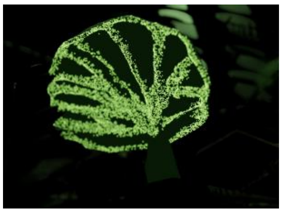


Figure 8: *Resinomycena fulgens* (night).

4.5. *Mycena lux-coeli*

The umbel is bell-shaped or conical, rarely flattened, 15-25 mm in diameter, purple brown in color, not viscous, and powdery. The folds are perpendicular to the petiole, slightly separated from each other, white with a purplish-brown margin.

The base of the petiole is not old and is often white and coarsely hairy with a hollow, cartilaginous interior. Folds and scapes strongly glabrous in the dark (Figures 9 and 10).

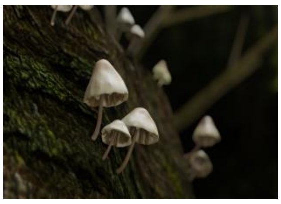


Figure 9:
***Mycena lux-coeli* (day).**

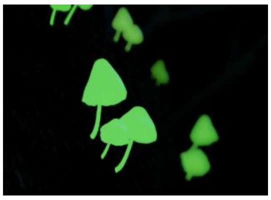


Figure 10:
***Mycena lux-coeli* (night).**

4.6. *Panellus pusillus*

Grows in clusters on dead trees. The umbels are cream-colored and 2-7 mm in diameter. The inside of the umbrella has numerous small tube pores. Light intensity is low, making it difficult to detect (Figures 11 and 12).

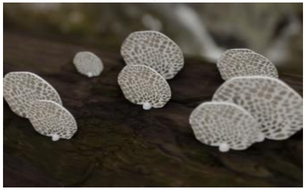


Figure 11:
***Panellus pusillus* (day).**

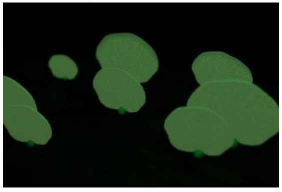


Figure 12:
***Panellus pusillus* (night).**

4.7. *Omphalotus guepiniformis*

The umbrella is semi-circular in shape, 5 to 30 cm in diameter, and its surface is sticky on damp days, such as after rain. The surface is sticky on damp days, such as after rain. When small, it is orange-brown to yellowish brown, and as it ages, it turns purple-brown or yellowish brown and becomes dull and shiny. The umbrella and the stem do not glow, neither on the surface nor inside. The folds that have been damaged by contact with a hard object will not glow. The peak of luminescence is about 2 to 3 days after the umbrella is fully opened (Figures 13 and 14).

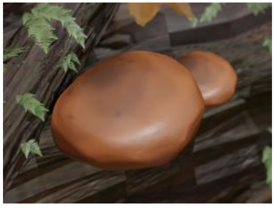


Figure 13: *Omphalotus guepiniformis* (day).

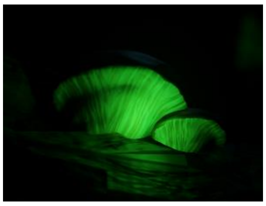


Figure 14: *Omphalotus guepiniformis* (night).

4.8. *Xeromphalina campanella*

Bristle-like mycelia are present at the base (Figures 15 and 16).

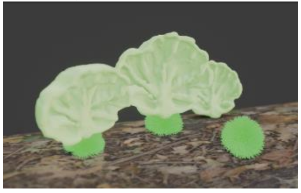


Figure 15:
***Xeromphalina campanella* (day).**

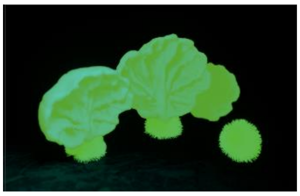


Figure 16:
***Xeromphalina campanella* (night).**

4.9. *Favolaschia peziziformis*

Favolaschia pezizformis can be observed on Aoshima Island, a tourist attraction in Miyazaki Prefecture. Aoshima is home to approximately 4,300 *Livistona chinensis*, and it is known that *Favolaschia peziziformis* occurs year-round, mainly on dead *Livistona chinensis* leaves. The diameter of the bulbs is about 6 mm, and there are tubular pores on the surface where spores are produced. It is also known to be a food source for insects and small animals such as slugs. We modeled *Favolaschia peziziformis* as the main axis of an educational PV (Figures 17 and 18).

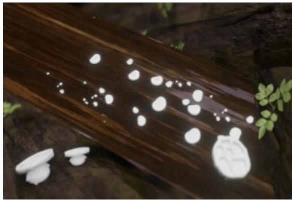


Figure 17:
***Favolaschia peziziformis* (day).**

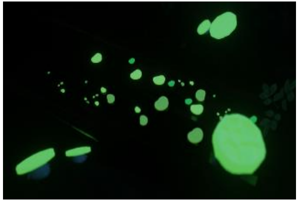


Figure 18:
***Favolaschia peziziformis* (night).**

4.10. Lightning Bug

Fireflies in Japan were also modeled, as glowing mushrooms (*Mycena lux-coeli*) and Lightning Bug are often seen together. Currently, there are three species of fireflies in Japan: genji botaru, heike botaru, and hime botaru. Figure 19 shows the actual modeling. When using the radiant node in the shader editor, the problem arose that its mesh was brightened but the light had little effect on its surroundings, so the light solves this problem by setting the light source directly near the mesh, which is illuminated by the radiant node. An image of a Lightning Bug shining is shown in Figure 20.

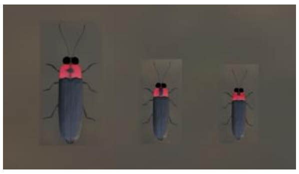


Figure 19:
Fireflies in Japan

Figure 20:
Fireflies glowing

4.11. Yurukyara

Yurukyara, short for "loose mascot character," is a mascot character that promotes regional awareness and tourism through events and campaigns. The benefits of using Yurukyara in PV include: reducing the rigid impression of educational PV, making it easier for viewers to become familiar with Favolaschia peziziformis itself, and providing appropriate information by conveying a message through Yurukyara. In addition, conveying a message through Yurukyara can lead to appropriate information dissemination, and furthermore, if it attracts media attention, it can be expected to have a significant PR effect. Figures 21 and 22 show the model we created. The umbrella part of the head emits light in the dark.

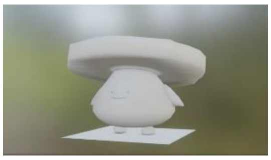


Figure 21:
Yurukyara (day)

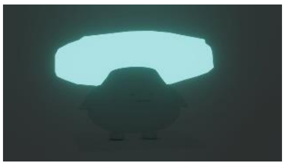


Figure 22:
Yurukyara (night)

5. Extension to VR

The created PV can be run after being built on an Android device, and when the device is set on Google Cardboard, a VR experience can be had (Figure 23). Tilting the headset horizontally or vertically is reflected in the virtual world.

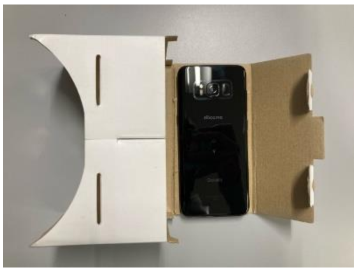


Figure 23: Terminal Set

The purpose of this project was to create an educational PV about *Favolaschia peziziformis* for public awareness. PowerDirector 1.8 was used for the production. The structure of the video is as follows. First, the video introduces Qingdao Island, mentioning the sea and the *Livistona chinensis*. Next, *Favolaschia peziziformis* during the daytime are shown and their characteristics are described. Then, the glowing *Favolaschia peziziformis* is described. Next, a slug eating an *Favolaschia peziziformis* is shown, and the presentation ends with a panoramic view of Qingdao. The following screen shows the extension to VR (Figure 24)..

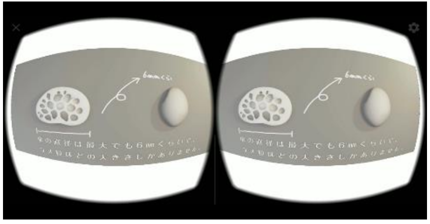


Figure 24:Favolaschia pezizaeformis and rice grains

6. Questionnaire Results

There were five questions, and all questionnaires about the PVs were rated on a scale of 1 to 5 (1: poor to 5: good). The results are shown in Figure 25.

Question 1: Were you interested in the glowing mushrooms?
Question 2: Would you like to visit the tourist attractions?
Question 3: Did you enjoy the VR experience?
Question 4: How would you rate the PV?
Question 5: Did you feel screen sickness?

7. Consideration

The average score for the question "1. Were you interested in the glowing mushrooms?" was 4 points. The average score was 4 points for the question "Were you interested in the glowing mushrooms? and "Evaluation Item 4: Evaluation of PV," the average score was 4.5.

In addition, the average response to the question "Would you like to visit the sightseeing spot in the evaluation item 2?" the average score was 3.8, which is lower than the other items, but about 70% of the respondents said they would like to go sightseeing.

In addition, the average score for the question, "Did you feel screen-sick?" was 3.8, which is lower than the average score for the other items. the average score was 3.3 points. In mobile VR, especially when using it for the first time, the focus may not match the movement of the screen, which may cause screen sickness. Therefore, we felt that it is necessary to devise ways to prevent screen sickness and to design the device in such a way that it does not cause eye fatigue.

Future tasks are to make the 3DCG model more realistic while receiving feedback from the Faculty of Agriculture and museum staff, and to improve the quality of each scene while referring to the various opinions received in the open-ended comments section of the questionnaire. We also need to think in detail about how to promote the glowing mushrooms in the Corona Disaster and how to use the 3DCG models we created this time.

8. Conclusion

Based on the feedback we received from the Faculty of Agriculture and museum staff, our future tasks are to make the 3DCG models more realistic and to improve the quality of each scene by referring to the various opinions we received in the free-text sections of the questionnaire. In addition, since we are also conducting research on AR-based exhibition support, we can expect to see AR development of the glowing mushroom PV. Finally, we would like to express our deepest condolences to our collaborator, Dr. Etsuko Harada (Faculty of Agriculture, University of Miyazaki), who passed away in October of the year before last. The Faculty of Agriculture and museum employees will be consulted in order to improve the 3DCG models' realism. Additionally, each scene will be improved in light of the feedback provided in the questionnaire's open-ended comments area. Additionally, we must carefully consider how to employ the 3DCG models we produced this time as well as how to promote the luminous mushrooms in the Corona Disaster.

References

1. 令和2年版観光白書 [Online]. https://www.mlit.go.jp/statistics/file000008.html
2. UMK テレビ宮崎　青島活性化への取り組み [Online].https://www.umk.co.jp/udoki/-2019518-1.html

Authors Introduction

Mr. Bidesh Biswas Biki

He is a master student at Department of Engineering, University of Miyazaki. His major is Mechanical Systems and Informatics. He works with Machine Learning Classification algorithms and trying to go in depth of Artificial Intelligence.

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. JSET.

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.

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 The Institute of Electronics, Information and Communication Engineers (IEICE), a delegate of the Kyushu branch of Institute of Electrical Engineers of Japan (IEEJ), a director of Japan Society for Simulation Technology (JSST).

Prof. Makoto Sakamoto

Makoto Sakamoto received the Ph.D. degree in computer science and systems engineering from Yamaguchi University. He is presently an associate professor in the Faculty of Engineering, University of Miyazaki. 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, augmented reality, entertainment computing, complex systems and so on.

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..

Parallel wave sound analysis
based on hierarchical domain decomposition method

Amane Takei*
Faculty of Engineering, University of Miyazaki,1-1, Gakuen Kibanadai-Nishi
Miyazaki, 889-2192, Japan

Akihiro Kudo
Tomakomai Collage, 443, Nishikioka, Tomakomai
Hokkaido, 059-1275, Japan

Makoto Sakamoto
Faculty of Engineering, University of Miyazaki,1-1, Gakuen Kibanadai-Nishi
Miyazaki, 889-2192, Japan

＊Corresponding Author

E-mail:takei@cc.miyazaki-u.ac.jp
http://www.miyazaki-u.ac.jp/

Abstract

We are investigating a large-scale non-steady wave sound analysis method based on the parallel finite element method by developing ADVENTURE_Sound as an opensource software. The iterative domain decomposition method is employed in the analysis code as a parallel technique. We have confirmed that the non-steady wave sound analysis code is very high-accuracy with errors within the allowable range in a numerical analysis.

Keywords: Wave sound analysis, Finite element method, Domain decomposition method, Huge-scale analysis.

1. Introduction

There is growing demand for advanced sound design, such as noise reduction and high-quality indoor and outdoor acoustic environment. It is thus necessary to understand the sound pressure distribution with high accuracy. There is also a need to reduce the costs of designing acoustic spaces [1] and electrical equipment for noise-suppression equipment. Acoustic analysis techniques have been used for the acoustic design of concert halls and noise suppression equipment due to improvements in computer hardware and software performance.

2. Governing equations and algorithm for parallel computing

In ADVENTURE_Sound, the wave-sound analysis is considered. To derive a weak form, the Galerkin method is applied to the Helmholtz equation [1]. The finite element approximation and discretized, the following equation is obtained:

$$\iiint_{\Omega_e} \nabla \Phi_h \cdot \nabla {\Phi_h}^* d\Omega_e - \frac{j\omega\rho}{Z_n} \iint_{\Gamma_e} \Phi_h {\Phi_h}^* d\Gamma_e - k^2 \iiint_{\Omega} \Phi_h {\Phi_h}^* d\Omega_e = 0. \quad (1)$$

where Φ is the speed potential that is the unknown function. k and ω are the wave number and angular frequency, ρ is the medium density, and Z_n is the specific acoustic impedance.

The equation contains complex numbers and becomes a complex symmetric matrix. In the present study, the speed potential Φ is obtained using the conjugate orthogonal conjugate gradient (COCG) method. The finite element approximation (1) is rewritten as $Ku = f$ by the coefficient matrix K, the unknown vector u, and the right-hand side vector f. Next, Ω is divided into N subdomains (Eq. (2)). Eq. (3) and (4) are obtained from Eq. (2) [2].

$$\begin{bmatrix} K_{II}^{(1)} & 0 & 0 & K_{IB}^{(1)}R_B^{(1)\mathrm{T}} \\ 0 & \ddots & 0 & \vdots \\ & & K_{II}^{(N)} & K_{IB}^{(N)}R_B^{(N)\mathrm{T}} \\ 0 & 0 & & \\ R_B^{(1)}K_{IB}^{(1)\mathrm{T}} & \cdots & R_B^{(N)}K_{IB}^{(N)\mathrm{T}} & \sum_{i=1}^{\mathrm{N}} R_B^{(i)}K_{BB}^{(i)}R_B^{(i)\mathrm{T}} \end{bmatrix} \begin{bmatrix} u_I^{(1)} \\ \vdots \\ u_I^{(N)} \\ u_B \end{bmatrix} = \begin{bmatrix} f_I^{(1)} \\ \vdots \\ f_I^{(N)} \\ f_B \end{bmatrix} \tag{2}$$

$$K_{II}^{(i)}u_I^{(i)} = f_I^{(i)} - K_{IB}^{(i)}u_B^{(i)} \quad (i = 1, \dots, N) \tag{3}$$

$$\left\{\sum_{i=1}^{\mathrm{N}} R_B^{(i)}\left\{K_{BB}^{(i)} - K_{IB}^{(i)\mathrm{T}}\left(K_{II}^{(i)}\right)^{-1}K_{IB}^{(i)}\right\}R_B^{(i)T}\right\}u_B = \sum_{i=1}^{\mathrm{N}} R_B^{(i)}\left\{f_B^{(i)} - K_{IB}^{(i)T}\left(K_{II}^{(i)}\right)^{-1}f_I^{(i)}\right\} \tag{4}$$

where ${f_B}^{(i)}$ is the right-hand vector for u_B, and $\left(K_{II}^{(i)}\right)^{-1}$ is the inverse matrix of $K_{II}^{(i)}$. Equation (4) is referred to as an interface problem and is an equation for satisfying the continuity between domains in the domain decomposition method. For simplicity, rewrite Eq. (5) as follows:

$$\begin{aligned} &Su_B = g, \\ &S = \sum_{i=1}^{N} R_B^{(i)}S^{(i)}R_B^{(i)T}, \quad S^{(i)} \\ &= K_{BB}^{(i)} - K_{IB}^{(i)T}\left(K_{II}^{(i)}\right)^{-1}K_{IB}^{(i)}. \end{aligned} \tag{5}$$

3. Numerical example

For examining ADVENTURE_Sound on a real-world problem, we model the environ-ment of acoustic experiments. The computations are performed on a 16-node (62-core) PC cluster (Intel(R) Xeon(R) CPU E5-2650L; 1.80 GHz; L2 20480 KB) with 32 GB RAM per node. The simulation statistics and the numerical are shown in Table 1 and 2, respectively. More detail thins will be shown at the conference.

Table 1 Simulation Statistic

Frequency	442[Hz]
No. of Elements	911,133
No. of DOF	1,249,959
Platform	10-node workstation cluster with Intel(R) Xeon(R) CPU E5-2650L, 1.8 GHz
No. of cores per node	16
No. of nodes	8
Main memory per node	32 [GB/node]

Table 2 Numerical result

Elapsed time	509.153 [s]
Memory requirements	0.26 [GB/node]

Acknowledgements

The present study was supported in part by a JSPS Grant-in-Aid for Scientific Research (Basic Research (B), 17H03256). The computer environment used in the present study was supported in part through a JST Adaptable and Seamless Technology transfer Program (A-STEP (Search Type), AS262Z02631H).

References

1. Amane Takei and Akihiro Kudo: Over 100 Million Degrees of Freedom Scale Wave Sound Analysis Based on Parallel Finite Element Method,.12:2(2020), 76-84. (in Japanese)
2. A. Takei, S. Sugimoto, M. Ogino, S. Yoshimura, H. Kanayama: Full Wave Analyses of Electromagnetic Fields with an Iterative Domain Decomposition Methodm, IEEE Transactions on Magnetics, 46:8 (2010), 2860-2863.

Authors Introduction

Prof. Amane Takei

Amane Takei 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 The Institute of Electronics, Information and Communication Engineers (IEICE), a delegate of the Kyushu branch of Institute of Electrical Engineers of Japan (IEEJ), a director of Japan Society for Simulation Technology (JSST).

Prof. Akihiro Kudo

Akihiro Kudo was received was received Ph.D. of engineering from Nagaoka University of Technology, 2007. He started working as an Assistant Professor in the Department of Electrical and Electronic Engineering at Tomakomai National College of Technology in 2007, and has held the position of Associate Professor there since 2011, up to the present.

His research field is acoustic engineering, and he is engaged in research on the localization of virtual sound source using headphones. He is a member of The Institute of Electronics, Information and Communication Engineers (IEICE) and Acoustical Society of Japan (ASJ).

Prof. Makoto Sakamoto

Makoto Sakamoto received the Ph.D. degree in computer science and systems engineering from Yamaguchi University. He is presently an associate professor in the Faculty of Engineering, University of Miyazaki. 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, augmented reality, entertainment computing, complex systems and so on.

Parallel full-wave electromagnetic field analysis based on hierarchical domain decomposition method

Amane Takei*
Faculty of Engineering, University of Miyazaki,1-1, Gakuen Kibanadai-Nishi
Miyazaki, 889-2192, Japan

Nanako Mizoguchi
Faculty of Engineering, University of Miyazaki,1-1, Gakuen Kibanadai-Nishi
Miyazaki, 889-2192, Japan

Kento Ohnaka
Graduate School of Engineering, University of Miyazaki,1-1, Gakuen Kibanadai-Nishi
Miyazaki, 889-2192, Japan

Makoto Sakamoto
Faculty of Engineering, University of Miyazaki,1-1, Gakuen Kibanadai-Nishi
Miyazaki, 889-2192, Japan

＊Corresponding Author

E-mail:takei@cc.miyazaki-u.ac.jp
http://www.miyazaki-u.ac.jp/

Abstract

In this presentation, a parallel full-wave electromagnetic field analysis code based on an iterative domain decomposition method is explained that is named ADVENTURE_Fullwave. A stationary vector wave equation for the high-frequency electromagnetic field analyses is solved taking an electric field as an unknown function. Then, to solve subdomain problems by the direct method, the direct method based on the LDL^T decomposition method is introduced in subdomains. The simplified Berenger's PML is introduced which these eight corners are given the average value of all PML's layers. And, we show a numerical example of a microwave. More detail will be shown in the conference.

Keywords: Electromagnetic field analysis, Finite element method, Domain decomposition method, Huge-scale analysis.

1. Introduction

Electromagnetic field analysis based on a numerical analysis method, such as the finite element method, has become widespread [1] due to recent improvements in computer performance and numerical calculation technology. In the case of accurately reproducing an analysis model of complicated shape, it is necessary to use many small the elements. In the case of analyzing the state of electromagnetic waves propagation in a wide range, a wide analysis domain is examined. Furthermore, to perform a high-accuracy analysis, it is necessary to model the analysis domain with a sufficiently small element for the wave-length, and, in this case, the number of elements also increases. Increasing the number of elements increases the scale of the problem. Therefore, a method that can calculate large-scale problems has come to be demanded. Moreover, large-scale problems must be solved with high accuracy. In the presentation, a large-scale analysis code: ADVENTURE_Fullwave is introduced, and detail of the parallel algorism is shown.

2. Governing equations and algorithm for parallel computing

In ADVENTURE_Fullwave, the full-wave analysis based on an E method [1] is considered. $\mathbf{E}_{\mathrm{h}}$ and $\mathbf{J}_h$ are finite element approximations of electric field $\mathbf{E}$ [V/m] and current density $\mathbf{J}$ [A/m^2], respectively. The permeability is given by $\mu = \mu_0\mu_r$ [H/m], μ_0 is the vacuum permeability [H/m], and μ_r is the relative permeability. The complex permittivity is given by $\varepsilon = \varepsilon_0\varepsilon_r - \sigma/j\omega$ [F/m], ε_0 is the vacuum permittivity [F/m], ε_r is the relative permittivity, and ω is the angular frequency [rad/s]. The following equation is the finite element equation to be solved:

$$\iiint_{\Omega} (1/\mu)\mathrm{rot}\mathbf{E}_{\mathrm{h}} \cdot rot\mathbf{E}_{\mathrm{h}}{}^{*}dv - \omega^2 \iiint_{\Omega} \varepsilon\mathbf{E}_{\mathrm{h}} \cdot \mathbf{E}_{\mathrm{h}}{}^{*}dv = j\omega \iiint_{\Omega} \mathbf{J}_h \cdot \mathbf{E}_h{}^{*}dv. \tag{1}$$

The equation contains complex numbers and becomes a complex symmetric matrix. In the present study, the electric field $\mathbf{E}$, which is unknown, is obtained using the conjugate orthogonal conjugate gradient (COCG) method. The finite element approximation (1) is rewritten as $Ku = f$ by the coefficient matrix K, the unknown vector u, and the right-hand side vector f. Next, Ω is divided into N subdomains (Eq. (2)). Eq. (3) and (4) are obtained from Eq. (2).

$$\begin{bmatrix} K_{II}^{(1)} & 0 & 0 & K_{IB}^{(1)}R_B^{(1)\mathrm{T}} \\ 0 & \ddots & 0 & \vdots \\ & & K_{II}^{(N)} & K_{IB}^{(N)}R_B^{(N)\mathrm{T}} \\ 0 & 0 & & \\ R_B^{(1)}K_{IB}^{(1)\mathrm{T}} & \cdots & R_B^{(N)}K_{IB}^{(N)\mathrm{T}} & \sum_{i=1}^{\mathrm{N}} R_B^{(i)}K_{BB}^{(i)}R_B^{(i)\mathrm{T}} \end{bmatrix} \begin{bmatrix} u_I^{(1)} \\ \vdots \\ u_I^{(N)} \\ u_B \end{bmatrix} = \begin{bmatrix} f_I^{(1)} \\ \vdots \\ f_I^{(N)} \\ f_B \end{bmatrix} \tag{2}$$

$$K_{II}^{(i)}u_I^{(i)} = f_I^{(i)} - K_{IB}^{(i)}u_B^{(i)} \quad (i = 1, \dots, N) \tag{3}$$

$$\left\{\sum_{i=1}^{\mathrm{N}} R_B^{(i)} \left\{K_{BB}^{(i)} - K_{IB}^{(i)\mathrm{T}} \left(K_{II}^{(i)}\right)^{-1} K_{IB}^{(i)}\right\} R_B^{(i)T}\right\} u_B = \sum_{i=1}^{\mathrm{N}} R_B^{(i)} \left\{f_B^{(i)} - K_{IB}^{(i)T} \left(K_{II}^{(i)}\right)^{-1} f_I^{(i)}\right\} \tag{4}$$

where $f_B{}^{(i)}$ is the right-hand vector for u_B, and $\left(K_{II}^{(i)}\right)^{-1}$ is the inverse matrix of $K_{II}^{(i)}$. Equation (4) is referred to as an interface problem and is an equation for satisfying the continuity between domains in the domain decomposition method. For simplicity, rewrite Eq. (5) as follows:

$$\begin{aligned} &Su_B = g, \\ &S = \sum_{i=1}^{N} R_B^{(i)}S^{(i)}R_B^{(i)T}, \quad S^{(i)} \\ &= K_{BB}^{(i)} - K_{IB}^{(i)T}\left(K_{II}^{(i)}\right)^{-1}K_{IB}^{(i)}. \end{aligned} \tag{5}$$

3. PML

3.1. *Berenger's PML*

The PML can be used to create an absorbing boundary by surrounding the analysis domain with a PML. From the viewpoint of the accuracy of the obtained solution, the PML is currently the most effective absorbing boundary condition. Although Berenger's PML is originally proposed as an absorbing boundary condition for the FDTD method, in the present study, we apply a finite element method dealing with an unstructured grid, we propose a simplified method omitting the directionality of electric conductivity given to the PML and confirm its effectiveness.

Berenger's PML [2] stacks several PMLs outside the analysis domain and gradually sets a large value of electric conductivity according to the outer layer so that the outermost wall can be surrounded with a perfect conductor wall without reflecting electromagnetic waves. Figure 1 shows a schematic diagram of Berenger's PML absorbing boundary.

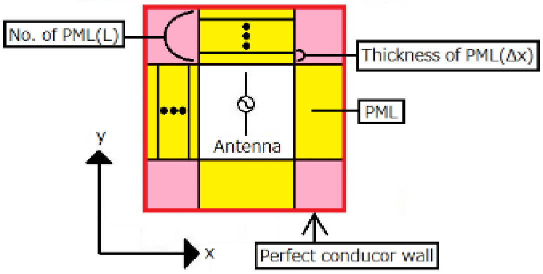


Fig. 1. PML absorbing boundary

In this paper the distribution of the electric conductivity for PML is expressed as follows:

$$\sigma = \sigma_{max}\left[\frac{\left(L - \hat{L}(x)\right)\Delta x}{L\Delta x}\right]^M \tag{6}$$

where Δx is the thickness of PML 1, L is the number of layers of the PML, $\hat{L}(x)$ is a coefficient determined by position x, and $\hat{L}(x)$ = 0 at the position of the Lth layer,

$\hat{L}(x)$ = 1 at the position of the (L-1)th layer, and $\hat{L}(x)$ = L -1 at the position of the first layer.

Moreover, σ_{max} is the maximum value of the electric conductivity for the PML, and M is the degree distribution of electric conductivity. This equation is used to determine the electric conductivity of each layer of the PML.

The parameters to be determined as the parameters of the PML are the thickness Δx of PML 1, the number L of PML layers, the maximum electric conductivity σ_{max} of the PML, the degree M distribution of the electric conductivity, the reflection coefficient R [dB] between the PML of the outermost layer, and the perfect conductor wall. The reflection coefficient R is approximated as follows:

$$|R(\phi)| \cong exp\left[-\frac{2\sigma_{max}L\Delta x}{(M+1)\varepsilon_0 c}\cos\phi\right] \qquad (7)$$

where ϕ is the incident angle of the electromagnetic wave, and c is the speed of light. Since we cannot decide the incident angle for an arbitrary incident wave, ϕ = 0, a reflection coefficient for perpendicular incidence is used as a reference. Moreover, since the M that gives the distribution of the electric conductivity causes the calculation accuracy to deteriorate if the change of the electric field in the PML is too steep, M is approximately 2 to 4. If the number of layers L is too large, more memory will be required, and if L is too small, it will not function adequately as an absorbing boundary. There are many cases where the concrete number of L is set to 4 to 16. The thickness Δx of PML 1 is a constant thickness of all layers.

We set the reflection coefficient R(0) according to the required accuracy. Upon determining the above parameters, the maximum electric conductivity σ_{max} is given as follows:

$$\sigma_{max} = -\frac{(M+1)\varepsilon_0 c}{2L\Delta x}\ln|R(0)| \qquad (8)$$

In the present study, we construct a PML using Eq. (6) through Eq. (8) with L = 9, M = 4, and Δx = λ/10. However, in order to reduce the analysis scale, we examine the optimum value of L in the next section.

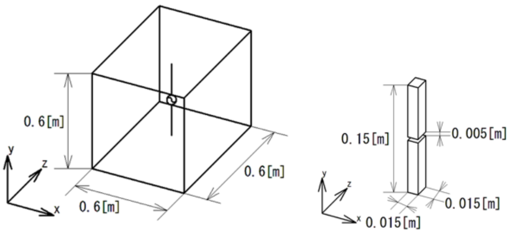
3.2. *Numerical results*

We assign the PML to the dipole antenna model. The analysis domain is a cube of length 0.6 [m] so that the distance from the antenna to the innermost PML matches the wavelength. The current density is applied to the antenna as a current source as follows:

$$I(y) = I_0 \cos\left(\frac{2\pi}{\lambda}y\right) \quad : -l \leq y \leq l \qquad (9)$$

where I_0 = 0.08 [A/m^2], λ is the wavelength, and l is the length from the feeding point to the antenna tip.

The analysis frequency is 1 [GHz], and the length of the antenna is 0.15 [m], which is the half wavelength. Here, mesh division is performed so that the maximum side length of the element is 1/20 of the wavelength. The analysis domain's boundary is a perfect conductor. Figure 4 shows a schematic diagram of the dipole antenna model.



(a) Analysis domain (b) Antenna

Fig. 2. Dipole antenna model

We assign PMLs to the domain boundary as shown in Fig. 2(a). The plane portion of the PML at the domain boundary overlaps a number of flat plates according to the number of layers, and the corner portion of the PML is one rectangular parallelepiped or cube. The boundary of the outermost layer of the PML is a perfect conductor wall. We perform performance evaluation by setting the thickness of one layer to be 0.03 [m] and the PML to have L = 9 (hereinafter a PML with L layers is abbreviated as PML(L)). Table 1 lists the number of elements and the degree of freedom of the analysis model.

Table 1. Number of elements and DOFs of the dipole antenna model

	PML(0): Perfect conductor wall	PML(9)
No. of Elements	4,669,759	26,899,669
DOFs	5,506,368	31,703,550

In (8), we set L = 9, Δx = 0.03, M = 4, and R(0) = -120 [dB], which yields the maximum electric conductivity σ_{max} to PML(9). In addition, we decide the electric conductivity of each layer using (6). In this study, we set the average value of each layer to the electric conductivity of the corner portion. We evaluate the performance of the PML based on the reflection coefficient obtained using the S_{11} parameter[3]. The observation point of the S_{11} parameter is on the x-axis 1 cm inside of the PML. The computing environment in the present study is a 25-PC cluster with Intel Core i7-2600K multi-core CPUs (total: 100 cores) and 32 GB memory. Table 2 lists the reflection coefficient, the CPU time, and the memory size.

Table 2. Results for reflection coefficient, CPU time, and memory size

	PML(0): Perfect conductor wall	PML(9)
Reflection coefficient [dB]	0	-18.65
CPU time [s]	1,278	18,787
Memory size [MB/core]	44.3	227.3

When the domain boundary is PML(0), i.e., when it is a perfect conductor wall, S_{11} = 1, so that the reflection coefficient is 0 [dB]. On the other hand, when the domain boundary is PML(9), the reflection coefficient is -18.65 [dB]. The design target reflection coefficient of the antenna, for example, is generally approximately -10 to -20 [dB], and in the present study, we use a reflection coefficient of approximately -10 to -20 [dB]^3. Thus, PML(9) can obtain sufficient absorption performance. On the other hand, in comparing with PML (0), PML (9) increases the amount of memory used and computation time, depending on the absorbing layer applied. Figure 3 shows a visualization diagram of the electric field obtained by analysis.

In Figure 3, the left-hand side shows PML(9) at the boundary edge and the electric field propagates from the dipole antenna to the free space. On the other hand, the right-hand side of Figure 3 shows the mode when the dipole antenna is enclosed by a perfect conductor wall.

Next, we perform the directivity evaluation of the dipole antenna by error evaluation using the theoretical solution in the far field. The error evaluation of the far field uses the E plane.

The theoretical solution^3 of the far field of the E plane is as follows:

$$E_\theta = j60I\frac{e^{-jkr}}{r}\cdot\frac{\cos\left(\frac{\pi}{2}\cos\theta\right)}{\sin\theta} \tag{10}$$

where j is the imaginary unit, I is the current, and r is the distance from the feeding point. The approximate distance r to the far-field peak of the Fresnel's region ($2\,l^2/\lambda\, <\, r$) is 0.250 [m], if the dimension l (= 0.150 [m]) of the dipole antenna is not ignored. Moreover, k is the wave number and is given by $k = 2\pi/\lambda$. The directivity evaluation is performed by comparing the numerical analysis solution with the theoretical solution on the E plane. Figure 4 shows a plot of the numerical analysis solution e_θ and the theoretical solution E_θ in increments of 1 [deg].

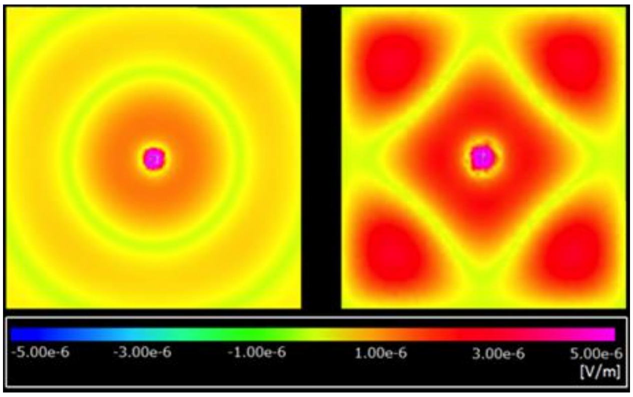


Fig. 3. **V**isualization of the analysis result (electric field) (Left: PML(9), Right: PML(0) (perfect conductor wall))

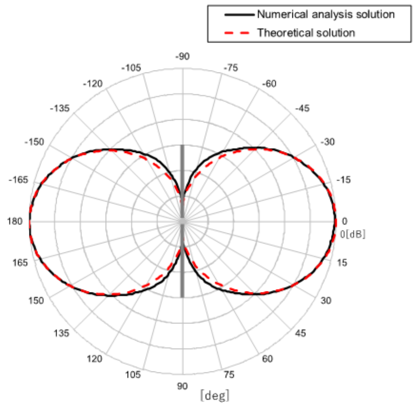


Fig. 4. Numerical and theoretical solutions in the E plane

The directivities of the numerical and theoretical solutions agree very well. The range of θ, which is the far field far beyond the Fresnel's region, can be expressed by (11). The lower limit θ_{Min} is $\arcsin(\,2\,l^2\,/\,r\lambda) + 90 \cong -57$ [deg], and the upper limit θ_{Max} is $90 - \arcsin(\,2\,l^2\,/\,r\lambda) \cong 53$ [deg]. The average error rate E_{err} in this range is obtained by (12). As a result, the average error rate is 1.70 [%], and it is shown that a highly accurate solution can be obtained.

$$\arcsin\left(\frac{2\,l^2}{r\lambda}\right) + 90 \leq \theta \leq 90 - \arcsin\left(\frac{2\,l^2}{r\lambda}\right) \tag{11}$$

$$E_{err} = \frac{\sum_{i=\theta_{Min}}^{\theta_{Max}} \frac{|e_i - E_i|}{E_i}}{\theta_{Max} - \theta_{Min} + 1} \times 100 \quad [\%] \tag{12}$$

In the calculations shown in Figure 2, we used a dipole antenna model with PML(9). Here, we find the optimum *L* from the average error rate in the far field and the reflection coefficient of PML(*L*) by a parameter study using the number of PMLs. Table 3 shows the number of elements for each *L*, the number of degrees of freedom of the edge, the error rate, the reflection coefficient, the calculation time, and the number of iterations of the COCG method applied to the interface problem.

Table 3. Numerical model data and results

	PML(9)	PML(8)	PML(7)
No. of elements	26,899,669	24,184,687	21,533,641
DOFs	31,703,550	28,506,352	25,383,890
Average error rate [%]	1.70	3.81	12.87
Reflection coefficient [dB]	-18.65	-15.79	-15.04
CPU time [h]	5.22	3.77	2.81
No. of iterations	46,508	37,755	30,695
Memory size [MB/core]	227.3	204.6	182.6

From Table 3, PML(9) is the case with the best far field accuracy. When the allowable range of the error rate is less than 5 [%], which is the allowable range of numerical analysis error, since PML(7) has a reflection coefficient of less than -15 [dB], the PML functions sufficiently. However, the error rate exceeded the allowable range. We can find that PML(8) is optimal because it has a better calculation time and iteration count than PML(9).

4. Analysis for Convergence Comparison

4.1. *Results of Closed-space model A*

The number of iterations and computation time for the closed-space model A with subdomain sizes of 100, 200, 500, and 1,000 are shown in Table 4. The convergence decision value is 1.0e-03. As shown in Figure 5, the convergence performance improves as the size of the subdomain increases.

Table 4. Result of closed-space model A(2,549,500 Elemens.)

Subdomain size	Num of iterations	Computation time [sec]
100	10,459	453
200	9,202	414
500	6,847	342
1,000	4,848	294

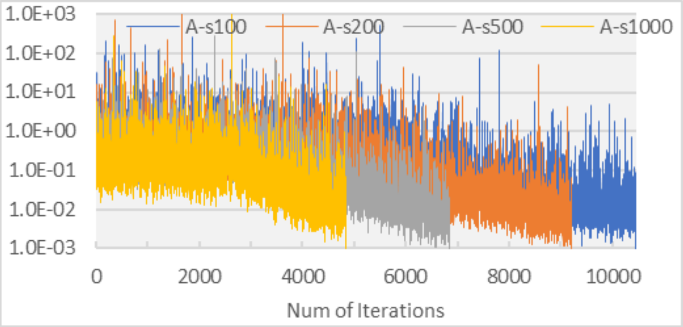


Fig. 5. The num. of iterations and Convergence in model A

4.2. *Results of liberated-space model B*

The number of iterations and computation time for the liberated-space model B with subdomain sizes of 100, 200, 500, 1000, 2000, 5000, and 10000 are shown in Table 5. The convergence decision value is 1.0e-03. As shown in Figure 6, the convergence performance improves as the size of the subdomain increases. More details will be shown at the conference.

Table 5. Result of closed-space model B(15,712,684 Elemens.)

Subdomain size	Num of iterations	Computation time [sec]
100	6,578	1,039
200	10,130	1,636
500	3,581	712
1,000	5,324	1,385
2,000	2,916	1,141
5,000	5,854	3,627
10,000	1,002	973

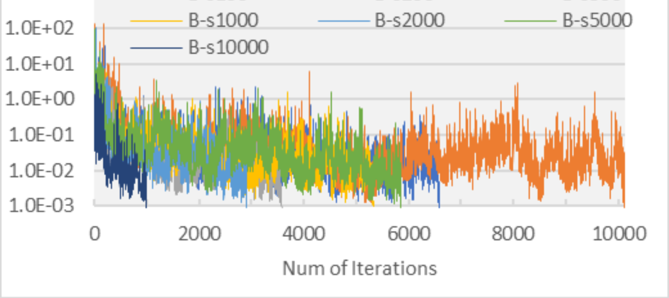


Fig. 6. The num of iterations and Convergence in model B

Acknowledgements

The present study was supported in part by a JSPS Grant-in-Aid for Scientific Research (Basic Research (B), 17H03256). The computer environment used in the present study was supported in part through a JST Adaptable and Seamless Technology transfer Program (A-STEP (Search Type), AS262Z02631H).

References

1. A. Takei, S. Sugimoto, M. Ogino, S. Yoshimura, H. Kanayama: Full Wave Analyses of Electromagnetic Fields with an Iterative Domain Decomposition Methodm, IEEE Transactions on Magnetics, 46:8 (2010), 2860-2863.
2. A. Takei, I. Higashi, M. Aikawa and T. Yamada, "Microwave analysis based on parallel finite element method," Journal of Advanced Simulation in Science and Engineering, 6:1 (2019), 215-233.

Authors Introduction

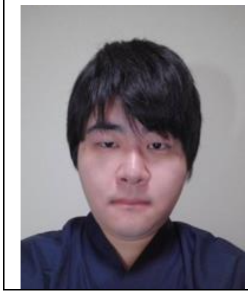
Prof. Amane Takei

Amane Takei 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 The Institute of Electronics, Information and Communication Engineers (IEICE), a delegate of the Kyushu branch of Institute of Electrical Engineers of Japan (IEEJ), a director of Japan Society for Simulation Technology (JSST).

Ms. Nanako Mizoguchi

She received her B.S. degree in Engineering in 2023 from the Faculty of Engineering, University of Miyazaki in Japan. She is a master course student of Graduate school of Engnnering, University of Miyazaki in Japan.

Mr. Kento Ohnaka



He received his B.S. degree in Engineering in 2022 from the Faculty of Engineering, University of Miyazaki in Japan. He is a master course student of Graduate school of Engnnering, University of Miyazaki in Japan.

Prof. Makoto Sakamoto

Makoto Sakamoto received the Ph.D. degree in computer science and systems engineering from Yamaguchi University. He is presently an associate professor in the Faculty of Engineering, University of Miyazaki. 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, augmented reality, entertainment computing, complex systems and so on.

Changes in the Behavior of a Small Number of Molecular Systems

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

Molecular systems in chemical reaction systems have been considered continuous systems. However, chemical reactions in living organisms involve small molecules and cannot be considered a continuous system in some cases. In this study, we examine the behaviors of the two-party Lotka-Volterra model with a small number of molecules. We then show that there are cases in which intrinsic fewness is dominant.

Keywords: systems with small number of elements, Abstract Rewriting System on Multisets

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; colour 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.

1.1. *Abstract Rewriting Systems on Multisets, ARMS*

Abstract Rewriting System on Multisets, ARMS is a model of computation of chemical re- actions, in which floating *molecules* can interact with each other according to given reaction rules. Techni- cally in ARMS, a chemical solution is a finite multiset of elements denoted by symbols from a given al- phabet, $A = \{a, b, \ldots, \}$; these elements correspond to *molecules*. Reaction rules that act on the molecules are specified in *ARMS* by reaction rules.

Let A be an *alphabet* (a finite set of abstract symbols). A *multiset* over a set of objects A is a mapping $M : A \longmapsto \mathbf{N}$, where $\mathbf{N}$ is the set of natural numbers, $\mathbf{N}$, 0, 1, 2,.... The number $M(a)$, for $a \in A$, is the *multiplicity* of object a in the multiset M. We denote $A^{\#}$ by the set of all multisets over A, including the empty multiset, $\emptyset$, defined by $\emptyset(a) = 0$ for all $a \in A$.

A multiset $M : A \mapsto N$, for $A = \{a_1, . . . , a_n\}$ is represented by the vector $w = (M(a_1)\ M(a_2) . . . M(a_n))$.

The union of two multisets $M_1, M_2 : A \mapsto N$ is addition of vectors w_1 and w_2 that represent the each multisets respectively. If $M_1(a) \leq M_2(a)$ for all $a \in A$, then we say that multiset M_1 is included in multiset M_2 and we write $M1 \subseteq M2$.

A reaction rule $u \rightarrow v, u, v \in A\#$ is a vector r, $r = -u+v$. Note that u and v can also be zero vector (empty). For example, the reaction a $b \rightarrow c$ is the vector of (-1 -1 1)=-(1 1 0) + (0 0 1).

A reaction is the addition of vectors $M \in A\#$ and $r \in R$, and it can be defined only when $r \subseteq M$. We can define a relation over $A\#$ a : ($\rightarrow$): for $M, M_l \in A\#, r \in R$. ARMS is a simple and conventional model of computation, but it can describe complex behaviors [3].

Because ARMS is rule based, modeling is done by creating "*if-then*" rules. There is no need to transform to differential equations. Less computation time. The computation in ARMS we write $M \rightarrow M_l$ *if* $M_l = (M + r)$.

1.2. *Lotoka-Volterra*

Mathematical ecology has used the Lotka-Volterra, LV equation is the basic equation1). LV equation is a differential equations with the predator as y and the prey as x; which is shown in the equation (1) and (2).

$$\frac{dx}{dt} = k_1x - k_2xy, \quad (1)$$

$$\frac{dx}{dt} = k_1x - k_2xy, \quad (2)$$

The equilibrium point of the LV equation can be found as the point where the derivative of x and y is zero; as shown in equation (3) and (4);

$$0 = \frac{dx}{dt} = x(k_1 - k_2y), \quad (3)$$

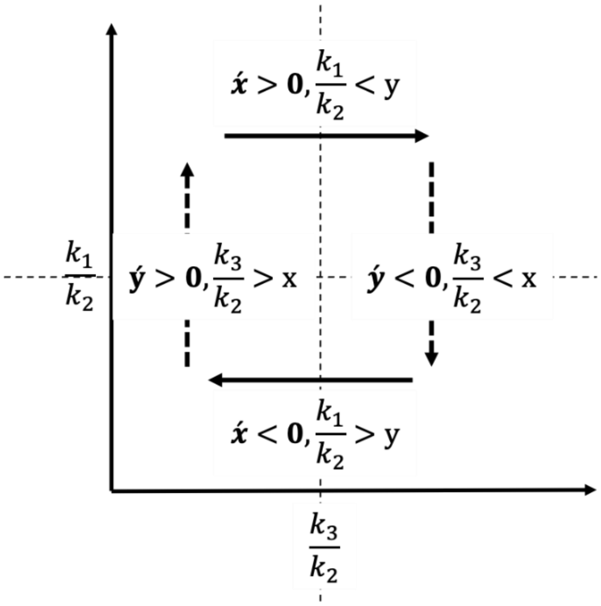
$$0 = \frac{dy}{dt} = y(k_2x - k_3). \quad (4)$$

We obtain the equilibrium points,

$$(x, y) = \{(0,0), \left(\frac{k_3}{k_2}, \frac{k_1}{k_2}\right)\} .$$

The positive and negative differential values of x or y around equilibrium points give the behavior of the LV equation.

If y is larger than k_1 / k_2, x decreases, and if y is smaller, it increases. On the other hand, if x is larger than k_3 / k_2, y increases, and if it is smaller, y decreases. Therefore, x and y oscillate around the equilibrium point; the solution to this equation is periodic [1] (Fig. 1).



$\dot{x} > 0, \frac{k_1}{k_2} < y$
$\frac{k_1}{k_2}$
$\dot{y} > 0, \frac{k_3}{k_2} > x$
$\dot{y} < 0, \frac{k_3}{k_2} < x$
$\dot{x} < 0, \frac{k_1}{k_2} > y$
$\frac{k_3}{k_2}$

Fig.1 Behavior of LV

2. Lotoka-Volterra with small number elements

Reaction rules of Lotoka-Volterra LV is { x $\rightarrow x, x: r_1; x, y \rightarrow y, y: r_2 ; y \rightarrow$ nil: r_3}, where reaction constants of each rule is k_1, k_2, k_3 and nil stands for empty set. The rules of VAS for LV are (1, 0), (-1, 1) and (0, -1). According to chemical rate equation, probability of applying each rule is xk_1/R, k_2xy/R and yk_3/R, where $R = xk_1 + xyk_2 + yk_3$.

In case x, y >>, the behaviors of LV with VAS is equal to the differential equations in the section 1.2. However, x, y <<, the system shows different behaviors. For such a case, stochastic ARMS is suitable for examining.

In LV with stochastic ARMS, when the initial point (*x*, *y*) is large, the behavior is almost same as the differential equations (Fig.2 (Top)). While, initial point is small, the behavior is different (Fig.2 (Below)).

This instability caused by the initial point has pointed out by using stochastic model (Gillespie method) or Markov process. However, the mechanism of how this instability emerges have not known well.

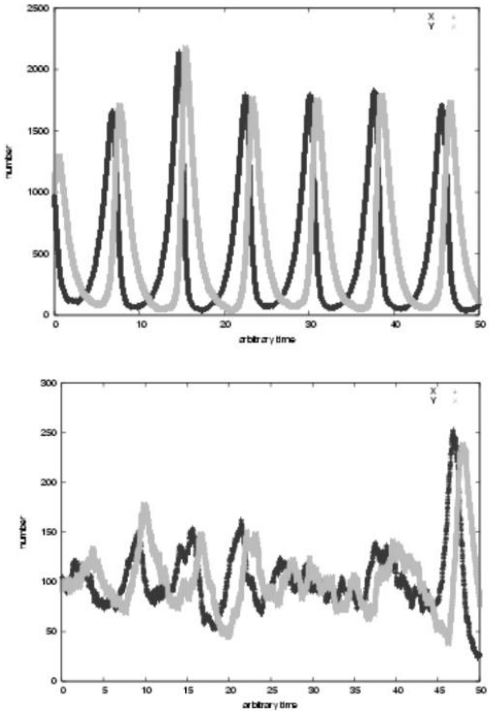


Fig. 2. Behavior of LV: (Top) x and y are large, (Below) x and y are small.

References

1. Suzuki, Y. (2010). Fluctuation Induced Structure in Chemical Reaction with Small Number of Molecules. In: Peper, F., Umeo, H., Matsui, N., Isokawa, T. (eds) Natural Computing. Proceedings in Information and Communications Technology, vol 2. Springer, Toky

Authors Introduction

Yasuhiro Suzuki

He is Associate professor of Graduate School of Informatics, Nagoya University

Retrieval by Sensory Information

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

Until now, information retrieval has been conducted by language. In recent years, information retrieval using smart microphones has also developed. However, voice-based retrieval is conducted by converting voice into linguistic information. In this study, we propose a search based not on linguistic information but on language sensitivity information. Language sensitivity information is a way of saying things. We perceive differences in the way we say the same word. In other words, the difference in how we say a word is the sensory information of the language. This study proposes a method for extracting and retrieving sensitive information from language.

Keywords: Sensitivity, Kansei, Tactile Score, Information retrieval

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.

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).

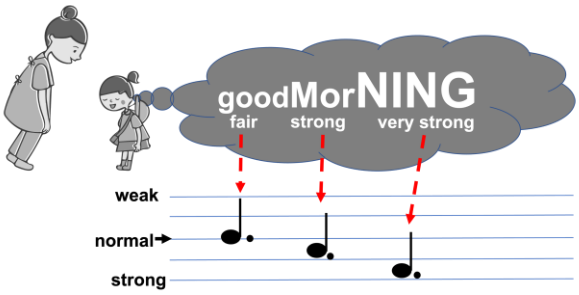


Fig. 1 Tactile Score

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 (Fig.2). 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.3). If we are walking on the grass in a park and the hardness of the grass is constant, we will feel "smooth".

3. Tactile tag

Any object with spatiotemporal variation can be converted into a tactile score. Hardness and roughness can be used as tactile tags.

By using tactile qualities, various objects can be tagged with tactile tags. For example, the tactile quality of a piece of music can be used to attach a tactile tag to the music. We created a music database for our experiments. The music used was music that was past its copyright protection period.

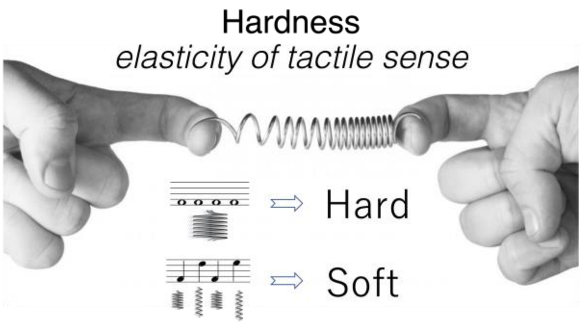


Fig. 2 Hardness, read from Tactile Score

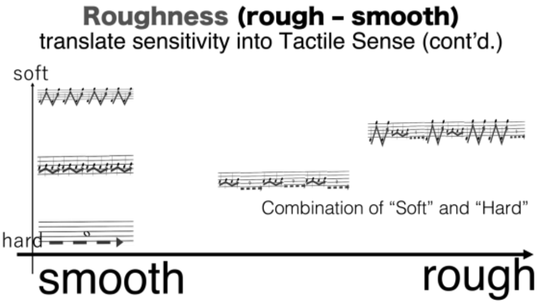

Roughness (rough - smooth)
translate sensitivity into Tactile Sense (cont'd.)
soft
hard
Combination of "Soft" and "Hard"
smooth
rough

Fig. 3 Roughness, read from Tactile Score

The data structure was a standard music database with the addition of the average Hardness and the average roughness values; the data structure is the usual music database, add the average Hardness and the average roughness;

<music name, information about the music (performer and so on), average hardness, average roughness>.

4. Tactile Query

Music can be retrieved by touching the audio or force-sensitive touch panel. Convert the time variation of vertical force to voice or force-sensitive touch panel into a tactile tag. The converted tactile tags can be used to search the database. In this search method, the search is not performed by the name of the music piece. The search is performed by way of saying or touching (Fig.4).

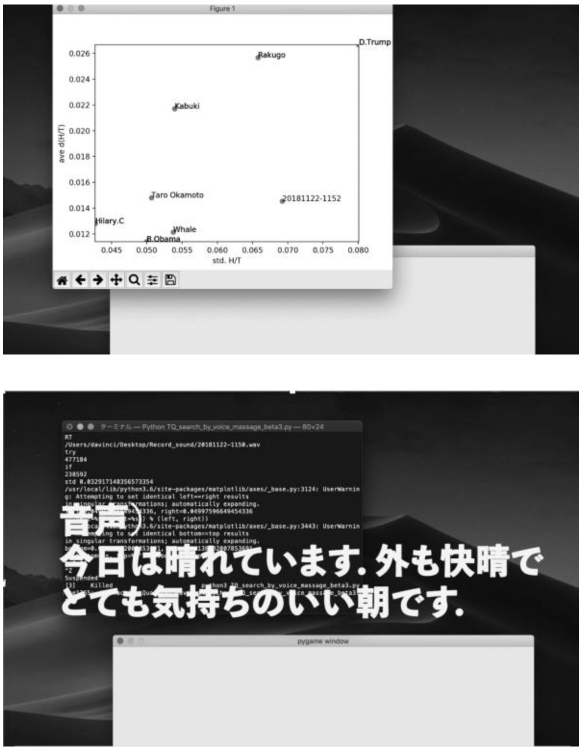


Figure 1
D.Trump
Rakugo
Kabuki
Taro Okamoto
20181122-1152
Hilary.C
Whale
B.Obama
0.026
0.024
0.022
0.020
0.018
0.016
0.014
0.012
ave d(H/T)
0.045
0.050
0.055
0.060
0.065
0.070
0.075
0.080
std. H/T
ターミナル — Python TQ_search_by_voice_massage_beta3.py — 80×24
pygame window

```
RT
/Users/davinci/Desktop/Record_sound/20181122-1150.wav
try
477184
if
238592
std 0.032917148356573354
/usr/local/lib/python3.6/site-packages/matplotlib/axes/_base.py:3124: UserWarnin
g: Attempting to set identical left==right results
in singular transformations; automatically expanding.
[illegible]336, right=0.04997596649454336
[illegible]) % (left, right))
[illegible]3.6/site-packages/matplotlib/axes/_base.py:3443: UserWarnin
[illegible]ing to set identical bottom==top results
in singular transformations; automatically expanding.
[illegible]
^Z
Suspended
[3]    Killed                  python3 TQ_search_by_voice_massage_beta3.py
[illegible]
```

音声）
今日は晴れています. 外も快晴で
とても気持ちのいい朝です.

Fig. 4 Tactile Query and result of retrieval; (Top) The map of Tactile quality, "20181122-1152" is the tactile quality of inputted voice and others are references.(Below) Japanese language in the middle, inputted language by voice ("it is a fine day and very good morning) and the system plays a music whose tactile quality is similar to the inputted voice

References

1. Y. Suzuki, R. Suzuki, Tactile score: a knowledge media for tactile sense. Springer Japan, 2014.

Authors Introduction

Yasuhiro Suzuki

He is Associate professor of Graduate School of Informatics, Nagoya University, Japan

Report of a robot competition on the problem of garbage in the sea and verification of learning effects

Takayuki Matsuo
National Institute Technology, Kitakyushu College,
2-20-1, Shii, Kokuraminami-ku, kitakyushu-shi, Fukuoka 802-0985, Japan

Masanori Sato
Nagasaki Institute of Applied Science,
536, Aba-machi, Nagasaki-shi, Nagasaki 851-0193, Japan

Masayoshi Ozawa
Kobe City College of Technology,
8-3, Gakuenhigashi-mach, Nishi-ku, Koube-shi, Hyogo 651-2102, Japan

Seiichiro Miura
National Institute of Technology, Tokuyama College
3538 Gakuendai, Shunan-shi, Yamagichi, 745-8585, Japan

Masakazu Arima
Osaka Metropolitan University,
1-1 Gakuen-machi, Naka-ku, Sakai-shi, Osaka, 599-8531, Japan

Kazuo Ishii
Department of Life Science and System Engineering, Graduate School of Kyushu Institute of Technology,
2-4, Hibikino, Wakamatsu, Kitakyushu-city, Fukuoka 808-0196, Japan

E-mail: matsuo@kct.ac.jp, sato_masanori@nias.ac.jp, ozawa-m@kobe-kosen.ac.jp, miura@tokuyama.ac.jp,
marima.marine@omu.ac.jp, ishii@brain.kyutech.ac.jp

Abstract

In the Junior League of the Eighth Underwater Robot Competition held in 2022, a marine cleaning robot competition was held under the theme of "life under the surface," the 14th goal of the SDGs. Specifically, we held a poster session on the problem of garbage in the sea and a robot competition to compete in garbage collection ability, and competed for points. In this paper, we give an overview of the junior league and discuss its learning effects.

Keywords: Robot competition, Education of robotics

1. Introduction

In recent years, the problem of ocean debris has become more serious. Among them, coastal pollution caused by stray litter and the harmful effects of microplastics on marine life are important issues that need to be solved quickly by humankind. The SDGs adopted by the United Nations Summit in September 2015 set 17 goals for a sustainable and better world [1]. The 14th goal, "Life under water," includes the problem of ocean debris. We are working on robot education for middle and high school students through robot competitions, and the 8th

Underwater Robot Junior League is a competition with the theme of underwater robots. Specifically, the robot competition was held on the theme of ocean debris. The junior league is divided into poster presentation and robot competition, and the winner is decided by the total score. The poster presentation investigated the problems that society has at present about the 14th goal of SDG, and proposed a robot to solve them, and in the robot competition, the proposed robot was embodied, and ocean debris was picked up.

In this paper, the outline of the junior league is explained, and the results of this competition are shown, and the learning effect of this competition is verified by analyzing the questionnaire conducted before and after the competition.

2. Outline of Junior League

The junior league was ranked by the total score of the poster presentation and the robot competition[2]. Here, the poster presentation and the outline of the robot competition are explained.

2.1. *Poster Presentation*

Students from different cultures, such as technical college, regular high school and technical high school, participate in the junior league. Poster presentations were held with the aim of encouraging people to find new values by sharing and absorbing different ways of thinking. The poster presentations are shown in Fig.1 The following items were required to be included in the poster and were subject to review:

- Poster title, team name, school name, team member name, teacher name
- Investigations into debris problems in the ocean and rivers,
- Issues to be solved in garbage collection,
- Ideas to solve the issues,
- Proposal for garbage collection robot,
- Conclusions,
- References.

Fig. 1 Poster presentation

2.2. *Robot Competition*

Two teams compete in a competitive format, with robots starting from each team's territory to pick up debris in the debris area. After the competition, players compete for overall points by adding points according to the type and amount of debris collected. The competition field uses a pool 2.2m long, 4.5m wide as shown in Fig. 2. The depth of the pool is about 60cm to 70cm. The competition field is divided into team area for each team and debris areas. The team area is the area where the robot starts, and the debris area is the area where the debris, located in the center of the competition field. And, the short side of the competition field in the team area will be the end line (red line). At the start of the competition, debris floats randomly in the debris area of the competition field, and there are three types of debris: "PET bottles", "jelly containers" and "Styrofoam balls". The type and number of floating debris, and the score when collected are shown in Table 1. Debris floating on the competition field is collected by robots, and if the condition of the debris satisfies the following conditions, points will be added as "collected debris".

- Debris that the robot lifts above the water surface is called "collected debris". If part of the debris touches the surface of the water or is submerged in water, it will not be considered as "collected debris".

- During the competition time, if a part of the robot touches the end line of the team area, the team members standing on the end line side may place the "collected debris" outside the competition field from the robot.

- At the end of the competition, the robot raises it above the water surface, and the debris that is still loaded on the robot

Figure 3 shows the scene during the competition.

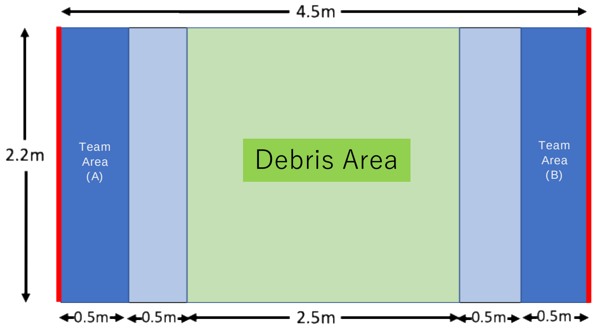


Fig. 2 Competition field

Table1 Debris type and score per piece

Debris type	Quantity	Score per piece
PET bottles	10	10 point
jelly containers	30	5 point
Styrofoam balls	50	3 point

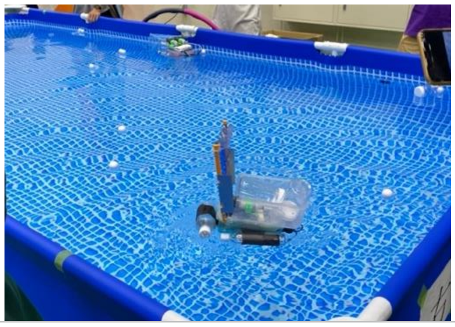


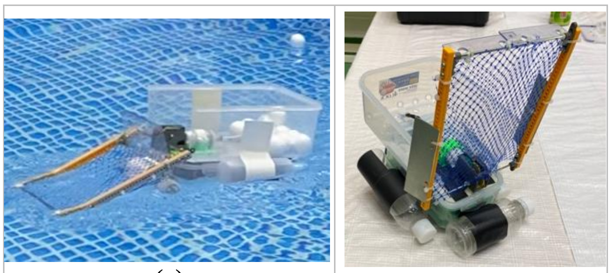
Fig. 3 Scene during the competition

2.3. *Result of Junior League*

The results of the competition of Junior League are shown in Table 2. The 8th competition was attended by 18 teams with a total of more than 120 participants. A lively exchange of ideas was observed in the poster session, and robots with various ideas were active in the robot competition. Most of the winning teams were industrial schools, and the teams with many students with basic knowledge such as microcomputers had an advantage in robot development. However, prior learning was conducted remotely for schools without knowledge, and there were few teams where robots did not work at all. The winning and runner-up robots are shown in Figure 4. Both teams were robots that scooped up debris with their arms and collected it.

Table 2 Award team list

Award	Team Name	Affiliation
Winner	Noukou G's	Tabuse Agricultual Technical Highschool
Runner-up	Know Sea's	Tabuse Agricultual Technical High school
Third place	Kokkaginoushi no tsudoishi Minamata-shibu	Minamata High school
Robot competition 1st place	Noukou G's	Tabuse Agricultual Technical High school
Robot competition 2nd place	Know Sea's	Tabuse Agricultual Technical Highschool
Poster presentation 1st place	Umi no Gomi wo Nakushitai	Jyoto High School
Poster presentation 2nd place	M・E・C Suirobo Doujinkai	Osaka Metropolitan University College of Technology
Governor of Yamaguchi Prefecture Award	Tokuyama Kosen Mechatro System Club	Natinal Inst. of Tech., Tokuyama College
Mayer of Iwakuni Award	Gankou “Damashii”	Iwakuni Technical High school
Special Award	Dept. of Electronic Mechanics	Mifune High school
Fiting-spirit Award	Team 0	Individual participation



(a) (b)

Fig. 4 The Winning and Runner-up robots (a) Noukou G's (b) Know Sea's

3. Verification of Learning Effects

We conducted a questionnaire about changes in understanding before and after the competition, and verified the educational effect. Answers to five questions about robots, circuits, and SDGs were aggregated and summarized in a graph. The contents of the responses are as follows.

A1: No understanding
A2: If they are teached they can understand
A3: they can understand on their own
A4: they can understand and teach to others

Figure 5 shows some of the responses to the question about understanding microcomputer programming. Response A1 decreased significantly, and responses A2 and A3 increased. Figure 6 shows some of the responses to the question about understanding of circuit making. A1 decreased significantly, and A3 increased significantly. Figure 7 shows some of the responses to the question about understanding of robot mechanism making. Overall, there is no big change, but answer A1 is the least among questions about understanding of microcomputer programming, circuit making and robot mechanism making. Figure 8 and Figure 9 show some of the responses to the question about understanding of SDGs and ocean debris. Among all the questions, the response A1 is the least. About circuit production and microcomputer programming, the response A1 decreased significantly before and after the convention, which is considered to have had a great educational effect. It is considered that the pre-learning we conducted online for each team two months before the competition led to their understanding of circuit construction and microcomputer programming.

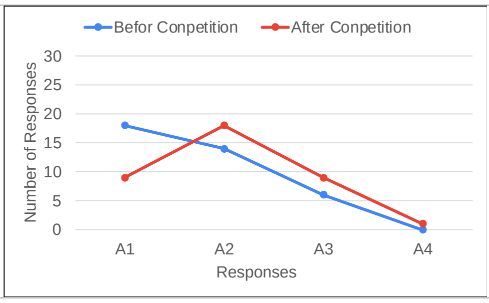


Fig. 5 Changes in understanding of microcomputer programming before and after competitions

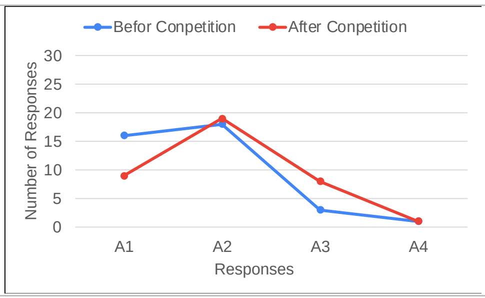


Fig. 6 Changes in understanding of circuit making before and after competitions

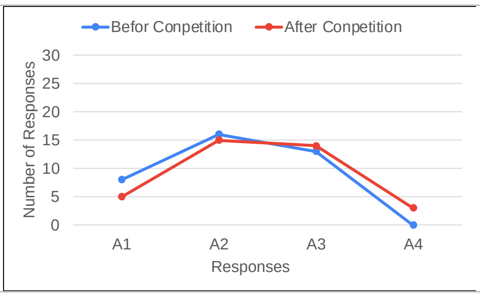


Fig. 7 Changes in understanding of robot mechanism making before and after competitions

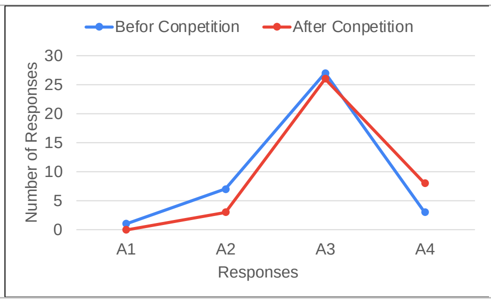


Fig. 8 Changes in understanding of SDGs before and after competitions

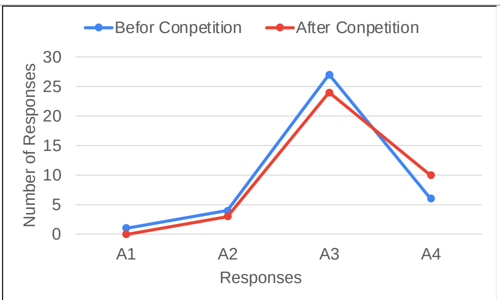


Fig. 9 Changes in understanding of ocean debris problem before and after competitions

4. Conclusions

The 8th Underwater Robot Competition Junior League held a robot competition on the theme of ocean debris problems. The competition was participated 18 teams and realized the active exchange of ideas among teams through poster presentations and the proposal of debris collection mechanism through robot competition.

From the point of view of robot education, the educational effect was confirmed mainly on microcomputer programming and circuit making.

Acknowledgment

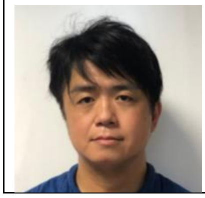
This work was supported by Yamaguti Prefecture, Iwakuni City, the Japan Society of Naval Architects and Ocean Engineers, IEEE/Oceanic Engineering Society (OES) Japan Society, Marine Technology Society Japan Society and Techno-Ocean Network.

References

1. Ministry of Foreign Affairs of Japan website.
2. Underwater robot festival in iwakuni Junior League website.

Authors Introduction

Dr. Takayuki Matsuo



He is an Associate Professor at Department Creative Engineering, National College of Technology (KOSEN), Kitakyushu College, Japan. His research area is underwater robots and biomimetic robots.

Dr. Masanori Sato

He is a Professor at the Faculty of Applied Information Technology in Nagasaki Institute of Applied Science, Japan. His main interests are agricultural robots, underwater robots, and life support robots.

Dr. Masayoshi Ozawa

He is an Associate Professor of Department of Mechanical Engineering at Kobe City College of Technology in Japan. He received his Doctor's degree in 2019 from Tokyo University of Marine Science and Technology. His research interest is Underwater Robotics and its operability.

Dr. Seiichiro Miura

He is an Associate Professor of National Institute of Technology, Tokuyama College in Japan. He graduated from the Dept. of Physics, Okayama University in 2002. He received his Dr. Sci. degree in Physics from Okayama University in 2002. His research interest is assistive technology and pattern formation.

Dr. Masakazu Arima

He is a Professor of Osaka Metropolitan University in Japan. He graduated from the Dept. of Naval Architecture, Osaka Prefecture University in 1989. He received his D. Eng. degree in Naval Architecture from Osaka Prefecture University in 1994. His research interest is underwater robotics and human factors.

Prof. Kazuo Ishii

He is a Professor at Graduate School of Life Science and System Engineering, Kyushu Institute of Technology, Japan. His research area is about field robots and intelligent robot systems

Image-based navigation of Small-size Autonomous Underwater Vehicle “Kyubic” in International Underwater Robot Competition

Yusuke Mizoguchi, Daiki Hamada, Riku Fukuda, Irmiya R. Inniyaka, Kaito Kuwata, Keisuke Nishimuta, Akihiro Sugino, Rikuto Tanaka, Yoshiki Tanaka, Yuya Nishida and Kazuo Ishii
Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka 808-0196, Japan
E-mail: yusuke.mizoguchi323@mail.kyutech.jp , ishii@brain.kyutech.ac.jp
www.brain.kyutech.ac.jp/~underwater-robotics/

Abstract

An International underwater robot competition, “RoboSub” is held in USA to demonstrate robot’s autonomy by completing underwater tasks, with a new theme each year. Student project team, “Kyutech Underwater robotics” developed autonomous underwater vehicle (AUV) “KYUBIC” with built-in image processing module for the RoboSub. Sensors on the AUV are connected via ethernet and its automatic navigation program is built using ROS. The AUV moves through gate to the target objects based on self-localization using doppler velocity log and IMU. When approaching the object, the AUV aligns its direction to the object based on the red line detection by image processing as well as the type, position, and altitude of the target objects are detected by deep learning on the image processing module.

Keywords: Autonomous Underwater Vehicle, Robot competition, Image processing, Deep neural network.

1. Introduction

Kyutech Underwater Robotics Team competed in an international robotics competition called RoboSub2022[1]. RoboSub tasks students to provide solutions to simplified versions of challenges facing the maritime industry today. For example, these include marine survey and mapping, pipeline recognition and tracking, and object detection and manipulation[1]. Our team has previously competed in the Okinawa Ocean Robotics Competition[2] and the Techno-Ocean Kobe Competition[3]. Therefore, we decided to compete in RoboSub2022 because it is an international robotics competition and the difficulty level of the image processing tasks is quite different from that of domestic competitions, making it an attractive opportunity to learn new skills. Our team was also the only team from Japan to participate. RoboSub2022 is organized by RoboNation, an organization that organizes 9 other international programs[4]. In this paper, we report on the system configuration of our AUV "KYUBIC" and our strategy and results at RoboSub2022, focusing on image-based navigation.

2. The Hovering Type AUV “KYUBIC” and RoboSub2022

2.1. *Specifications of KYUBIC*

"KYUBIC" as shown in Figure 1 is a hovering AUV

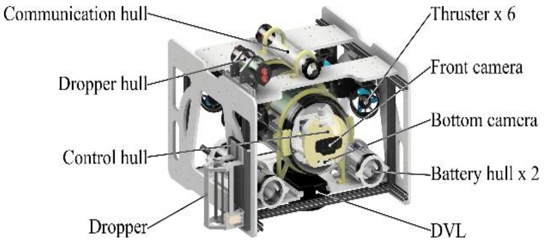


Fig.1. AUV” KYUBIC”

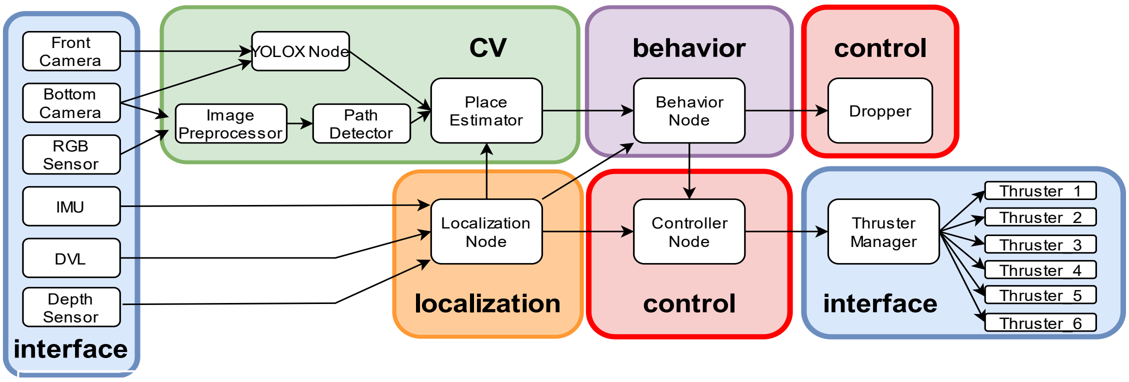


Fig.2. designed KYUBIC's node diagram

Table 1. Specification of AUV KYUBIC

Size (L, W, H)	570, 550, 400 [mm]
Weight (in air)	30 [kg]
Max. depth	15 [m]
Thrusters	Blue Robotics T200 x6
Computer1	ASRock 4X4 BOX-4800U
Computer2	NVIDIA Jetson Nano(4GB)
DVL	Teledyne Marine Pathfinder
IMU	CSM-MG200
Camera	Logitech c920n x2

platform designed for educational use. KYUBIC's hardware is modular in structure, allowing easy replacement of units (payloads) to suit experiments and robotic competitions. The hardware framework consists of some engineering plastic outer panel and an aluminum frame, providing both durability and portability. It is also equipped with various sensors such as DVL and IMU, six thrusters, and two computers, as shown in Table 1.

2.2. *KYUBIC's software configuration*

As described in the previous section, "KYUBIC" is equipped with various sensors and thrusters. While the AUV is in operation, the CPU continuously acquires information from various sensors, analyzes the AUV status, calculates the target thrust value for each thruster, and sends the target thrust value to the thrusters. Since this system is complex, our team has tried to improve the ease of development and debugging by dividing the system processing into packages for each function. As a result, the system is divided into five packages for interface, localization, CV (Computer Vision), control, and behavior. The relationship between the packages and nodes in them are shown in Figure 2. The behavior package contains the Behavior Node, which analyzes the situation and sends instructions to the Control Node on the next action to be taken, and the interface package contains the sensor drivers such as DVL, depth sensor, and IMU, and the communication interface to the microcontroller that controls the robot's actuators including the thrusters. The Localization package contains the Localization Node that calculates the self-position using dead reckoning from the DVL, IMU, and depth sensor values. The Control package includes a Control Node that calculates the target thrust by P-PID control according to the direction and distance the AUV should move based on the difference between its self-position and target position. This enables waypoint tracking to be performed. The CV package also includes PathDetector, YOLOXNode, PlaceEstimator, and ImagePreprocessor. PathDetector is a program that recognizes paths (orange boards) placed at the bottom of the pool and implemented in a rule-based algorithm. The YOLOXNode is a node that performs image recognition using YOLOX[5], one of the models of DNN (Deep Neural Network). The PlaceEstimator is a node that calculates the position of the target object in the world coordinate system by combining the information of object recognition nodes from the YOLOXNode and PathDetector with the current position of the AUV. Since the size of all objects appearing in this competition are known, the distance to the object can be calculated by the camera equation and the area of the detected object, and that the tasks required in this competition could be accomplished only by a position-based control program. The PathDetector, which is considered to have insufficient object detection performance for bluish images underwater, as such an ImagePreprocessor node is necessary to apply correction processing to the image

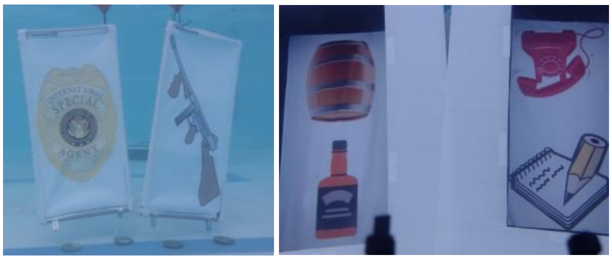


Fig.3. Mock objects photographed underwater

using an RGBSensor that can acquire the RGB components of ambient light.

3. Image recognition methods

Previously, rule-based image processing had mainly been used in our student projects, and we had been able to use MATLAB to find single-colored objects with straight lines and spheres[3]. However, to successfully complete the buoy touch and marker drop missions in this competition, our team needed to recognize complex images, which would be difficult to achieve using rule-based algorithms. Therefore, we decided to implement image recognition using DNN as a strategy to accomplish the missions.

Although more and more newer methods of DNN based image recognition are being proposed, our team selected the YOLOX model. YOLOX is chosen because it is licensed under the Apache-2.0 license, there are publicly available examples of its implementation using the "TensorRT SDK" (described below), and other users have also released implementation examples of ROS2 nodes. Therefore, our team decided to use the YOLOX model as it was suitable for integration into the AUV.

3.1. *Custom dataset generation and training*

When using DNNs to recognize their own objects, it is necessary to annotate and create their own dataset. To prepare the data set, our team printed the patterns published by the competition organizer on waterproof fabric and created mock objects by combining PVC pipes and metallic weights. Since the competition is held in an indoor pool underwater, we thought it would be appropriate to train the weights of the model based on images taken of the objects placed in the indoor pool with the same size as in the actual competition to make the conditions as close as possible. However, since we do not have an indoor pool easily accessible to us, we used an outdoor pool. The depth of the outdoor pool was shorter than the vertical length of the buoys for the actual competition. As a result, the mock objects created were reduced to approximately 70% the size of the actual dimensions for the competition. These were placed in an outdoor pool on campus and underwater video was captured using the AUV. The video was divided into frame-by-frame images, and the images were annotated to create a unique data set. The image of the mock objects taken underwater are shown in Figure 3. The annotation targets were 9 labels including "gate", "gun", "badge", "bootlegger", "g-man", "bottle", "barrel", "phone", and "paper" which included patterns that were not used in the buoy touch and marker drop missions. For training, the YOLOX-s weights pre-trained on the COCOtrain2017 dataset[6] were used as initial weights and fine-tuned with a batch size of 16 and a learning rate of 1/64,300 epochs. The learned weights were evaluated using AP (Average precision), AR (Average Recall), and IoU (Intersection Over Union). The AP was 80.7% and the AR was 83.6%. The number of training and validation images for each label, AP and AR are shown in Table 2. IoU was calculated between 0.50~0.95 for both AP and AR.

Table.2 The number of images of teacher data per label and AP and AR

Label	Train data	Validation data	AP [%]	AR [%]
gate	101	35	89.2	90.5
gun	165	49	83.2	86.3
badge	187	44	84.5	86.3
bootlegger	69	21	85.1	86.1
g-man	50	15	79.6	82.6
bottle	55	9	79.0	82.2
barrel	68	14	70.6	77.1
phone	58	14	75.2	80.7
paper	57	14	80.0	81.4
average	90	23.8	80.7	83.6

3.2. *Inference*

In determining the size of the network model, a policy was first determined. The policy was to select a model that could be inferred in about 100 [ms] and that was as accurate as possible. The inference time was determined since the information acquisition cycle of the AUV's onboard DVL is 3.5 [Hz] and that the AUV's motion has a large time constant. As a result, a network model of size YOLOX-s was selected within the YOLOX model. As mentioned above, the computational resource for inference built into "KYUBIC" is the "NVIDIA Jetson Nano (4GB)," which is inexpensive, small, and easily accessible, but not powerful. When inference was performed using Pytorch weight files and python, the median inference time on the Jetson Nano was 10.3 [sec] out of 10 experiments. This was far from satisfying our requirements. However, after optimization with the TensorRT SDK provided by Nvidia, the median inference time on the ROS node implemented in C++ was 0.110 [sec] out of 10 experiments, which was a significant reduction. This enabled us to infer YOLOX-s, calculate the center position of the buoy, and judge whether the relationship between the buoy and the AUV is stable, all in the target period of approximately 0.1 [sec].

4. Results

4.1. *Results of image recognition*

To evaluate the weights trained on the home-made dataset, a test dataset was created by annotating videos taken in the competition environment in the same way as when the home-made dataset was created. We executed inference on that dataset and calculated the Average Precision (AP) and Average Recall (AR). The results along with AP and AR on the dataset at training are shown in Table 3. This time, only the "badge" and "gun" labels were calculated due to the videos that could be taken in the competition environment. The number of test data images for each label was 170 and 106. Both AR and AP decreased in the test data set, suggesting that overlearning occurred in the training data set. Quantitative evaluation on the test data confirmed that recognition accuracy decreased. However, inference in the competition environment confirmed that when an object is in the image, it is always recognized as one object in succession without causing False Positive or False Negative, which recognizes it as a separate object.

Table.3 Comparison of AP and AR
on training dataset and competition dataset

Label	AP(@train) [%]	AR(@train) [%]	AP(@test) [%]	AR(@test) [%]
gun	83.2	86.3	73.5	78.5
badge	84.5	86.3	74.4	78.9
Average	83.8	86.3	74.4	78.7

4.2. *Result of Competition*

Our team was not able to fully execute our strategy in this competition.

There were three main reasons for this. First, the Place Estimator could not be implemented. Second, the Image Preprocessor could not be implemented, so the Path Detector could not fully demonstrate its performance. Finally, the integration test between the "YOLOX Node" and the "Control Node" was not sufficiently completed. However, during the competition, the strategy was reconfigured and reimplemented, focusing on the position controller, which yielded high performance for the applied missions. As a result, our team successfully completed the missions our team concentrated on (gate pass and surfacing missions) and were ranked 8th out of 19 teams in the qualifying round.

Conclusion

In this paper, we have introduced the international robotics competition RoboSub2022, as organized by RoboNation and AUV "KYUBIC". We have also described the strategy of the competition, our implementation, and the image recognition module using DNN that was necessary to accomplish the strategy. Regarding the image recognition module, it was able to recognize pictures, but was not able to score points. This was because the integration test between the image recognition module and the position controller had not been sufficiently performed. However, we confirmed that the image recognition module's performance was adequate. In the future, our team would like to expand the integration test, develop a target tracking program based on the image recognition results, and develop a program to calculate three-dimensional positions using an RGBD camera instead of an RGB camera.

References

1. “ RoboSub 2022 official website ” ,2022.(Accessed Dec.19,2022)
2. Irmiya R.I,et al.,"Underwater Acoustic Positioning Based on MEMS Microphone for a Lightweight Autonomous Underwater Vehicle "Kyubic"", Proceedings of

International Conference on Artificial Life & Robotics (ICAROB2022),pp. 349-353, 2022.

3. K.Harada, et al.," Autonomous Underwater Vehicle with Vision-based Navigation System for Underwater Robot Competition", Proceedings of International Conference on Artificial Life & Robotics (ICAROB2022),pp. 354-359, 2022.
4. "RoboNation official website",2022.(Accessed Dec.19,2022)
5. Ge, et al., "YOLOX: Exceeding YOLO Series in 2021", arXiv preprint arXiv:2107.08430,2021.
6. Lin, TY. et al.," Microsoft COCO: Common Objects in Context", European Conference on Computer Vision, pp.740-755,2014.

Authors Introduction

Mr. Yusuke Mizoguchi

He received his bachelor's degree from the Department of Electronics, Kyoto Institute of Technology, Japan in 2021. He is currently a Master's course student at the Department of Human Intelligence Systems, Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology, Japan.

Mr. Daiki Hamada

He received his master's degree from Department of Human Intelligence Systems, Kyushu Institute of Technology, Japan, in 2022. He is pursuing the PhD at Kyushu Institute of Technology, Department of Life and Systems Engineering, under the supervision of Prof. Kazuo Ishii.

Mr. Riku Fukuda

He received his bachelor's degree from the Department of Engineering, Electronic Information Engineering, Nishinippon Institute of Technology, Japan in 2021. He is currently a Master's course student at the Department of Human Intelligence Systems, Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology, Japan.

Mr. Irmiya R. Inniyaka

He received his master's degree in Automotive Mechatronics from the Department of Mechatronics, Cranfield University, England in 2016. He is currently pursuing his Doctoral program at Kyushu Institute of Technology, Japan. His research interests are in Control, Underwater Robotics navigation and localization.

Mr. Kaito Kuwata

He completed the production and electrical system technology course at Kyushu Polytechnic College in 2022. Currently enrolled in a master's course in Human Intelligence Systems, Graduate School of Life Science and Technology, Kyushu Institute of Technology, Japan.

Mr. Keisuke Nishimuta

He received his bachelor's degree from the Department of Mechanical and Control Engineering, the Kyushu Institute of Technology School of Engineering, Japan in 2022. He is currently a Master's course student in Kyushu Institute of Technology, Japan.

Mr. Akihiro Sugino

He received his bachelor's degree from the Faculty of Science and Engineering, Kyushu Sangyo University, Japan in 2021. He is currently a Master's course student at the Department of Human Intelligence Systems, Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology, Japan.

Mr. Rikuto Tanaka

He completed the production and electrical system technology course at Shikoku Polytechnic College in 2021. Currently enrolled in a master's course in Human Intelligence Systems, Graduate School of Life Science and Technology, Kyushu Institute of Technology, Japan.

Mr. Yoshiki Tanaka

He received his master's degree from Department of Human Intelligence Systems, Kyushu Institute of Technology, Japan, in 2019. He is pursuing the PhD at Kyushu Institute of Technology, Department of Life and Systems Engineering, under the supervision of Prof. Kazuo Ishii. His research area is underwater robots, its application.

Dr. Yuya Nishida

He is currently an Associate Professor at the Department of Human Intelligence Systems, Kyushu Institute of Technology, Japan. He obtained his master's degree in engineering in 2008 and his D. Eng. degree in 2011 at the same university. His research interests are in the field of Underwater Robotics, Field Robotics, and Intelligent Systems.

Dr. Kazuo Ishii

He is currently a Professor at the Department of Human Intelligence Systems of Kyushu Institute of Technology, Japan. He obtained his M.S.degree in 1993 and his D. Eng. degree in 1996 at The University of Tokyo. His research interests are in the fields of Underwater Robotics, Field Robotics, Neural Networks and Intelligent Systems.

Development of Harvesting Robot for Tomato Robot Competition 2022 and Its Evaluation

Takeru Ohshige
Dept. of Intelligent Mechanical Engineering, Hiroshima Institute of Technology,
2-1-1 Miyake, Saeki, Hiroshima,731-5193, Japan
Moeko Tominaga
Dept. of Integrated Systems Engineering, Nishinippon Institute of Technology,
2-1-1 Miyake, Saeki, Hiroshima,731-5193, Japan
Takuya Fujinaga
Dept. of Electronics Engineering and Computer Science, Fukuoka University,
8-19-1 Nanakuma, Jonan-ku, Fukuoka 814-0180, Japan
Yasunori Takemura
Dept. of Integrated Systems Engineering, Nishinippon Institute of Technology,
2-1-1 Miyake, Saeki, Hiroshima,731-5193, Japan
Jonghyun Ahn
Dept. of Intelligent Mechanical Engineering, Hiroshima Institute of Technology,
2-1-1 Miyake, Saeki, Hiroshima,731-5193, Japan,
E-mail: ad19018@cc.it-hiroshima.ac.jp, tominaga@nishitech.ac.jp, t.fujinaga.wz@fukuoka-u.ac.jp,
takemura@nishitech.ac.jp, j.ahn.h2@it-hiroshima.ac.jp

Abstract

Agriculture is one of the most important industries for human food production. Recently, the number of farmers in Japan is decreasing, and the age of farmers is increasing. Therefore, the automation of agriculture using robot systems is highly required. The “Tomato Robot Competition” is held every year in Fukuoka, Japan, to arouse students’ interest in the field of agriculture, and to promote the development of agricultural automation technology by robot systems. In this paper, we introduce the developed tomato harvesting robot to participate in the “Tomato Robot Competition 2022”, and its harvesting performance evaluation. The developed tomato harvesting robot consists of 3 linear arms, an end-effector, a rail movement system, and an electronic system of communication and control. In the evaluation, the developed tomato harvesting robot took approximately 1.1 [min] to harvest a single tomato.

Keywords: Plant factories, Harvesting robot, Tomato robot competition 2022

1. Introduction

Agriculture is one of the most important industries for human food production. Recently, the number of farmers in Japan is decreasing, and the age of farmers is increasing [1]. Therefore, the automation of agriculture using robot systems is highly required.

Tomatoes are the most widely produced agricultural crop in the world [2], and one of the vegetables designated by Japan's Vegetable Production and Shipping Stabilization Law [3]. In countries developing agriculture well such as the Netherlands, "plant factories" that control the cultivation environment are also popular [4]. However, a series of operations such as plucking, leaf cutting, and harvesting must be performed in an environment where the facility is kept hot and humid, placing a heavy burden on the workers. Against this background, the “Tomato Robot Competition 2022”, a competition aimed at technological innovation related to robots living in harmony with nature and the introduction of robots into agricultural work, was held from November to December 2022 [5]. In this paper, we introduce the tomato harvesting robot that we developed to participate in the “Tomato Robot Competition 2022”, and its harvesting performance evaluation.

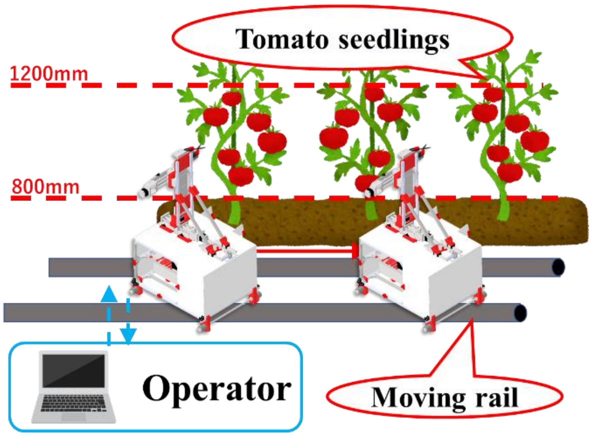

Tomato seedlings
1200mm
800mm
Operator
Moving rail

Fig. 1 Overview of tomato harvest work using a robot

2. Development of Harvesting robot

2.1. *Harvest strategy of this robot*

As shown in Figure 1, the robot is installed on top of a pipe in the plant factory where it operates. The robot is moved to the top of the pipe and moved to the front of the tomatoes to be harvested. By repeating this operation, harvesting tomatoes growing at a height of 800 [mm] ~ 1200 [mm] from the ground is completed. Only tomatoes in the harvest stage of tomato fruit are harvested based on their color components.

2.2. *Design details of harvesting robot*

The tomato harvesting robot must satisfy the following four elements to be developed in this study.

(1) Operability
(2) Scalability
(3) Manageability
(4) Safety

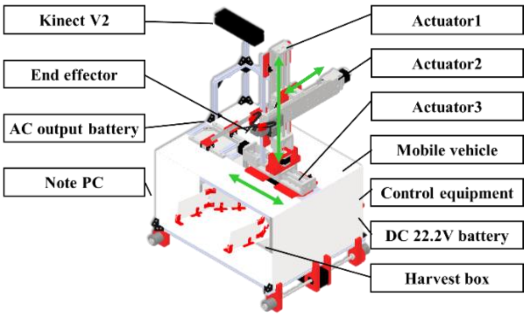

Kinect V2
End effector
AC output battery
Note PC
Actuator1
Actuator2
Actuator3
Mobile vehicle
Control equipment
DC 22.2V battery
Harvest box

Fig. 2 Overview of developed harvesting robot

Regarding (1), one of the concepts of this machine is that it is easy to operate and can be operated by anyone. Therefore, As shown in Figure 2, an orthogonal arm was selected to allow the controller's sensory manipulation of the robot's arm.

Regarding (2), this is to ensure expandability when making improvements to the machine. The orthogonal type of arm was adopted from the viewpoint of scalability for automation. In addition, the robot's frame was also made of a 20 [mm] x 20 [mm] aluminum frame for expandability.

Regarding (3), This means that the robot can be easily maintained in the event of trouble. The control equipment, batteries, and harvesting box are all in one location for easy management. In addition, the robot is equipped with a cover that covers the entire outside for dust and drip-proofing, but the cover around the area where the control equipment is located can be opened and closed.

Regarding (4), The robot can be stopped by the user in case of an irregularity. The unit is equipped with an emergency stop button.

2.3. *Electrical and Controls Design*

The system configuration of the robot is shown in Figure 3. Two power sources were used a DC lithium polymer battery and an AC output battery. The AC output battery supplies power to the Kinect V2, the Wi-Fi router, and the Note PC, to which the Real sense is connected.

The lithium polymer battery is initially connected to the emergency stop. Next, supplies power to each device requiring a 22.2 [V] supply voltage. Also, the servo motor requires a 7.5 [V] power supply and the Arduino requires a 5 [V] power supply, so we use a DC/DC converter to step down the voltage.

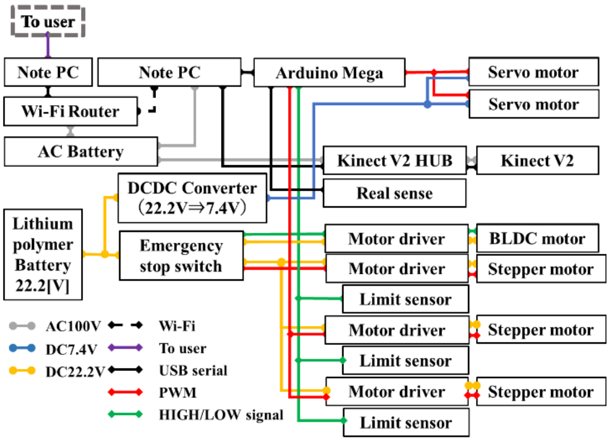

To user
Note PC
Note PC
Arduino Mega
Servo motor
Servo motor
Wi-Fi Router
AC Battery
Kinect V2 HUB
Kinect V2
DCDC Converter (22.2V⇒7.4V)
Real sense
Lithium polymer Battery 22.2[V]
Emergency stop switch
Motor driver
BLDC motor
Motor driver
Stepper motor
Limit sensor
Motor driver
Stepper motor
Limit sensor
Motor driver
Stepper motor
Limit sensor
AC100V
DC7.4V
DC22.2V
Wi-Fi
To user
USB serial
PWM
HIGH/LOW signal

Fig. 3 Electrical and system configuration diagram

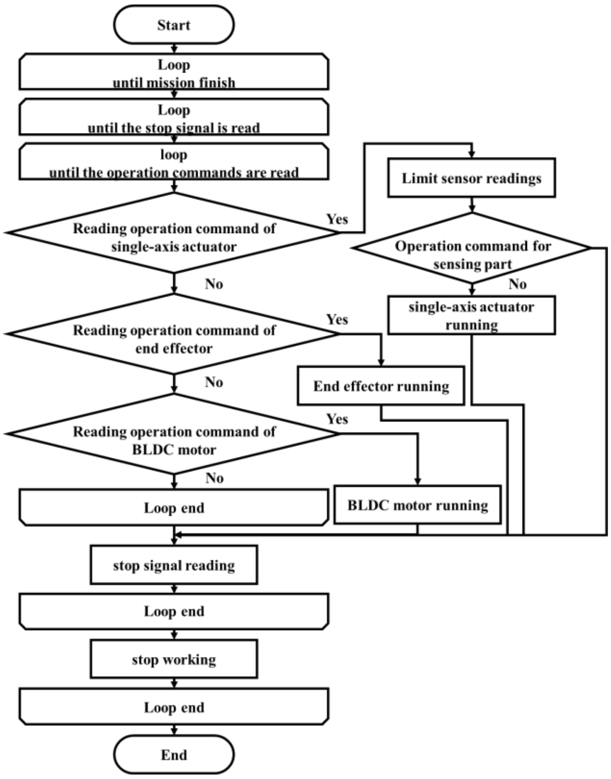

Start
Loop until mission finish
Loop until the stop signal is read
loop until the operation commands are read
Reading operation command of single-axis actuator
Yes
No
Limit sensor readings
Operation command for sensing part
No
single-axis actuator running
Reading operation command of end effector
Yes
No
End effector running
Reading operation command of BLDC motor
Yes
No
BLDC motor running
Loop end
stop signal reading
Loop end
stop working
Loop end
End

Fig. 4 Flowchart for tomato harvesting robot

Next, the system architecture is described. This robot operates by controlling motors and sensors through a remote desktop connection by Wi-Fi to a laptop PC equipped with the tomato robot. The moving cart is operated by a BLDC motor. The Cartesian arm is equipped with a motor driver, stepping motor, and limit sensors. The orthogonal arms are equipped with three single-axis actuators. The end effectors are operated by servo motors. These represent the gripping mechanism of the scissors and the rotation mechanism of the roll angle. All these motors are controlled by Arduino, a microcomputer that can handle PWM signals and HIGH/LOW signals.

2.3. *The control program of harvesting robot*

Figure 4 shows the operation program of the machine. First, the motion signals for each of the single-axis actuators, end effectors, and BLDC motors are read. For the three single-axis actuators, after reading the motion commands, the machine reads whether the limit sensor of the single-axis actuator has detected a slider. If not detected, the actuator moves according to the motion signal, and if the motion signal is a 2-servo motor or BLDC motor, the actuator simply moves according to the motion command. After the motion, the motion stops when the stop signal is read. This sequence of motion continues to operate as a loop until the robot's power is turned off.

3. Evaluation experiment of developed Robot

We experiment to evaluate the performance of the developed robot. This evaluation was done in the same condition as the competition. In addition, ten people were harvested for accurate evaluation. Operators 1~5 are not experienced operators and operators 6~10 are experienced operators. Figure 5 shows the results by watching the actual operation, and Figure 6 shows the results by watching the sensor video.
When watching the actual operation, the success rate is 96.7% and the harvesting speed is 0.9 [min/piece].
When watching the sensor video, the success rate is 80.4% and the harvesting speed is 1.1 [min/piece].

4. Conclusion

In this paper, we describe the design and development of our robot and the evaluation of its harvesting performance. The results of the experiment show that the

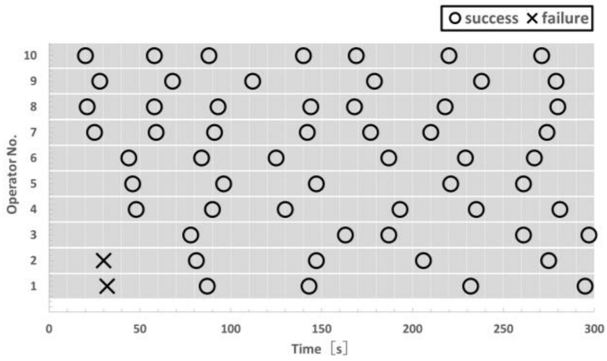


Fig. 5 Harvest results by watching the actual operation

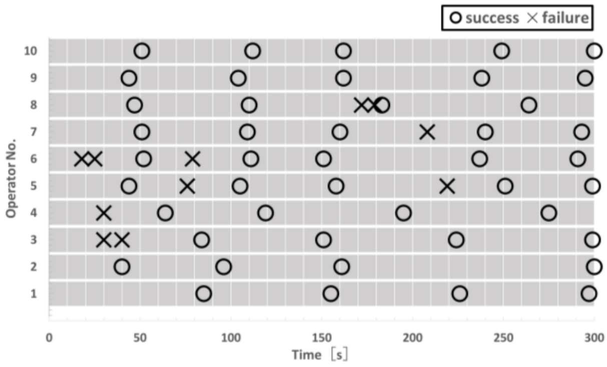


Fig. 6 Harvest results by watching the sensor video

success rate and harvesting speed decreased while watching the sensor images. Therefore, we should that the sensor mounting and the sensing system itself need to be improved.

In addition, the experienced operators had a slightly higher harvesting speed than the inexperienced operators. Although they are remotely operated now, autonomous robots are desirable for competition.

In the future, it is necessary to increase the operating speed and consider sensing for automation.

References

1. Census of Agriculture and Forestry, Ministry of Agriculture, Forestry, and Fisheries of Japan, (June 30, 2022).
2. Food and Agriculture Organization of the United Nations
3. Japan's Vegetable Production and Shipping Stabilization Law
4. Drago Cvijanović, Katarina Stojanović, Suzana Borović, “Impact of Digital Technology and Smart Systems on Mobility and Agriculture in Serbia”, International Conference on Logistics in Agriculture 2022
5. Tomato Robot Competition 2022

Authors Introduction

Mr. Oshige Takeru

He is acquiring the B.S. Department of Intelligent Mechanical Engineering, in Hiroshima Institute of Technology. He is interested in tomato robots and the Improvement of technology related to robots.

Dr. Moeko Tominaga

She received a Ph.D. degree from Kyushu Institute of Technology, Japan in 2021 and is now working as an Assistant Professor at Nishinippon Institute of Technology, in Japan. Her research area is about robotics and machine learning. In particular, she is interested in inter-agent cooperation in multi-agent systems.

Dr. Takuya Fujinaga

He received Ph.D. degree from Kyushu Institute of Technology, Japan in 2021 and now he is working as an Assistant Professor at Fukuoka University, in Japan. His research area is about field robotics.

Dr. Yasunori Takemura

He received Ph.D. degree from Kyushu Institute of Technology, Japan in 2010 and now he is working as an Associate Professor at Nishinippon Institute of Technology, in Japan. His research area is about machine learning, data mining and robotics.

Dr. Jonghyun Ahn

He received the Ph.D. degree in underwater robotics from the Kyushu Institute of Technology, Kitakyushu, Japan in 2017 and now he is working as an Assistant Professor at Hiroshima Institute of Technology. His research interests include underwater imaging system, intelligent sensing, and underwater communication system.

Development of a Tomato Harvesting Robot for Farm Field

Shunsuke Oda, Ryuma Fukumoto, Kensuke Hirata, Shu Tahara, Keisuke Yoshida,
Shinsuke Yasukawa, Kazuo Ishii
Dept. of Human Intelligence Systems, Kyushu institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka,808-0196, Japan
E-mail: oda.shunsuke299@mail.kyutech.jp

Abstract

The 9th Tomato Harvesting Robot Competition was held at the green house in Kitakyushu Science and Research Park. The competition consisted of two leagues for different field areas: rail-style for greenhouses and free-style for open fields. Our team competed in the free-style league. Participation in the free-style category required the following: tomato harvesting mechanism; mobility mechanism to move on the rough terrain; camera to photograph tomato; and self-location system. We developed a 3-axis cartesian coordinates manipulator tomato harvesting robot using crawler to move on soil and using suction and cutting to harvest the tomatoes. In this paper, we described the system architecture of tomato harvesting robot and the results of 9th Tomato Harvesting Robot Competition in 2022.

Keywords: *Tomato-Harvesting Robot, Crawler, move on soil*

1. Introduction

The current rapid aging of Japan's population and declining birthrate has raised concerns about labor shortages in the near future. In the agricultural sector, in particular, the number of key agricultural workers dropped to 1,363,000 by 2020, and the percentage of workers aged 65 or older was as high as 69.6%, making labor shortage a serious problem. In addition, the stable production and quality assurance of food in the near future are threatened, and an immediate solution to the problem and improvement of the situation are desired [1]. We aim to address these problems through the realization of smart agriculture, increasing efficiency while minimizing manual labor through productivity improvement and automation of production lines. Part of this effort is the development of automatic tomato harvesting machines for automating labor-intensive and manpower-intensive harvesting and transportation tasks in the agricultural field. Most tomatoes are either produced in greenhouses or in open fields, and the introduction of robots is desired because of the high temperature, high humidity, and long work hours required in harvesting them. Tomatoes are not encased by hard shells, thus requiring the use of end-effectors and harvesting techniques that can delicately handle them without damaging their surface or interior. In this context, with the aim of advancing automatic harvesting by developing and sharing basic technologies for harvesting robots, a competition for automatic tomato harvesting robots has been organized by the Center for Socio-Robotic Synthesis, Kyushu Institute of Technology since 2014 [2]. The Tomato Robot Competition evaluates the accuracy and speed of the harvesting robot's harvesting motion. This paper describes the design specifications of the “Tomato Vacumer,” a tomato harvesting robot developed for the competition, and the control technology used to ensure accurate harvesting.

2. Tomato harvesting robot

2.1. *Robot behavior and configuration*

Fig.1 shows the appearance of the tomato harvesting robot. The robot is equipped with a 3-axis Cartesian manipulator, a harvesting mechanism, a moving mechanism, and a vision sensor. Modular structure of the electrical circuits facilitates maintenance. The development of a tomato harvesting robot is described in [3]. The system configuration, image processing, and moving mechanism have been modified for the competition.

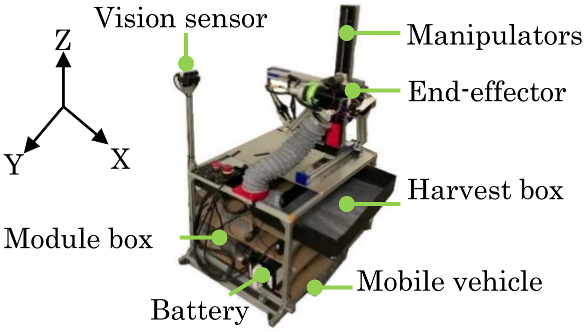

Vision sensor
Z
Y
X
Manipulators
End-effector
Harvest box
Module box
Mobile vehicle
Battery

Fig1. Tomato harvesting robot

2.2. *System configuration*

The tomato harvesting robot is controlled using Intel's NUC as the main PC and NVIDIA's Jetson AGX Xavier as the sub-PC. The NUC runs MATLAB and Simulink by MathWorks, while the Jetson runs Python and OpenCV. The main PC and the sub-PC are connected via the ROS network and send and receive data in both directions. Fig.2 shows the system configuration. The main PC mainly controls each manipulator, end-effector, and moving mechanism using Simulink models. In Simulink1, command values are sent to other Simulink models according to the harvesting strategy using Stateflow, a tool for designing state transition control schemes. Simulink 2, 3, and 4 receive command values from Simulink 1 and transfer them to the connected Arduino microcontroller via serial communication. The Arduino then processes and controls the connected manipulators, end-effectors, and moving mechanisms. Also, rosbag is used for logging. Two ROS Nodes are running on the sub-PC. Node1 publishes RealSenseD435 RGB and depth images to the ROS network at regular intervals. Node2 starts processing when it receives a command value to execute image processing from Simulink1 and publishes the result.

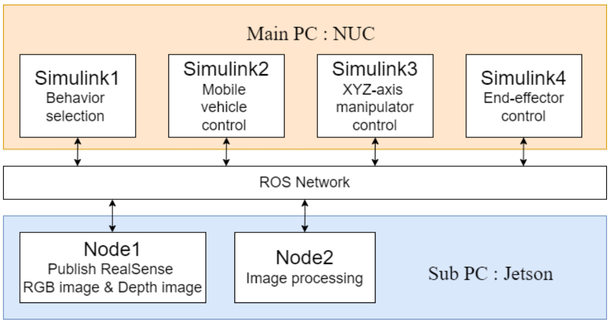

Main PC : NUC
Simulink1
Behavior selection
Simulink2
Mobile vehicle control
Simulink3
XYZ-axis manipulator control
Simulink4
End-effector control
ROS Network
Node1
Publish RealSense RGB image & Depth image
Node2
Image processing
Sub PC : Jetson

Fig2. System Configuration Diagram

2.3. *Mobile vehicle*

CuGo by CuGoRex was selected as the mobile mechanism for movement on soil. Crawlers were selected as the moving mechanism to enable back-and-forth movement and right-and-left turning in the soil. An encoder is installed to measure the move distance by acquiring the rotation speed and to control the rotation speed by PID control. Fig.3 shows the appearance of the crawler used in the tomato harvesting robot.

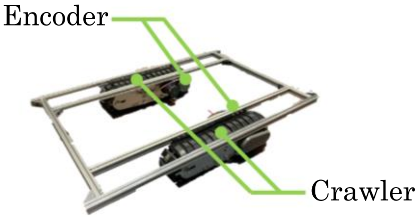

Encoder
Crawler

Fig3. Mobile vehicle

2.4. *Orientation control*

An IMU (Adafruit BNO055) is installed to maintain the straightness of the tomato harvesting robot. On-off control is performed based on the difference between the yaw angle output by the IMU and the yaw angle at the time the move command is received. If the difference is greater than the threshold, the speed of one crawler is increased to correct the heading.

2.5. *Faster image processing*

Conventional image processing using only Simulink models was deemed to be time-consuming. The reason is that MATLAB and Simulink do not support alignment of RGB and depth images, so the 3D point cloud is converted once and the depth image is restored again. Therefore, image processing is done by the sub-PC and was developed in Python. The average processing time per image using the conventional system was 318 ms. Using the modified system significantly reduced this to 7 ms.

3. Harvest strategy

Fig.4 shows a flowchart of the harvesting operation. First the tomato harvesting robot searches for tomatoes. If a tomato is detected, the robot stops moving and finds the coordinates of the center of the tomato in the image. The coordinates obtained are in the image coordinate system, so they are converted to world coordinates. The YZ-axis manipulator is then moved. After that, suction is started, and the X-axis manipulator is moved. A laser and an optical sensor are located inside the harvesting mechanism, and if the tomato is inside the harvesting mechanism, the laser is intercepted, and the tomato is judged to be within grasp. The blade at the lower jaw of the harvesting mechanism closes upward to cut the fruit stalk. The X-axis manipulator returns to the initial position before the suction stops. The lower opens and releases the tomato, which falls and rolls along the hose connected to the harvesting box to complete the harvesting process.

4. Competition rules

The operating environment for the tomato harvesting robot in the freestyle category was inside a greenhouse where tomatoes are grown. The robot was placed on a 2000×8000 mm soil area as shown in Fig.5. 10 minutes, two rounds of competition. The robot is operated from an arbitrary position within it [4].

5. Experiment in soil

To evaluate the performance of the tomato harvesting robot, an experiment was conducted in a field like the freestyle area. The robot's first step was to move to the right to detect tomatoes. Harvesting operation would then be performed after detecting tomatoes. After harvesting, it would start moving again by 4 m before reversing direction and start searching again. Fig.6 shows the motion data of the robot with respect to time (from top to bottom) – number of crawler revolutions; yaw, roll, and pitch angles; movement direction (0:stop, 1:left, 2:right); mode number within the harvest strategy (0: stop, 10 to 19: prepare for operation, 20 to 29: search for tomatoes, 30 to 39: harvest tomatoes, 40 to 49: harvest complete, 50 to 59: induce fruit, 60 to 69: harvest complete); flag for tomato detection (0: not detected, 1: detected); X-coordinate self-position; Y-coordinate self-position; and self-position in XY-coordinates. The robot started to move at around 8 seconds and detected the tomatoes after around 11 seconds before stopping to start harvesting them.

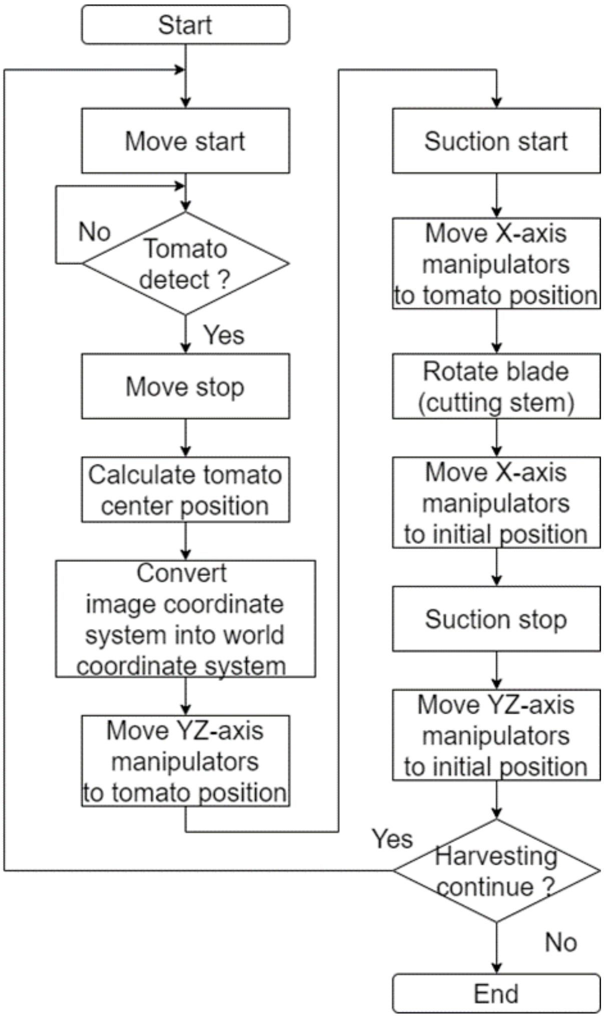

Start
Move start
No
Tomato detect ?
Yes
Move stop
Calculate tomato center position
Convert image coordinate system into world coordinate system
Move YZ-axis manipulators to tomato position
Suction start
Move X-axis manipulators to tomato position
Rotate blade (cutting stem)
Move X-axis manipulators to initial position
Suction stop
Move YZ-axis manipulators to initial position
Yes
Harvesting continue ?
No
End

Fig. 4 Flowchart of the harvesting operation

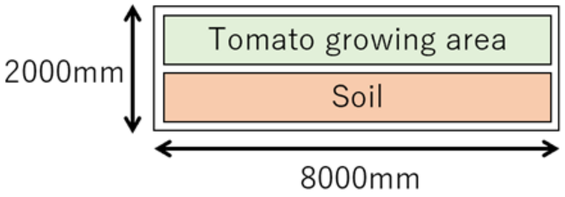
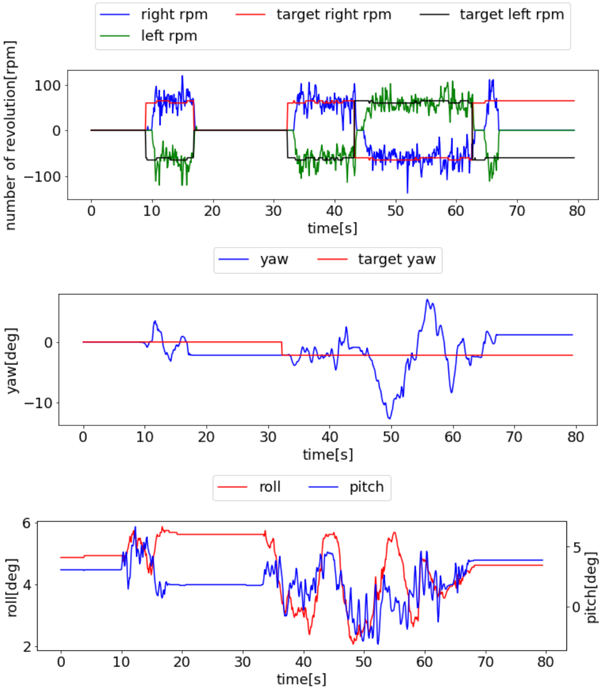

2000mm
Tomato growing area
Soil
8000mm

Fig. 5 Free-style competition area

Harvesting was finished at 30 seconds and the robot started searching again. The direction of movement was reversed because it advanced to the set distance of 4 meters after 45 seconds. The movement was reversed again after 60 seconds. Crawler control by PID control overshot due to poor gain adjustment relative to the target speed. The moving mechanism was smaller than that of the tomato robot, so the robot would sway to the left and right when it moved. However, the manipulator moved during harvesting, but the robot itself did not sway due to the movement of the manipulator. In addition, the heading was corrected based on the yaw angle during the movement, but the swing could have been so large that the difference from the target angle exceeded 10 degrees. The crawlers were found to be particularly sensitive to soil conditions, and would not run if there were stones between them, but would rotate only on one side, resulting in a misorientation. As a result, the azimuth was off, so the robot moved forward by about 0.1 m.



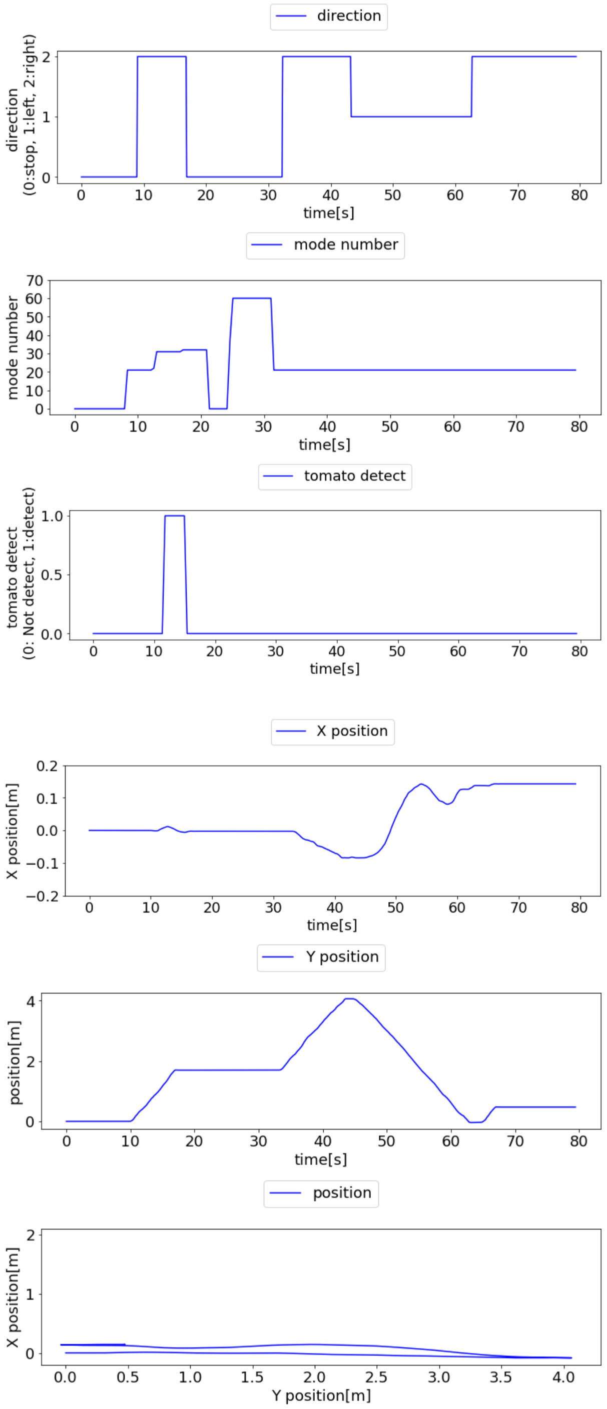


Fig 6. Tomato robot motion data

6. Conclusion

We improved the movement mechanism and the program to compete in the Tomato Robotics Competition. For image processing, speed was improved by migrating from Simulink to Python without changing the content of the process. The mobile mechanism was able to run on the soil, but it was not able to correct the orientation and moved forward. In the future, we will revise the algorithm of orientation correction and develop a system that is less affected by soil conditions.

References

1. Ministry of Agriculture, Forestry and Fisheries, “Key agricultural workers,” accessed on December 5, 2022 https://www.maff.go.jp/j/wpaper/w_maff/r3/r3_h/trend/part1/chap1/c1_1_01.htmlMini
2. “competition for automatic tomato harvesting robots official website, ”,https://www.lsse.kyutech.ac.jp/~sociorobo/tomato-robot2022/
3. Fujinaga T., Yasukawa S., Ishii K., "Development and Evaluation of a Tomato Fruit Suction Cutting Device,"2021 IEEE/SICE International Symposium on System Integration, SII 2021, pp.628 - 633, 2021.
4. 9th Tomato Robotics Competition Home Page, “Senior Section Consultation Rules ” https://www.lsse.kyutech.ac.jp/~sociorobo/wp-content/uploads/2022/10/31ffd783d8f4c55dcf758e0ce30d2d86.pdf

Authors Introduction

Mr. Shunsuke Oda

He received his B. S. Intelligent Mechanical Engineering degree at Hiroshima Institute of Technology in 2022. He is currently a Master’s course student in Kyushu Institute of Technology, Japan.

Mr. Ryuma Fukumoto

He completed the production and electronic information system technology course at Kyushu Polytechnic in 2021. He is currently enrolled in the Master's course in Human Intelligence Systems, Graduate School of Life Science and Technology, Kyushu Institute of Technology, Japan.

Mr. Keisuke Hirata

He received his bachelor’s degree from the Department of Engineering, Electronic Information Engineering, Nishinippon Institute of Technology, Japan in 2021. He is currently taking his Master’s Program in Kyushu Institute of Technology, Japan.

Mr. Shu Tahara

He received his B.S., Mechanical Engineering Course from the Department of Engineering, University of the Ryukyus, 2022. He is currently taking his Master’s Program in Kyushu Institute of Technology, Japan.

Mr. Keisuke Yoshida

He received his Bachelor's Degree from the Department of Systems Engineering, Faculty of Science and Technology, Saga University in 2022. He is currently his Master’s Program in Kyushu Institute of Technology, Japan..

Dr. Shinsuke Yasukawa

He received his Ph.D. degree in electrical, electronic, and information engineering from Osaka University, Japan, in January 2017. Currently, he is an Associate Professor in the Department of Human Intelligence Systems, Kyushu Institute of Technology, Japan. His research interests include information processing in biological systems and their applications in robotics.

Prof. Kazuo Ishii

He is currently a Professor at the Department of Human Intelligence Systems of Kyushu Institute of Technology, Japan. He obtained his M.S. degree in 1993 and his D. Eng. degree in 1996 at The University of Tokyo. His research interests are in the fields of Underwater Robotics, Field Robotics, Neural Networks and Intelligent Systems.

Report on the 8th Tomato-Harvesting Competition toward Smart Agriculture

Kazuo Ishii
Dept. of Human Intelligence Systems, Kyushu institute of Technology, 2-4 Hibikino, Wakamatsu, Kitakyushu,
Fukuoka,808-0196, Japan,
ishii@brain.kyutech.ac.jp

Takayuki Matsuo
National Institute Technology, Kitakyushu College, 2-20-1, Shii, Kokuraminami-ku, kitakyushu-shi, Fukuoka, Japan

Yasunori Takemura, Takashi Sonoda, Atsushi Sanada, Moeko Tominaga
Department of Engineering, Nishinippon Institute of Technology, 2-11, Aratsu, Kanda-town, Miyako-gun, Fukuoka 800-0396, Japan

Yuya Nishida, Shinsuke Yasukawa, Kanako Shirahashi
Dept. of Human Intelligence Systems, Kyushu institute of Technology, 2-4 Hibikino, Wakamatsu, Kitakyushu,
Fukuoka,808-0196, Japan

Takuya Fujinaga
Faculty of Engineering, Fukuoka University, 8-19-1 Nanakuma, Jonan, Fukuoka, 014-0133, Japan

Daisaku Arita
Faculty of Information Engineering, University of Nagasaki, 1-1-1 Manabino, Naganocho, Nagasaki, 851-2195, Japan

Kazushi Kawajiri
Hibikinada Green House Inc. 4 Yanasaki, Wakamatsu, Kitakyushu, Fukuoka, 808-0111, Japan

Kenich Ohshima, Masayuki Okada
Kyushu Polytechnic College, 1665-1, Shii, Kokura Minami, Kitakyushu, Fukuoka, 802-0985, Japan

Abstract

Tomato is one of the important fruit vegetables and most tomatoes are produced in the greenhouses, or large-scale farms, where the high temperature and humidity, and long harvest age force the farmer heavy works. To develop the tomato harvesting robot, many research issues exist such as manipulator design, end-effector design, collaborative behavior, artificial intelligence, motor control, image processing, target recognition and so on. With an aim to promote the automation of tomato harvesting, we have organized the tomato harvesting robot competition since 2014, and currently changed the competition field to the greenhouse in 2020. In this paper, we discuss the results of 8th tomato harvesting robot competition in 2021.

Keywords: Tomato-Harvesting-Robot Competition, End-effector, Mobile Mechanism

1. Introduction

According to statistical information by Ministry of Agriculture, Forestry and Fisheries of Japan (MAFF) [1], the agricultural workforce in 2020 is 1.4 million workers, which is 22.5% of 2015 in spite of world population increase. In addition, the rate of elder workers (over 65 years old) increased 4.9% to 69.8%. The decrease of the number of agricultural workers and their aging have become one of social problems in Japan, whose reasons are heavy work, knowledge transfer from experts to young workers, the large initial investment when starting farming, and a small income compared to the investment, etc.

Recently smart agriculture using AI, IoT, big data and robot technologies has been attracting attention, and it is expected to improve the agricultural works in efficiency and contribute to labor-saving by automated operations, parameterization and visualization of expert farmers and sharing. In the automation of agricultural work, research on agricultural harvesting robots has been actively carried out, e.g., Hayashi et al. developed a strawberry harvesting robot with obstacle avoidance control algorithm to approach the target fruit without damaging the neighboring fruits [2], Henten et al. developed an autonomous robot for removing unwanted cucumber leaves [3] and Lehnert et al. realized autonomous harvesting of sweet pepper using mobile robot with 6-DOF robot arm, end-effector and RGB-D camera[4].

We have been organizing the Tomato Harvesting Robot Competition since 2014 with the aim of developing agri-robots, arousing the interests of young researchers in robotics into agriculture, and giving back the developed technology to the agricultural field [5]. The competition is divided into two leagues, Junior league for outreach activities to junior and senior high school students and Senior league for actual harvesting demonstrations. In the junior league, each team develops a robot that harvested small tomatoes using LEGO Mindstorm. In the senior league, competitors demonstrate the harvesting performances of their developed robot(s) using tomato plants for commercial production. Fujinaga et al. also reports the mosaic image of tomato plants using IR and RGB-D images for estimation of maturity of tomato fruits, which are also developed through the competitions [6] [7].

In this paper, we report the results of the 6th Tomato-Harvesting-Robot Competition and discuss the regulation changes toward actual applications.

2. Smart Agriculture and Tomato-Harvesting-Robot Competition

2.1 Smart Agriculture

Smart agriculture is a new style of agriculture that utilizes robot technology and information and communication technology (ICT) to promote labor saving, precision, and high-quality production. By utilizing smart agriculture that makes full use of advanced technology, it is possible to overcome the issues in agricultural work, secure new farmers, and improve cultivation technology.

Japanese agriculture is facing a serious labor shortage due to the aging of individual farmers. It is required to support such hardships in Japanese agriculture by utilizing ICT. The smart agriculture is expected for labor saving and labor reduction of agricultural work.

The second issue is the transfer of cultivation technology to new farmers. In the old system, the knowledge is transferred in families or small groups. The shortage of human resources makes difficult to inherit expert knowledge in agriculture in the old fashion.

Japan's food self-sufficiency rate (calorie basis) was about 40% in FY2018, and imports far exceed domestic production. In order to increase the yield and increase the self-sufficiency rate in the face of the above-mentioned shortage of human resources, automation by IoT, AI and robots is indispensable for reliably growing agricultural products with a small number of workers. We are expecting to contribute Smart Agriculture through the Tomato-Harvesting-Competitions.

2.2 Tomato-Harvesting-Robot Competition

The Tomato-Harvesting-Robot competition started in 2014 and the 8th competition was held in winter of 2021, which consists of two leagues, the Senior League and the Junior League. The target competitors for Senior League are supposed to have and automated and self-contained robots, and the Junior League are for young students, where they make teams with 2 to 4 students and make a robot with LEGO Mindstorms or Spike, or their own mobile robots.

In the Senior League, each team developed one or two robots that harvests medium tomatoes (approximately 60 to 120 grams) provided by Hibikinada green farm Ltd.. Each team chooses one of two areas depending on mobility, the rail-style area assuming the operation in the greenhouse with rails or the free-style area assuming the outdoor cultivation.

The regulation is revised as shown on Table 1. From 2020, the place of senior league is changed to our Greenhouse(Fig.1) in our campus, Kitakyushu Research & Science Park. The field of free-style area is soil field of the greenfield. So that all teams are required to adapt the change of lighting condition depending on time and weather, which will make the image processing to adapt outdoor environment

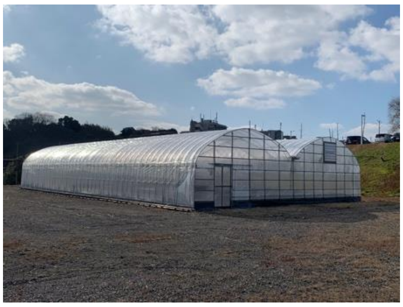


Fig. 1 Overview of Greenhouse

Table 1 History of regulation revisio

N-th	Year	Rule & Changes
1	2014	Senior League started with 6 categories from T1 (manual operation) to T6 (autonomous) and 2 kinds of working fields (rail and artificial grass)in indoor field (Gymnastium). The black boards are set behind tomato plants.
2	2015	Senior: No rule change. Junior: Junior League started. Lego mindstorm is used for basic platform.
3	2016	Senior: Bump is placed in the center of grass field. Junior: No rule chage.
4	2017	No rule change.
5	2018	Senior: Slope is placed instead of Bump in the grass field. The black boards are removed. In scoring, success rate is added. Junior: Original arm made of stationery is allowed. Tomato box removed.
6	2019	No rule change.
7	2020	Senior: Video evaluation instaed of 1st and 2nd rounds. The cometition field is changed to the ourdoor experimental green house. The grass field is changed to the soil field. Junior: Online cometition in each school.

2.3 Competition site, Greenhouse in Our Campus

The greenhouse developed for smart agriculture is shown in Fig.2. The house size is 10m x 20m and the half area is designed for cultivation with soil and the other is for hydroponics. Currently, as agricultural IoT Sensors, temperature, moisture data, pH regular measurement of soil, regular measurement of temperature, humidity and illuminance in the greenhouse, photography of plants, have been implemented. As actuators for control air condition and soil condition, opening and closing the greenhouse walls, spraying mist, water supply control by solenoid valves, water circular system in hydroponics, the system for purification of rainwater and control of nutrients are under implementation.

Using the experimental house, we will encounter to various problems in agriculture. For example, recognition of green worms imitating the leaf vein, disease of plants, harvesting and so on. Transmitting data to cloud data base is also one research topics with limited communication.

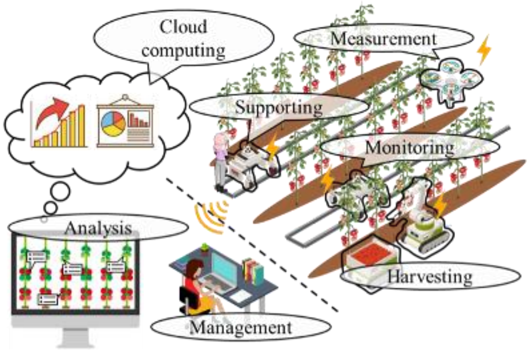


Fig.2 Overview and concept of the Greenhouse

2.4 Results

Table 2 shows the result of 9th competition. The winner team, HAYAHI-LAB (Fig.3)succeeded to harvest 18(=14+1+4) tomatoes in 10 min and the success rate was 72%. The second team (Fig.4) got 5 tomatoes and the rate is 40%. The speed of harvesting is getting faster and the success rate is improved.

Table 2 The results of the competition

Team	Tomato Condition					Damage to Stem (-5)	Harvest Rate	Points	Rank
	Get	Get & Unripe	Get & Dameages	Drop	Remain & Damages				
Hibikino-Toms	2	0	1	2	0	0	0.40	12.00	2
	3	0	2	2	0	0	0.43	23.43	
Tomastar	0	0	0	4	0	0	0.00	-8.00	
	0	0	0	0	0	0	0.00	0.00	
Mizurin	0	0	0	1	0	0	0.00	-2.00	
	1	0	0	5	0	0	0.17	-7.33	
SugarLab	3	1	3	2	0	1	0.33	17.67	3
	3	3	2	1	2	0	0.27	18.00	
HAYASHI-LAB	13	1	4	0	0	0	0.72	179.11	1
	10	1	2	2	1	0	0.63	109.00	
Iida Lab	2	0	1	0	2	0	0.40	12.00	
	1	0	1	0	3	0	0.20	-1.20	
HSRL-Tomato	0	0	0	0	0	0	0.00	0.00	
	0	0	0	0	0	0	0.00	0.00	
TOMASON	0	0	0	0	0	0	0.00	0.00	
	1	0	1	0	0	0	0.50	12.00	

Fig. 3 The winner Team HAYASHI-LAB

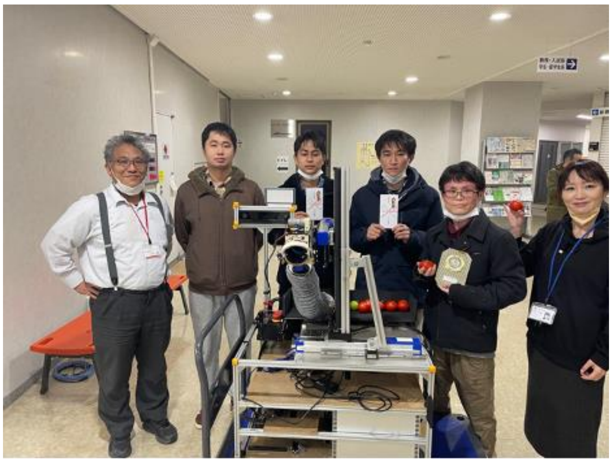


Fig. 4 The second team Hibikino-Toms

3. Conclusion

In this paper, we report on the 9th Tomato Robot Competition, which aimed at the social synthesis of robots. We also summarized the changes in the rules of the Senior League since the first competition. The operating environment was modified to be similar to that of a real farm, and the calculation method of the score was modified to aim at harvesting accurately. We have been working for experimental greenhouse. The agricultural IoT devices are implemented toward smart agriculture. The house system will include IoT sensors and actuators to control environment, robots for monitoring, harvesting, and transportation. Next regulation change will be usage and combination of IoT devices toward smart agriculture.

References

1. 2020 CENSUS of Agriculture and Forestry in Japan, 2020
2. Hayashi et. al, Collision-free control of a strawberry-harvesting robot by recognition of immature fruits, Journal of Science and High Technology in Agriculture, Vol.25, No.2, pp. 29-37, 2013
3. Henten et. al, An Autonomous Robot for De-leafing Cucumber Plants grown in aHigh-wire Cultivation System, Biosystems Engineering, Vol.94, No.3, pp. 317–323, 2006
4. Lehnert et. al, Autonomous Sweet Pepper Harvesting for Protected Cropping Systems, IEEE Robotics and Automation Letters, Vol.2, No.2, pp. 872–879, 2017
5. Matsuo et. al, Toward Smart Tomato Greenhouse: The Fourth Tomato Harvesting Robot Competition, Journal of Robotics, Networking and Artificial Life, Vol. 6, No.2, pp. 138-142, 2019
6. Takuya Fujinaga et al., Image Mosaicing Using Multi-Modal Images for Generation of Tomato Growth State Map, J. Robot. Mechatron., Vol.30, No.2, pp. 187-197, 2018.
7. Takuya Fujinaga, et al., Tomato Growth State Map for the Automation of Monitoring and Harvesting, J. Robot. Mechatron., Vol.32, No.6, pp. 1279-1291, 2020.

Authors Introduction

Prof. Kazuo Ishii

He is currently a Professor at the Department of Human Intelligence Systems of Kyushu Institute of Technology, Japan. He obtained his M.S. degree in 1993 and his D. Eng. degree in 1996 at The University of Tokyo. His research interests are in the fields of Underwater Robotics, Field Robotics, Neural Networks and Intelligent Systems.

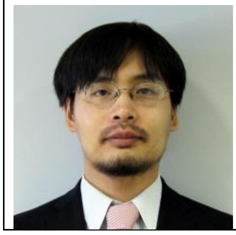
Dr. Takayuki Matsuo

He is an Associate Professor at Department Creative Engineering, National College of Technology (KOSEN), Kitakyushu College, Japan. His research area is underwater robots and biomimetic robots.

Dr. Yasunori Takemura

He is an Associate Professor at Department Integrated System Engineering, Nishi-Nippon Institute of Technology, Fukuoka, Japan. His research area is about machine learning, data mining and Robotics

Dr. Takashi Sonoda

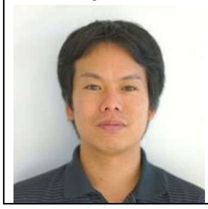


He is an Associate Professor at Department Integrated System Engineering, Nishi-Nippon Institute of Technology, Fukuoka, Japan. His research are underwater robotics and robot manipulator systems

Dr. Moeko Tominaga

She received Ph.D. degree from Kyushu Institute of Technology, Japan in 2021 and now she is working as an Assistant Professor at Nishinippon Institute of Technology, in Japan. Her research area is about robotics and machine learning. In particular, she is interested in inter-agent cooperation in multi-agent systems.

Dr. Yuya Nishida



He is an Associate Professor at Graduate School of Life Science and System Engineering, Kyushu Institute of Technology, Japan. His research area is about filed robotics, its application, and data processing.

Dr. Shinsuke Yasukawa

He received his Ph.D. degree in electrical, electronic, and information engineering from Osaka University, Japan, in January 2017. Currently, he is an Associate Professor in the Department of Human Intelligence Systems, Kyushu Institute of Technology, Japan. His research interests include information processing in biological systems and their applications in robotics.

Ms. Kanako Shirahashi

She belongs to the Center for Socio-Robotic Synthesis (CSRS), Kyushu Institute of Technology. She supports and manages the projects of CSRS, such as Tomato-Harvesting Robot Cometition, Underwater Robot Comepition, Ksrp Ekiden Tournament. Etc.

A Modeling of Sphere Considering Slipping Adapted Three-Rollers

Kenji Kimura
Department of Control Engineering, National Institute of Technology, Matsue College
14-4 Nishi-ikuma-cho, Matsue-shi,Shimane,690-8518, Japan
*E-mail: k-kimura@matsue-ct.jp**

Kouki Ogata
Department of Physic, Faculty of Science and Engineering, Saga University
1 Honjo, Saga 840-8421, Saga, Japan
E-mail: ikokatago@gmail.com

Hiroyasu Hirai
Graduate School of Life Science and engineering, Kyushu Institute of Technology
2-4 Hibikino,Wakamatsu-ku, Kitakyushu-shi 808-0196, Fukuoka, Japan
E-mail: multipletitan@gmail.com

Takumi Ueda
Graduate School of Life Science and engineering, Kyushu Institute of Technology
2-4 Hibikino,Wakamatsu-ku, Kitakyushu-shi 808-0196, Fukuoka, Japan
E-mail: ueda@brain.kyutech.ac.jp

Kazuo Ishii
Graduate School of Life Science and engineering, Kyushu Institute of Technology
2-4 Hibikino,Wakamatsu-ku, Kitakyushu-shi 808-0196, Fukuoka, Japan
E-mail: ishii@brain.kyutech.ac.jp

Abstract

Many types of spherical robots use friction-drive systems for locomotion because such systems enable omnidirectional movement and are more capable of climbing steps than mobile robots equipped with multiple omni-wheels. Slipping between spheres and rollers is a remarkable issue with friction-driven mechanisms. However, the previously established sphere kinematic models do not consider slipping, and kinematic models consider slipping in only two constraint rollers. In this study, we propose a mathematical model that allows for slipping on three constraint rollers and simulate the angular velocity vector of the sphere and slip speed at each contact point.

Keywords: Angular velocity vector of the sphere, Motion analysis of the sphere, Slip velocity of the sphere

1. Introduction

A sphere is one of the main shapes of a robot. It is used not only as a multifingered fingertip mechanism for hand robots but also as an actuator transmission mechanism for omnidirectional movement and drive in mobile robots. Spheres are also used as driving rollers for omnidirectional movement mechanisms, with various arrangements and sphere structures depending on the application of the movement mechanism. **Figure 1** shows the roller contact type for the number of actuators ($\boldsymbol{N_R}$) per sphere.

In the case of $N_R = 2$, ACROBAT-S [1], wheel chair [2] have sphere kinematics (**Figures 1**(a) and (b)). The omnidirectional condition is that the rollers are arranged on the equator of the sphere[3]. Furthermore, the angular velocity vector of the sphere has two degrees of freedom. Theoretically, it is considered in [4].

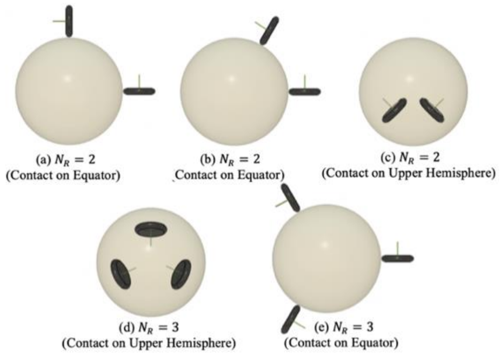


Figure 1 Type of roller arrangement for sphere mobile robot

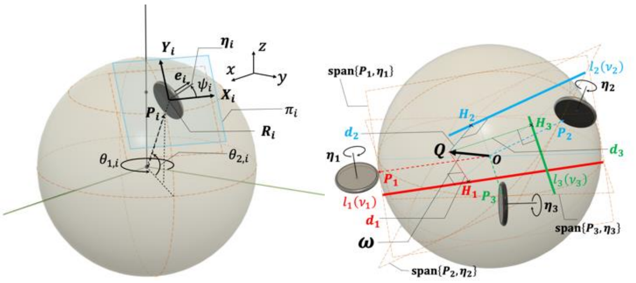
As shown in **Figure 1**(c), the ball-holding mechanism [5] is designed to transport the ball. All robots in the RoboCup middle-size league (MSL) use a ball-dribbling mechanism to control the rotation of the ball, which is implemented with two rollers on the upper half of the ball. Most designs employ slip-roller arrangements, which are determined heuristically in experiments in the absence of suitable mathematical models because of their strong friction force and enhanced ball-holding ability. Here, the roller is arranged in the upper hemisphere with a slip at the contact point between the roller and the sphere.

In a previous study, we used two constraint rollers that allow for slipping to derive a mathematical model of sphere rotational motion [6]. This model is included in the kinematics of [4]. Furthermore, we employed experiment [7] to validate the model of [6].

In the case of $N_R = 3$, omnidirectional wheeled mobile platform (OWMP) [8] has three constraint rollers and a ball-balanced robot [9] has three unconstraint rollers (**Figures 1**(d)). The sphere rotational dimensions are different because of the roller structure. Each constrained roller has less rotational diversity than the unconstrained rollers. However, the holding force is stronger than that of the unconstrained roller. The stability of the sphere is higher in the case of three rollers than in the case of two rollers.

OWMP [8] is kinematic with a roller arrangement restricted to the equator; we extend this to an arbitrary arrangement discussion.

In this study, we modify the previously developed kinematic model [6] in the case of three constraint rollers and present a mathematical model of sphere rotational motion.



(a) (b)

Figure 2 (a). Roller axis vector $\boldsymbol{\eta}_i$ at contact point $\boldsymbol{P}_i$ on the sphere. (b). The existence of sphere angler velocity vector i

2. The sphere forward kinematics by three constraint rollers

In this section, we derive the angular velocity vector of the sphere to geometrically model.

2. 1 *The existence of angular velocity vector of the Sphere*

As shown in **Figure 2**(a), the center $\boldsymbol{O}$ of a sphere with radius r is fixed as the origin of the coordinate system $\Sigma - xyz$. The *i-th* constraint roller have mass $\boldsymbol{R}_i$ and contact point $\boldsymbol{P}_i$ with respect to sphere.

$\boldsymbol{\eta}_i$ denotes the unit vector along the rotational axis of the constraint roller. it has a starting point at $\boldsymbol{R}_i$ ($\boldsymbol{O}$, $\boldsymbol{P}_i$ and $\boldsymbol{R}_i$ are on the same line). $\boldsymbol{\omega}$ denotes the angular velocity vector of the sphere. ν_i denotes the peripheral speed of the constraint roller. The velocity vector of the sphere $\boldsymbol{\mathcal{V}}_i^S$ at $\boldsymbol{P}_i$ can be represented as $\boldsymbol{\mathcal{V}}_i^S = \boldsymbol{\omega} \times \boldsymbol{P}_i$. And $\boldsymbol{\mathcal{V}}_i^R$ denotes velocity vector of the roller. $\boldsymbol{e}_i \in \mathbf{span}\{\boldsymbol{P}_i, \boldsymbol{\eta}_i\}$ denotes unit normal vector along $\boldsymbol{\mathcal{V}}_i^R$. $\boldsymbol{e}_i$ and $\boldsymbol{\mathcal{V}}_i^R$ are satisfy $\nu_i = \langle \boldsymbol{\mathcal{V}}_i^R, \boldsymbol{e}_i \rangle$ ($\boldsymbol{\mathcal{V}}_i^S = \boldsymbol{\mathcal{V}}_i^R$: nonslip condition) Thus, $\boldsymbol{\omega}$ can be satisfied as follow.

$$\langle \boldsymbol{\eta}_i, \boldsymbol{\omega} \rangle = -\frac{\nu_i}{r} \tag{1}$$

$\boldsymbol{\omega}$ must be on $\mathbf{span}\{\boldsymbol{\eta}_i, \boldsymbol{P}_i\}$ and can be represented as a following line $l_i(\nu_i)$ that is parallel to $\boldsymbol{P}_i$ and passes through the end point of $-(\nu_i/r)\boldsymbol{\eta}_i$.

$$l_i(\nu_i) = \left\{\boldsymbol{\omega} \middle| \left(-\frac{\nu_i}{r}\right)\boldsymbol{\eta}_i + t(1/r)\boldsymbol{P}_i, t \in \mathbb{R}\right\} \tag{2}$$

$\boldsymbol{R}_i$ is located along the plane π_i parallel to the tangent plane of the sphere at $\boldsymbol{P}_i$ (polar coordinate). We put vectors as normal orthogonal base $\{\boldsymbol{X}_i, \boldsymbol{Y}_i\}$ on π_i at start point $\boldsymbol{R}_i$. Thus, $\boldsymbol{\eta}_i$ is linear combination of Eq. (6) and rotates counterclockwise with respect to ψ_i.

$$\boldsymbol{\eta}_i = [\boldsymbol{X}_i \cos\psi_i + \boldsymbol{Y}_i \sin\psi_i] \tag{3}$$

Where

$$\boldsymbol{X_i} = \begin{bmatrix} -\sin\theta_{1,i} \\ \cos\theta_{1,i} \\ 0 \end{bmatrix},\ \boldsymbol{Y_i} = \begin{bmatrix} -\sin\theta_{2,i}\cos\theta_{1,i} \\ -\sin\theta_{2,i}\sin\theta_{1,i} \\ \cos\theta_{2,i} \end{bmatrix} \quad (4)$$

2.2 Calculation *of Optimal point in sum of the squared distances*

we calculate the optimal point $\boldsymbol{Q_o} = (x_0, y_0, z_0)(\in \mathbb{R}^3)$, which is determined such that the sum of the squared distances between $\boldsymbol{Q} = (x, y, z)(\in \mathbb{R}^3)$ and $l_i(v_i)(i = 1, 2, 3)$ is minimized.

As shown in Figure 2(b), the distances between $\boldsymbol{Q}$ and $l_i(v_i)$ in each line $l_i(v_i)$ is represented. Therefore, the sum of the squared distances is represented as follow:

$$L(x, y, z) = {d_1}^2 + {d_2}^2 + {d_3}^2 \quad (5)$$

where

$$d_i = \left\| \left(-\frac{v_i}{r}\right)\boldsymbol{\eta}_i + \frac{\langle \boldsymbol{P}_i, \boldsymbol{Q}\rangle}{r^2}\boldsymbol{P}_i - \boldsymbol{Q} \right\| \quad (6)$$

$(x, y, z) = (x_0, y_0, z_0)$ such that $L(x, y, z)$ is minimal value is satisfy as following.

$$z_0 = -\frac{E_9}{2E_3}, y_0 = \frac{D_5E_9 - 2D_8E_3}{4D_2E_3}$$

$$x_0 = \frac{1}{8C_1D_2E_3}(-C_4D_5E_9 + 2C_4D_8E_3 + 2C_6D_2E_9 - 4D_2E_3C_7) \quad (7)$$

Where

$$E_3 = D_3 - \frac{{D_5}^2}{4D_2}, E_9 = D_9 - \frac{D_5D_8}{2D_2}, E_{10} = D_{10} - \frac{{D_8}^2}{4D_2} \quad (8)$$

Where

$$D_2 = C_2 - \frac{{C_4}^2}{4C_1}, D_3 = C_3 - \frac{{C_6}^2}{4C_1}, D_5 = C_5 - \frac{C_4C_6}{2C_1}$$

$$D_8 = C_8 - \frac{C_4C_7}{2C_1}, D_9 = C_9 - \frac{C_6C_7}{2C_1}, D_{10} = C_{10} - \frac{{C_7}^2}{4C_1}$$

(9)

Where

$$C_{10} = L(0,0,0)$$

$$C_1 = \frac{1}{2}\big(L(1,0,0) + L(-1,0,0) - 2L(0,0,0)\big)$$

$$C_2 = \frac{1}{2}\big(L(0,1,0) + L(0,-1,0) - 2L(0,0,0)\big)$$

$$C_3 = \frac{1}{2}\big(L(0,0,1) + L(0,0,-1) - 2L(0,0,0)\big)$$

$$C_7 = \frac{1}{2}\big(L(1,0,0) - L(-1,0,0)\big)$$

$$C_8 = \frac{1}{2}\big(L(0,1,0) - L(0,-1,0)\big)$$

$$C_9 = \frac{1}{2}\big(L(0,0,1) - L(0,0,-1)\big) \quad (10)$$

$$C_4 = \frac{1}{4}\big(L(1,1,0) - L(-1,1,0) - L(1,-1,0) + L(-1,-1,0)\big)$$

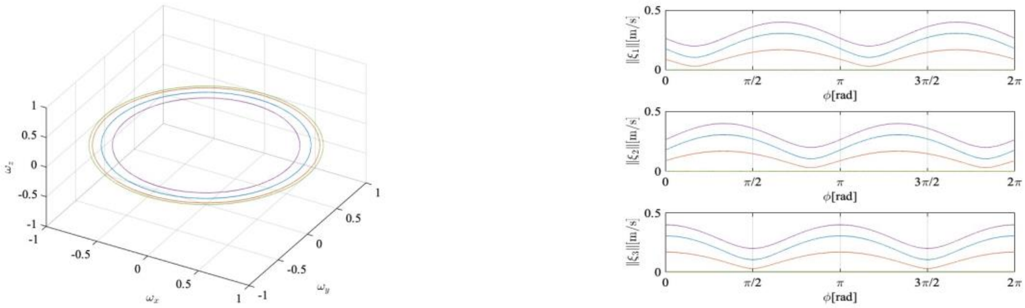
$$C_5 = \frac{1}{4}\big(L(0,1,1) - L(0,-1,1) - L(0,1,-1) + L(0,-1,-1)\big)$$

$$C_6 = \frac{1}{4}\big(L(1,0,1) - L(1,0,-1) - L(-1,0,1) + L(-1,0,-1)\big)$$

3. Simulation

This section presents the simulation results, including the trajectory of the endpoint of the angler velocity vector $\boldsymbol{\omega}_k$ and slip speed of sphere and roller $\|\boldsymbol{\zeta_k}\|$ in *k-th* pattern of roller arrangement ($k = 1,2,3,4$) in the case in which a regular triangle ($\theta_{1,1}$, $\theta_{1,2}$, $\theta_{1,3}$) = $(30°, 150°, 270°)$and $\psi_i = 0°(i = 1,2,3,4)$. The patterns are set up by $\theta_{2,i}$ ($i = 1,2,3,4$) as follows: Pattern I ($k = 1, \theta_{2,i} = 0°$), Pattern II ($k = 2, \theta_{2,i} = 10°$), Pattern III ($k = 3, \theta_{2,i} = 20°$), Pattern IV ($k = 4, \theta_{2,i} = 30°$).

As input, we define function $\nu_1(\varphi) = \sin(\varphi + 240°)$,$\nu_2(\varphi) = \sin(\varphi + 120°)$ and$\nu_3(\varphi) = \sin\varphi$.



(a) (b)

Figure 3 Simulation and comparisons in *k-th* pattern of roller arrangement ($k = 1,2,3,4$) (a). Trajectory of end point of angular velocity vector $\boldsymbol{\omega}_k$ (b). Slip speed of sphere and roller $\|\boldsymbol{\zeta}_{1,k}\|$, $\|\boldsymbol{\zeta}_{2,k}\|$, $\|\boldsymbol{\zeta}_{3,k}\|$.

As output, $\boldsymbol{\omega}_k$ and $\|\boldsymbol{\zeta}_k\|$ $(k = 1,2,3,4)$ were indicated, such as Pattern I [$k = 1$; green curve], Pattern II [$k = 2$; red curve], Pattern III [$k = 3$; blue curve], and Pattern IV [$k = 4$; violet curve] (**Figure 3**).

As shown in **Figure** 3(a), $\boldsymbol{\omega}_k$ $(k = 1,2,3,4)$ draws circle trajectories and gets a small radius in turn.

As shown in **Figure** 3(b), duo to Pattern I (nonslip case), $\|\boldsymbol{\zeta}_{1,1}\|$, $\|\boldsymbol{\zeta}_{2,1}\|$, and $\|\boldsymbol{\zeta}_{3,1}\| = 0$ [m/s]. $\|\boldsymbol{\zeta}_{1,k}\|$, $\|\boldsymbol{\zeta}_{2,k}\|$, $\|\boldsymbol{\zeta}_{3,k}\|$ have minimal values of 0.03, 0.10, and 0.19 [m/s] and maximal values of 0.16, 0.30, and 0.40 [m/s], respectively.

4. Conclusion

In this study, we considered the existence of an angular velocity vector for the sphere and proposed a sphere forward kinematics model that allows for slipping. Furthermore, we demonstrated the trajectory of the endpoint of the angler velocity vector and behavior of slip speed. and obtained maximal and minimal values.

In future research, this model will be verified experimentally. It could also be applied to a mobile robot.

Reference

[1] M.Wada, K.Kato, "Kinematic modeling and simulation of active-caster robotic drive with a ball transmission (ACROBAT-S)". 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems. Daejeon, 2016-12-9/25, IEEE Robotics and Automation Society,

[2] S.Ishida , H.Miyamoto,"Holonomic Omnidirectional Vehicle with Ball Wheel Drive Mechanism, 2012 The Japan Society of Mechanical Engineers.Vol.78,No.790, pp.2162-2170, 2012.

[3] K. Kimura, K. Ishii, Y. Takemura, M. Yamamoto, Mathematical Modeling and motion analysis of the wheel based ball retaining mechanism, *SCIS & ISIS*, pp.4106-4111, 2016.

[4] K. Kimura, S. Chikushi, et al, Motion Analysis of a Sphere Driven by Rollers, Journal of the Robotics Society of Japan. Vol.38, No.5, pp.485-495, 2020.

[5] S. Chikushi, M. Kuwada, et al., Development of Next-Generation Soccer Robot"Musashi150"for RoboCup MSL,*30th Fussy System Symposium*, pp. 624-627,2014.

[6] K. Kimura, K. Ogata, K. Ishii, Novel Mathematical Modeling and Motion Analysis of a Sphere Considering Slipping, Journal of Robotics, Networking and Artificial Life, Vol.6, issue 1, pp. 27-32, 2019.

[7] K. Kimura, S. Chikushi, K. Ishii, Evaluation of the Roller Arrangements for the Ball-Dribbling Mechanisms adopted by RoboCup Teams, Journal of Robotics, Nrtworking and Artificial Life, Vol.6, issue 3, pp. 183- 190, 2019.

[8] M. Kumagai, T. Ochiai: "Development of a robot balanced on a ball – Application of passive motion to transport",Proc. ICRA IEEE(2009), pp. 4106-4111, 2009.

[9] Lee,Y.C,Danny,Lee,D.V.,Chung,J.,andVelinky,S.A.,"Control of a redundant, reconfigurable ball wheel drive mechanism for an omunidirectional platform",Robotica, Cambridge University Press,Vol.25, pp.385-395, 2007.

Authors Introduction

Dr. Kenji Kimura

He is a Lecturer in Department of Control Engineering, National Institute of Technology, Matsue College, He received his ME (mathematics) from Kyushu University in 2002 and received his Ph.D. degree in engineering from Kyushu Institute of Technology in 2020.

Mr. Kouki Ogata

He received education in Fukuoka Daiichi high school until 2016, He is currently undergraduate School student at Saga University. His current research interests include sphere mobile robot kinematics, analytical dynamics.

Mr. Hiroyasu Hirai

He received his B.E. and M.E., in Computer Science from Nippon Bunri University, Japan, in 2013 and 2016, respectively. He is a 3rd year student in the doctoral program of the Kyushu Institute of Technology.

Mr. Takumi Ueda

He received his M.E., in Engineering from Kyusyu Institute of Technology, Japan, in 2017. He is a 3rd year student in the doctoral program of the Kyushu Institute of Technology. His research interest includes non-linear control by PID using database.

Dr. Kazuo Ishii

He is a Professor in the Kyushu Institute of Technology. He received his Ph.D. degree in engineering from University of Tokyo, in 1996. His research interests span ship marine engineering and Intelligent Mechanics.

Roller Arrangement Problem of Omnidirectional Mobil Robot Adapted Three Omni Rollers

Kenji Kimura
Department of Control Engineering, National Institute of Technology, Matsue College
14-4 Nishi-ikuma-cho, Matsue-shi,Shimane, 690-8518, Japan

Yuki Shigyo
Digital Solution Div, DX Engagement Dept, Fujitsu
3-5-20 Minamikamata. Ota City, Tokyo, 144-0035, Japan

Kazuo Ishii
Graduate School of Life Science and engineering, Kyusyu Institute of Technology
2-4 Hibikino,Wakamatsu-ku, Kitakyushu 808-0196, Fukuoka, Japan
E-mail:k-kimura@matsue-ct.jp, shigyo.yuki@fujitsu.com , ishii@brain.kyutech.ac.jp

Abstract

Mobile robots adapted to omni rollers are required to have efficient mobility in areas like logistics. Since such systems are easily controlled and provide omnidirectional locomotion. However, theoretical research regarding the motion efficiency has not been conducted. In this study, to evaluate roller arrangements with respect to speed efficiency, a mechanism where the roller placement position can be changed arbitrarily on a round shape mechanism and evaluate the robot speed efficiency is designed. We consider existence domain of the robot speed vector using the theory of linear transformation.

Keywords: Omni-roller, transformation matrix, robot mobile speed

1. Introduction

Recently, efficient mobility has become a requirement for mobile robots in areas like logistics. Thus, omnidirectional movement with either non-holonomic or holonomic characteristics can produce a total of 3 degrees of freedom motion (sum of 2 degrees of freedom translational motion and 1 degree of freedom rotational motion). Robotic vehicle development is getting attention.

Among them, the holonomic movement mechanism is simple to control because of the independently driven wheels and offers outstanding omnidirectional mobility. Thus, A mobile robot arranged in an equilateral triangle has been developed[1].

RV-infinity [2], Musashi150 [3], and NuBot [4] have adopted a mechanism placed in RoboCup MSL, that has three omni rollers arranged in an equilateral triangle. **Figure 1** shows three roller arrangement patterns, and **Figure 1**(i) shows the adopted A-type. Additionally, the types shown in **Figures 1** (ii) and (iii) are also conceivable, but intuitively, they use the most symmetrical triangle arrangement, and theoretical research is unconducted.

In this study, we generalize of equilateral triangle arrangement by assuming a mechanism in which the roller placement position on a round shapes mechanism can be changed arbitrarily. We derive kinematics that generalizes the kinematics in [1] and derives a transformation matrix, which associates input and output. Additionally, we evaluate roller arrangement, focusing on speed efficiency. As approaches, employing the area of robot velocity vector

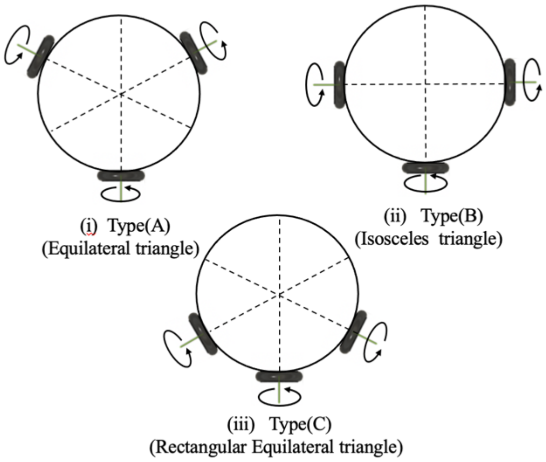


Figure 1 Omni rollers arrangement type of mobile robots in three rollers

as evaluation functions. Kinematics is derived and the linear transformation matrix, associating the input (roller speed) with the output (component of robot movement speed) is employed to measure the image size.

2. Kinematics of mobile robot adapted three omni rollers

This section introduces kinematics, which the roller placement position can be changed arbitrarily.

Assuming horizontal plane movement, the kinematics of both robots is similar. The primary focus lies in the connection between driving omni-wheels rotational speed and robot velocity.

2.1 *Inverse kinematics*

Figure 2 shows a top view of the mobile robot that has a common radius of all omni-wheels adapted The *i-th* rollers ($i = 1,2,3$) contact point $\boldsymbol{P_i}$ on a circle, which has radius R. X-Y is the global coordinate system (origin $\boldsymbol{O}$) and $\acute{X}$-$\acute{Y}$ is robot coordinate system(origin $\boldsymbol{\acute{O}}$). And. The robot orientation ϕ is referred as the angle between X-axis and $\acute{X}$-axis. and the distance from the robot center to the contact points between wheels and floor is R.

$\{\boldsymbol{e_1}, \boldsymbol{e_2}, \boldsymbol{e_3}\}$ represents the normal vector of rotational direction.

We compute the roller peripheral speed v_i that gives the robot translation speed $\boldsymbol{V} = [V_x, V_y]^{\boldsymbol{T}}$ and φ denotes the robot direction for the robot coordinate $\acute{X}$-$\acute{Y}$ and the robot rational speed $L\dot{\phi}$ ($\dot{\phi}$: robot angular velocity) and the roller peripheral speed v_i are decomposed translation and rotational components.

Rollers contact point $\boldsymbol{P_i}$ are the adapted angle θ_i measured from x-axis on the circle and has a radius L.

$$\boldsymbol{P_i} = L[\cos\theta_i, \sin\theta_i]^{\boldsymbol{T}} \tag{1}$$

Rollers velocity vector $\boldsymbol{e_i}$ is perpendicular to $\boldsymbol{P_i}$.

$$\boldsymbol{e_i} = [-\sin\theta_i, \cos\theta_i]^{\boldsymbol{T}} \tag{2}$$

v_i is represented as the sum of the translation component $\langle \boldsymbol{e_i}, [V_x, V_y]^{\boldsymbol{T}} \rangle$ and rotational component $L\dot{\phi}$.

$$v_i = \langle \boldsymbol{e_i}, [V_x, V_y]^{\boldsymbol{T}} \rangle + L\dot{\phi} \tag{3}$$

Thus, $[v_1, v_2, v_3]^{\boldsymbol{T}}$ is represented as follows.

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} -\sin\theta_1 & \cos\theta_1 & 1 \\ -\sin\theta_2 & \cos\theta_2 & 1 \\ -\sin\theta_3 & \cos\theta_3 & 1 \end{bmatrix} \begin{bmatrix} V_x \\ V_y \\ L\dot{\phi} \end{bmatrix} \tag{4}$$

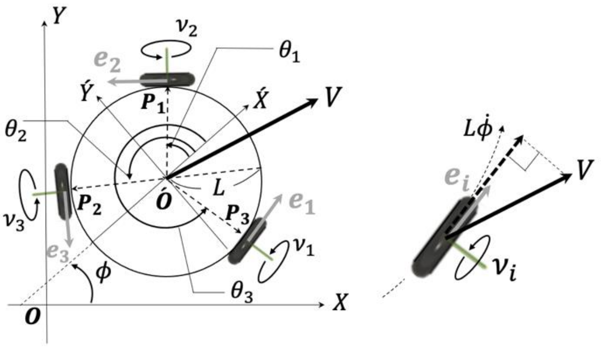


Figure 2 A top view of the mobile robot for kinematics.(i) Parameter list two coordinates (ii).Single rollers motion.

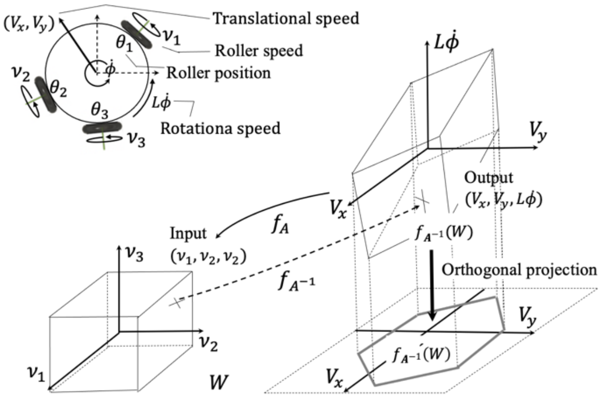


Figure 3 Structural transformation between the three rollers, robot mobile speed (V_x, V_y), and rotational speed $L\dot{\phi}$.

2.2 *Forward kinematics*

Eq.(4) is solved to $[V_x, V_y, L\dot{\phi}]^T$ as following;

$$\begin{bmatrix} V_x \\ V_y \\ L\dot{\phi} \end{bmatrix} = \frac{1}{\det\boldsymbol{A}} \cdot \begin{bmatrix} \cos\theta_2 - \cos\theta_3 & -\cos\theta_1 + \cos\theta_3 & \cos\theta_1 - \cos\theta_2 \\ \sin\theta_2 - \sin\theta_3 & -\sin\theta_1 + \sin\theta_3 & \sin\theta_1 + \sin\theta_2 \\ \sin(\theta_3 - \theta_2) & \sin(\theta_1 - \theta_3) & \sin(\theta_2 - \theta_1) \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \tag{5}$$

Where

$$\det\boldsymbol{A} = \sin(\theta_3 - \theta_2) + \sin(\theta_1 - \theta_3) + \sin(\theta_2 - \theta_1) \tag{6}$$

3. Evaluation function focus on robot speed vector area

In this section, the roller contact location is calculated using the orthogonal projection area of image as evaluation functions.

Figure 3 shows a cubic domain that has a square, 2 $[m/s]$ on each side and the volume, 8 $[m/s]^3$ as input.

$$W = \{(v_1, v_2, v_3) \| v_1|, |v_2|, |v_3| \leq 1\} \tag{7}$$

Here, from linear transformation mapping $f_{\boldsymbol{A}^{-1}} : [v_1, v_2, v_3] \rightarrow [V_x, V_y, \dot{\phi}L]$, $f_{\boldsymbol{A}^{-1}}(W) \in \mathbf{R}^3$ is a parallelepiped domain as follows (linear transformation matrix $\boldsymbol{A}$, $\det \boldsymbol{A}^{-1}$ is equal to $1/\det \boldsymbol{A}$. It is the volume ratio). Using $\boldsymbol{A}^{-1} = [\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_2]$, $f_{\boldsymbol{A}^{-1}}(W)$ is represented as follows.

$$f_{\boldsymbol{A}^{-1}}(W) = \tag{8}$$

$$\{(V_x, V_y, L\dot{\phi}) | v_1\boldsymbol{A}_1 + v_2\boldsymbol{A}_2 + v_3\boldsymbol{A}_3, \ |v_1|, |v_2|, |v_3| \leq 1\}$$

Next, Using the projection vector of $\boldsymbol{A}_i$ with respect to V_x, V_y-plane, we define the orthographic projection domain $f'_{\boldsymbol{A}^{-1}}(W)$ for horizontal plane as follows.

$$f'_{\boldsymbol{A}^{-1}}(W) = \tag{9}$$

$$\{(V_x, V_y) | \ v_1\acute{\boldsymbol{A}}_1 + v_2\acute{\boldsymbol{A}}_2 + v_3\acute{\boldsymbol{A}}_3, \ |v_1|, |v_2|, |v_3| \leq 1\}$$

Thus the orthographic projection area $D_{Are}(\theta_1, \theta_2, \theta_3)[m/s]^2$ is represented as sum of parallelogram area $I(\ ,\ )$;

$$D_{Are} = \begin{cases} 4I(\acute{\boldsymbol{A}}_1, \acute{\boldsymbol{A}}_2) + 4I(\acute{\boldsymbol{A}}_1 - \acute{\boldsymbol{A}}_2, \acute{\boldsymbol{A}}_3) \\ \qquad [\,(\acute{\boldsymbol{A}}_2 \times \acute{\boldsymbol{A}}_3)_z (\acute{\boldsymbol{A}}_1 \times \acute{\boldsymbol{A}}_3)_z < 0] \\ 4I(\acute{\boldsymbol{A}}_1, \acute{\boldsymbol{A}}_2) + 4I(\acute{\boldsymbol{A}}_1 + \acute{\boldsymbol{A}}_2, \acute{\boldsymbol{A}}_3) \\ \qquad [\,(\acute{\boldsymbol{A}}_2 \times \acute{\boldsymbol{A}}_3)_z (\acute{\boldsymbol{A}}_1 \times \acute{\boldsymbol{A}}_3)_z > 0] \end{cases} \tag{10}$$

where

$$\acute{\boldsymbol{A}}_1 = \begin{bmatrix} \cos\theta_2 - \cos\theta_3 \\ \sin\theta_2 - \sin\theta_3 \\ 0 \end{bmatrix}, \acute{\boldsymbol{A}}_2 = \begin{bmatrix} -\cos\theta_1 + \cos\theta_3 \\ -\sin\theta_1 + \sin\theta_3 \\ 0 \end{bmatrix}$$

$$\acute{\boldsymbol{A}}_3 = \begin{bmatrix} \cos\theta_1 - \cos\theta_2 \\ \sin\theta_1 + \sin\theta_2 \\ 0 \end{bmatrix} \tag{11}$$

Theoretically, we find out that D_{Are} was minimized in the equilateral triangle arrangement (Case of type(A)).

4. Simulation

This section presents the simulation findings, including the evaluation values of behavior and shape of orthogonal projection area D_{Are} and we assume that $L = 1$[m].

4.1 *Behavior of the evaluation function*

Figure 4 shows evaluation function D_{Are} when the symmetry arrangement is with respect to y-axis (substituting Eq.(10) with $\theta_1 = \theta$, $\theta_2 = 180° - \theta$, and $\theta_3 = 270°$, $-90° \leq \theta \leq 90°$).

D_{Are} takes the minimal value when $\theta = 30°$ (case of equilateral triangle arrangement).

4.2 *Shape of orthographic projection area*

Simulations were performed at the three various roller arrangement patterns that $\boldsymbol{P}_1$, $\boldsymbol{P}_2$ and $\boldsymbol{P}_3$ are arranged a symmetry triangle shaped(See Figure1).

They are set up θ_i $(i = 1,2,3)$ as follows:

Type(A): $(\theta_1, \theta_2, \theta_3) = (30°, 150°, 270°)$

Type(B): $(\theta_1, \theta_2, \theta_3) = (0°, 180°, 270°)$

Type(C): $(\theta_1, \theta_2, \theta_3) = (210°, 270°, 300°)$

Figure 5 shows an outlines of robot velocity distributions for Type(A), Type(B), and Type(C).

It shows that the orthogonal projections on the parallelepiped's horizontal planes are all hexagons.

Table 1 shows the area D_{Are} on the horizontal plane V_x– V_y of the parallelepiped as evaluation values.

Type (A) is the smallest. Additionally, since Type(C) is the largest, the moving area for the translational motion

Table 1 Comparison of the evaluation value

	Type (A)	Type (B)	Type (C)
D_{Are} $[m/s]^2$	4.61	6.00	13.85

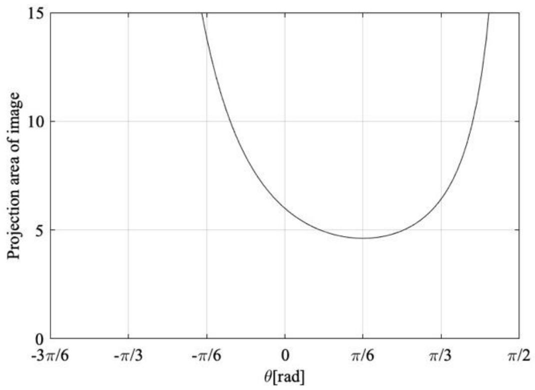
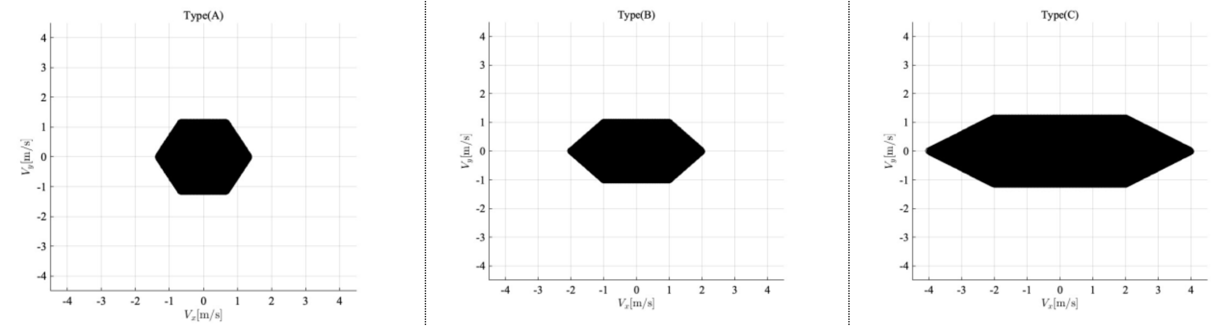


Figure 4 Behavior of evaluation function D_{Are} $[m/s]^2$



(i) Case of Type (A) (ii) Case of Type (B) (iii) Case of Type (C)

Figure 5 Projection domain of $f_{A^{-1}}(W)$ on horizontal plane (V_x, V_y)

is effective. Therefore, results were obtained for the evaluation.

5. Conclusions

In this research, we derived the kinematics of a mechanism that can arbitrarily change the roller placement position on a circle and considered the roller placement using the transformation matrix, associating input and output as an evaluation function.

As a result, theoretically, for Case of type(A), we find out that evaluation function was minimized in the equilateral triangle arrangement. Furthermore, in terms of translational motion, type(C) is the most efficient comprehensively among the three types.

References

[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.*

Authors Introduction

Dr. Kenji Kimura

He is a lecturer in the Department of Control Engineering, National Institute of Technology, Matsue College starting in 2022. He received the ME (mathematics) from Kyushu University in 2002 and received his Ph.D. degree in engineering from Kyushu Institute of in 2020. His research interests include spherical robot, control for object manipulation, and Riemannian geometry.

Mr. Yuki Shigyo

He is an engineer working for Fujitsu Ltd. He received the BE from Tokyo University of Science in 2016 and the ME (mathematics) from Nagoya University in 2018. His research interests are spherical mobile robot kinematics and machine learning.

Dr. Kazuo Ishii

He has been a professor at the Kyushu Institute of Technology since 1996. He received his Ph.D. degree in engineering from the University of Tokyo, Tokyo, Japan, in 1996. His research focuses on ship marine engineering and intelligent mechanics. He holds five patents derived from his research.

Underwater Live Video Streaming Experiment Using Radio Frequency Communication for AUVs

Raji Alahmad*
Department of Human Intelligence Systems, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka, 808-0196 Japan.†
E-mail: raji@brain.kyutech.ac.jp
https://www.kyutech.ac.jp/

Yuya Nishida[1], Kazuo Ishii[2]
Department of Human Intelligence Systems, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka, 808-0196 Japan
[1]*E-mail: ishii@brain.kyutech.ac.jp,* [2]*E-mail: y-nishida@brain.kyutech.ac.jp*

Yuki Fukumoto[3], Tohlu Matsushima[4]
Department of Electrical and Electronic Engineering, Kyushu Institute of Technology,
1-1 Sensuicho, Tobata, Kitakyushu, Fukuoka, 804-550 Japan.
[3]*E-mail: fukumoto@ele.kyutech.ac.jp,* [4]*E-mail: matsushima@ele.kyutech.ac.jp*

Abstract

Autonomous Underwater Vehicles (AUVs) require long-distance communication, especially in the deep sea. Radiofrequency (RF) communication provides a high data rate. However, electromagnetic wave is seriously limited by high attenuation in the water medium. In this study, we investigate Radiofrequency communication in seawater. The experiment results show the effects of the distance between the transmitter and receiver, and the stability of the antennas. We achieved HD video transmission with 25fps.

Keywords: AUV, Rf communication, Network.

1. Introduction

Developing Autonomous vehicles for underwater applications is an important topic in many sectors such as the exploration of marine resources, surveying seabed topography, and military applications [1].

The biggest challenge that faces researchers when developing AUVs is communication [2]. there are three main technologies used in underwater communication:

(i) Acoustics communication: this is the most common and used technology for underwater communication because it can provide long-distance propagation which can be in tens of kilometers. The major drawback of this technology is the latency in communication since the transmission baud rate is only up to several kilobits per second (kbps) [3].

(ii) Wireless Optical communication: by using wireless optical communication we can achieve a data rate of hundreds of Megabits per second (Mbps) with a distance of tens of meters, this is considered the best technology for a high data rate. However, the Line of

Sight (LOS) and the difficulty of aligning the transmitter and receiver are the disadvantages of this technology [4].

(iii) Radio Frequency (RF) communication: This technology provides a high data rate with less than 10 Mbps [5], but it is restricted by the high attenuation over a short distance.

In this study, we investigate the transmission speed of radio wave communication in seawater as well as analyze the underwater video framerate.

2. Experiment Setup and Design

2.1. *Experimental setup*

We used in this study for a stationary base station antenna a regular octagonal-shaped antenna with a diameter of 2m. And a rectangular antenna of 0.8 m by 0.5 m mounted on the AUV. The base antenna was mounted in the center of a pool with a size of (5.5m × 5.5m × 4.8m) as width, length, and depth respectively. Since the salinity of the water affected the conductivity, affected the connection as well, the experiment was conducted in seawater. Figure 1 illustrates the concept of the experiment.

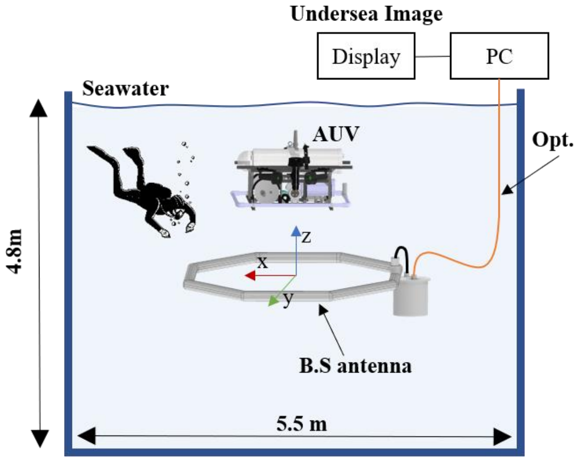

Undersea Image
Display
PC
Seawater
AUV
Opt.
4.8m
z
x
y
B.S antenna
5.5 m

Fig. 1. Experimental design concept.

The communication speed has been measured with different distances horizontal and vertical between the two antennas with the origin of the center of the base station antenna.

The AUV used in this study is called Darya Bird. It is an Autonomous Underwater Vehicle developed by a number of graduate students at Kyushu Institute of Technology [6]. The Robot consists of several high-pressure resistance hulls, they are connected using a T-slot frame designed in a way to easily add and remove components on demand. Figure 2 shows the AUV Darya Bird.

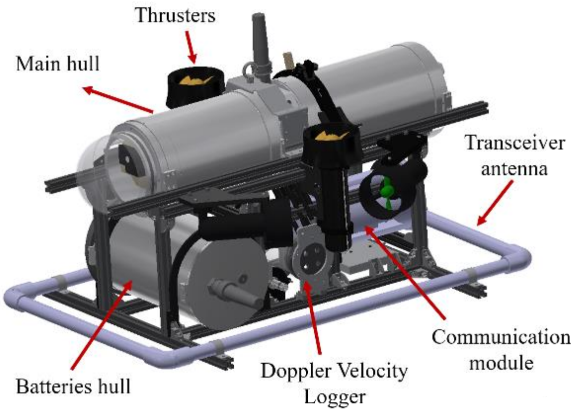

Thrusters
Main hull
Transceiver antenna
Communication module
Doppler Velocity Logger
Batteries hull

Fig. 2. The AUV used in the experiment (Darya Bird).

IP network camera used for video capturing. We used a Power over Ethernet (PoE) ejector to provide a 12 V PoE connection for the camera and the communication model. And we connect all to a LAN hub. The Ground PC was connected to the base antenna by an optical fiber cable, in addition, we used a wired optical connection to connect the AUV to the reference PC in order to capture the same streaming video. The reference video is used for the porpuse of comparison with the one taken by the wireless connection. The layout connection is shown in Figure 3.

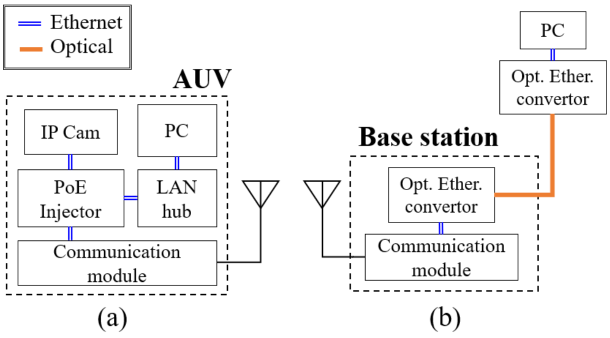

Ethernet
Optical
AUV
IP Cam
PC
PoE Injector
LAN hub
Communication module
Base station
PC
Opt. Ether. convertor
Opt. Ether. convertor
Communication module
(a)
(b)

Fig. 3. The layout connection (a) the connection inside the AUV, (b) the connection between the ground PC and the stationary antenna

Table 1 illustrate the camera setup parameters.

Table 1. The network camera setup parameter.

Parameter	Value
Resolution	1028 × 720
Frame rate	25 fps
Codec	H264

2.2. *Experiment design*

The measurement of the transmission speed was conducted in different placements of the AUV.

First, we place the center of the AUV with the center of the stationary antenna, then we moved the AUV along the z-axis (starting from the z=0 which is the depth of 2.4m) and measure the signal.Then we move horezental on the x-axis, afterward, we repeat moving on z-axis.

The Wavelet OFDM System is used with 2 to 28 MHz for the standard mode with a symbol length of 8.192 us.

3. Results and Discussion

The results of this study are divided into two parts:

3.1. *Transmission rate*

Table 2 illustrates the transmission rate of UDP and TCP. The transmission rate is higher when x=z=0, and goes lower when we moved on the z-axis far from the base antenna.

Table 2. The transmission rate of UDP and TCP.

x [m]	z [m]	UDP [Mbps]	TCP [Mbps]
0	0	6.8	4.7
0	1	4.8	3.3
0	1.65	0	0
1	0	6.8	4.5
1	1	0.4	1.7
1	1.2	0	0
2	0	6.5	4.6
2	1	0	0

As it is clear that the transmission rate in UDP is 6.8 Mbps at the center (x=z=0). This rate is higher than the TCP mode which is 4.7 Mbps.

It can be noticed that we could achieve communication even when the AUV was moved far from the base antenna on the x-axis.

3.2. *Video streaming framerate*

The video streaming was taken in two scenarios:

3.2.1. *The AUV is placed in x=0m, and z=0.5m*

During this experiment, a diver jumped into the pool and swam in front of the AUV to see the smoothness of the obtained video. The receiving video using the wireless channel was compared with a video has been obtained by the wired optical connection.

Figure 4 illustrates the framerate of the received videos when z=0.5m.

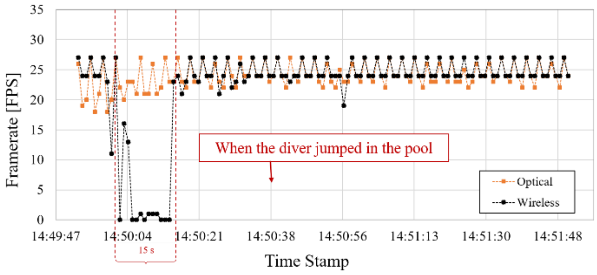


Fig. 4. The framerate of the received videos when z=0.5m.

It can be seen that the framerate decreased drastically to 0 fps for 15 sec when the diver jumped into the pool. While the rest of the framerate was stable with an average of 25 fps which is the same as the obtained one by wired connection, as well as the same as the camera's basic setting.

3.2.2. *The AUV is placed in x=0m, and z=1m*

In this experiment, the diver has jumped once again into the pool since the jumping has an effect on the received framerate. Figure 5 illustrates the framerate of the received video when z=1m.

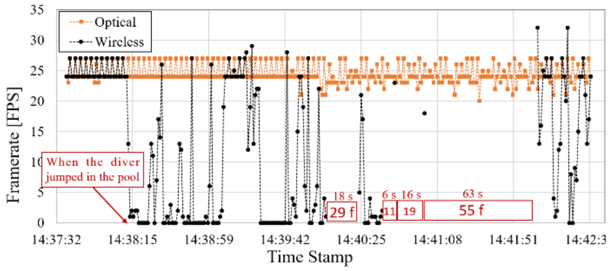


Fig. 5. The framerate of the received videos when z=1m.

In this scenario, the framerate has decreased drastically to 0 fps for about 10 seconds, however, the unstable connection continued for the rest of the video with a very bad connection in two parts: 29 frames in 18 seconds and 85 frames in 85 seconds.

Figure 6 shows snapshots from the received videos when z=1m. In the figure, we compare randomly obtained frames in both videos (wireless and wired).

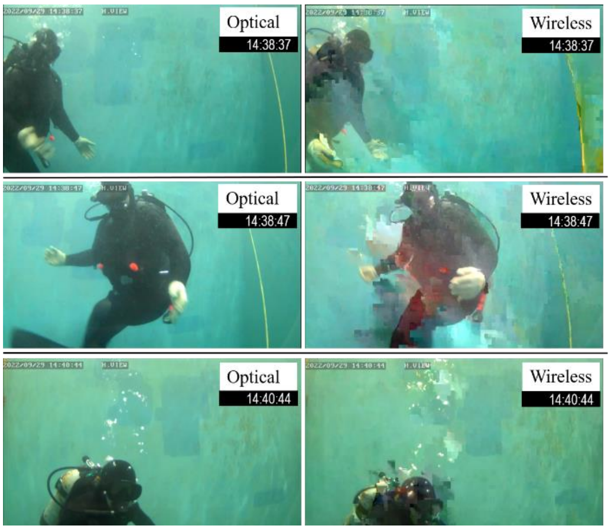


Fig. 6. Snapshots from the received videos when z=1m.

We noticed that even though the video streaming quality was not smooth, the snapshots of the same timestamp provide the same frame (the diver's position and his gesture are the same). However, two frames showed different timestamps for the same frame as shown in Figure 7.

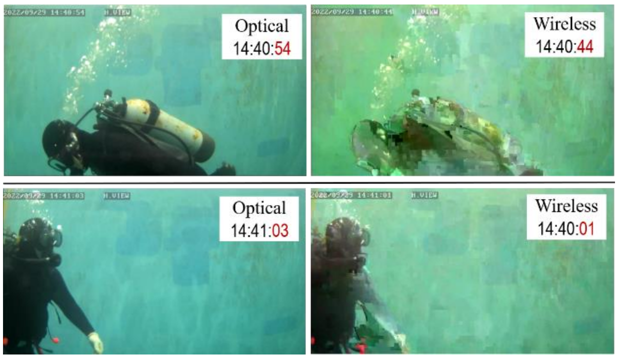


Fig. 7. Snapshots from the received videos when z=1m (Mismatch timestamp).

The reason for the missed frames can be attributed to two points:

(i) The instability of the antenna: when the diver jumped into the pool, he caused a strong vibration in the base antenna and the AUV. in the first scenario (z=0.5) this vibration continued for about 15 seconds which caused framerate =0 fps. While in the second scenario, where the distance between the two antennas was 1 meter, which made the slight vibration of the antenna effects strongly the signal. And we noticed from the underwater monitoring video that the diver hit the rope which fix the base antenna, which made more vibration in the base antenna.

(ii) The bubbles caused by the diver: when the diver jumped into the pool, a huge amount of bubbles were generated. These bubbles made additional attenuation of the non-line-of-sight signals.

4. Conclusion

This study introduced an experimental study of underwater live video streaming for AUVs applications. Loop antennas were used, stationary antenna and one mounted on the AUV. We achieved HD video streaming with 25fps on average within a distance of 1 meter. Further experiments are needed to confirm the effects of bubbles as well as the maximum distance to obtain stable streaming.

Acknowledgments

This study was part of commissioned research (No.02301) by National Institute of Information and Communications Technology (NICT), Japan.

References

1. Chen P, Li Y, Su Y, Chen X, Jiang Y. Review of AUV Underwater Terrain Matching Navigation. J Navig. 2015;68(6):1155-1172.
2. Heidemann J, Stojanovic M, Zorzi M. Underwater sensor networks: Applications, advances and challenges. *Philos Trans R Soc A Math Phys Eng Sci*. 2012;370(1958):158-175. doi:10.1098/rsta.2011.0214
3. Jiang W, Yang X, Tong F, Yang Y, Zhou T. A Low-Complexity Underwater Acoustic Coherent Communication System for Small AUV. *Remote Sens*. 2022;14(14):1-15. doi:10.3390/rs14143405
4. Al-Zhrani S, Bedaiwi NM, El-Ramli IF, et al. Underwater Optical Communications: A Brief Overview and Recent Developments. *Eng Sci*. 2021;16:146-186. doi:10.30919/es8d574
5. Ali MF, Jayakody DNK, Chursin YA, Affes S, Dmitry S. *Recent Advances and Future Directions on Underwater*

Wireless Communications. Vol 27. Springer Netherlands; 2020. doi:10.1007/s11831-019-09354-8

6. Tanaka Y, Alhamad R, Fujinaga T, et al. Multi Robot Strategy and Software Development of Robots for Underwater Survey. *Robonation's Annu 21th RoboSub Compet J*. Published online 2018:1-5.

Authors Introduction

Dr. Raji Alahmad

He is currently a researcher at the Department of Human Intelligence Systems of Kyushu Institute of Technology, Japan. He obtained his MEng degree in 2018 and his DEng in 2021 at Kyushu Institute of Technology. His research interests are AI algorithms, logistics, and underwater communication.

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 area is in the field of robotics, its application, and data processing.

Prof. Kazuo Ishii

He is currently a Professor at the Department of Human Intelligence Systems of Kyushu Institute of Technology, Japan. He obtained his M. S. degree in 1993 and his D. Eng. degree in 1996 at The University of Tokyo. His research interests are in the fields of Underwater Robotics, Field Robotics, Neural Networks, and Intelligent Systems.

Prof. Yuki Fukumoto

He is currently a Professor at the Department of Electrical Engineering and Electronics of Kyushu Institute of Technology, Japan. He obtained his Master's degree in 1988 at Kyoto Institute of Technology, his D. Eng. in 2001 at Okayama University, and his MBA in 2003 at McGill University, Canada. His research interests are Electromagnetic compatibility, Manufacturing Technology, and Electron device and electronic equipment.

Dr. Tohlu Matsushima

He is currently an Associate Professor at the Department of Electrical Engineering and Electronics of Kyushu Institute of Technology, Japan. He obtained his Master's degree in 2006 and his D. Eng. in 2009 at Okayama University. His research interests are EMC (environmental electromagnetic engineering, electromagnetic compatibility), and power electronics

Control strategy to change the locomotion mode of a reconfigurable wheel/track robot based on the soil conditions

Supaphon Kamon
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology Kitakyushu 808-196, Japan
Email: kamon.supaphon212@mail.kyutech.jp

Enrico di Maria
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology Kitakyushu 808-196, Japan
Email: e-dimaria@brain.kyutech.ac.jp

Kazuo Ishii
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology Kitakyushu 808-196, Japan
Email: ishii@brain.kyutech.ac.jp

Abstract

The use of agricultural machinery damages the soil by compaction and distortion, where the compaction is due especially to vehicle with small contact areas like wheels, while the distortion is caused especially by tracked systems. In this work we discuss a wheel/track reconfigurable robot able to adjust the contact area based on the soil conditions, to minimize soil damage, energy consumption and adapt travers ability. After giving an overview, we propose a control strategy for switching between the locomotion modes.

Keywords: Reconfigurable robot, Tracked robot, Wheeled robot, Agricultural robot

1. Introduction

Soil degradation is one of the main issue of using agricultural machinery, [1]. During their operations in the field, the vehicles damage the soil by compaction and distortion. The first type of damage occurs in both topsoil (0-50 cm) and subsoil, and determines a soil volume reduction by compressing the pore space. The occlusion of the pores makes it difficult for the water to infiltrate and for the plants to reach the nutrients by the roots. The second type of damage occurs at the topsoil because of the shear loads and it destroys the pores by shear deformation [2]. Once the soil is damaged, the farmers have to restore the original status, and this is an expensive process.

Wheels and tracks are the most common running gears used for the locomotion in the agricultural fields. The two systems are different in terms of soil traversability and soil damage. With a smaller contact area, the wheels can easily sink on a soft soil, while the larger area provided by the track can allow a better floatation. For this reason, wheeled vehicles can cause higher compaction than a tracked system comparable in size. At the same time, it was observed experimentally [3] as the tracks cause a higher distortion, because of the higher shear forces and the peaks of pressure under the rollers. In terms of performance, in [4] a comparison based on the Bekker model was made between small tracked and wheeled vehicles moving on cohesive and frictional soils. The result was that a tracked vehicle outperforms a vehicle on cohesive soils, where increasing the area improves not only the floatation but also increases the maximum drawbar pull. On frictional soil instead, the performances were comparable, especially if the wheels' size is large enough to not sink too much and ensure a low rolling resistance.

Moreover, on loose frictional soils, the higher normal stress exerted by a wheel, because of the smaller contact area, leads to a higher soil strength and higher draught.

A last aspect to consider is that, while steering (especially in skid steering), a shorter contact area sliding at the ground (as for wheels) can determine a reduction of the torque, and then the energy required.

In order to mitigate the problem of the soil damage and adjust the performance of the vehicle based on the soil conditions, we introduce our reconfigurable vehicle Hadrian as in Fig. 1. This system can adjust the contact area as conceptualized in [5], to vary the pressure at the contact with the ground [6]. In this paper we provide an overview of this reconfigurable locomotion system. The target of the proposed vehicle is to transport the grape in the vineyards during the harvesting.

After describing the vehicle, we will conceptualize a possible strategy for making the robot change its locomotion system. Finally, we will describe the control interface created to operate the vehicle.

2. Reconfigurable vehicle overview

The system we developed consists of a mobile base able to pass from a half-track configuration to a wheeled configuration. The front axle of the vehicle consists of two tracks. The sprockets are rigidly connected to the chassis, while each idler is connected by a shock absorber to a Scott-Russell mechanism anchored at the top of the chassis. This mechanism is driven by an electro-hydraulic actuator, and it can lift the two idlers. When the idlers are lifted, only the sprockets remain in contact with the ground, as if they were wheels. The reconfigurable system is described in Fig.2 and Fig. 3. The rear axle consists of two wheels, which are supporting the chassis through a trailing-arm suspension system. The vehicle is equipped with four traction motors. The disposition of the motors is in-wheel for the rear axle, while the front axle sees the motors positioned inside the chassis and connected to the sprockets' shafts by a chain. The battery pack is positioned on the rear of the vehicle. The total weight is 210 kg, with a carrying capacity of 200 kg.

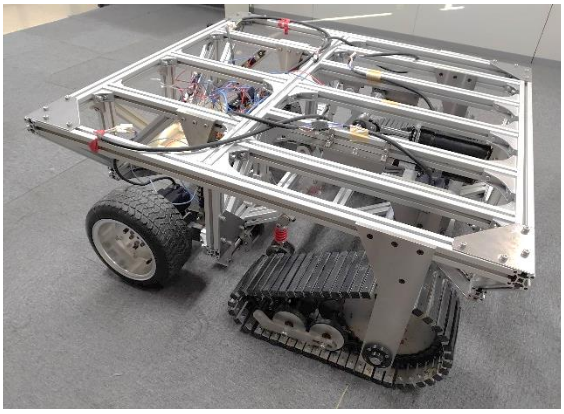


Fig. 1. Reconfigurable track/wheel robot Hadrian

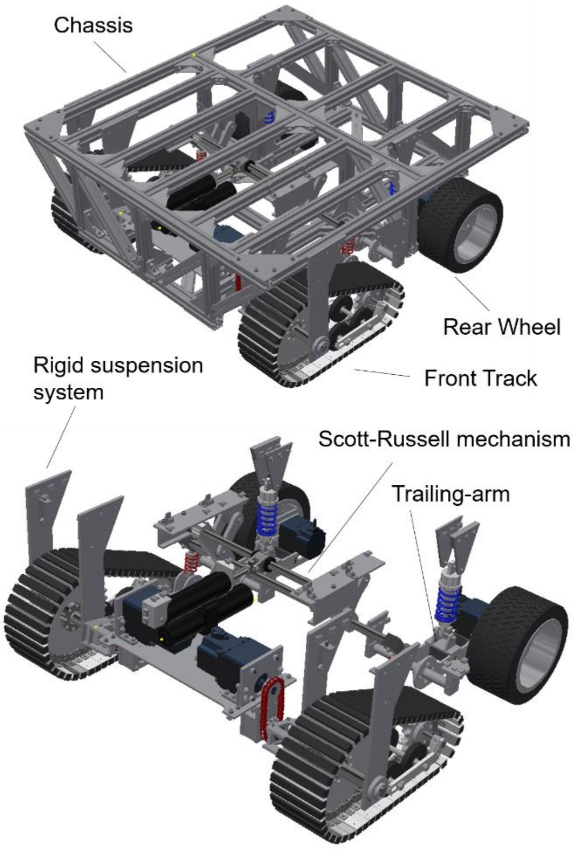


Fig. 2. Overview of Hadrian

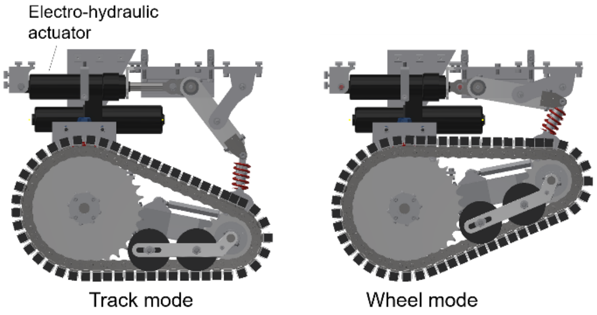


Fig. 3: Track mode (left) and wheel mode (right)

3. Soil damage and measurements for the experiments

The choice of the locomotion system by the vehicle is a difficult task, as various aspects must be taken into account, such as the state of the terrain, the payload of the vehicle, the damage caused and the performance in terms of rolling resistance (energy consumption) and traction.

The first approach to the problem is to test the vehicle on agricultural soil in several conditions, and compare the results to obtain the first thresholds to be used as a reference for the switching between the two locomotion modes.

The decrease in volume following the passage of a vehicle, and the consequent compaction can be measured by means of a device known as the Cone Penetrometer, [7]. This consists of a cone at the end of a rod, which is in turn connected to a load cell. By driving the cone into the ground, the load cell measures the pressure value on the cone and provides the so called Cone Index [MPa]. This is a measure of the state of soil compaction, where higher values indicate a greater state of soil compaction.

During our experiments, we will measure the state of soil compactness by using the Cone penetrometer.

Soil distortion is expressed as the horizontal displacement that particles of soil undergo due to the shear forces applied by the running gear. For a tracked vehicle, the shear displacement is expressed as in Eq. (1)

$$J = ix \tag{1}$$

where i is the slip and x is the position of the soil particle along the contact length (CL) in the moving direction. x can vary from 0 to the maximum length. The slip can be calculated as in Eq. (2)

$$i = \frac{\omega r - V}{\omega r} \tag{2}$$

where V is velocity of robot. ω is angular velocity of running gear or motor. For a wheeled system, the shear displacement can be calculated as Eq. (3)

$$J = r[(\theta_0 - \theta) - (1 - i)(\sin\theta_0 - \sin\theta)] \tag{3}$$

where angle θ is the rim angle where the wheel contact with terrain and angle θ_0 is the entry angle that defines the angle where a point on the wheel contact with the terrain as shown in Fig 4.

In our experiments, we will evaluate the shear displacement by using markers on the soil and measuring the markers displacement after the vehicle pass. As for the slip, we can measure it by knowing the vehicle velocity and running gear rotational velocity.

Finally, we will measure the energy consumption by monitoring the current used.

We are going to evaluate soil damage (compaction and distortion), energy consumption, rolling resistance and traction force under several soil conditions. We will consider firm and soft soil with different levels of soil moisture and compactness. The soil moisture will be monitored by using several soil moisture sensors in wireless communication with the robot.

Because the evaluation of the thresholds requires many experiments, we need a way to easily control the robot remotely. For this reason, we develop a user friendly interface for controlling the robot and switching locomotion mode, while also monitoring the torque, the velocity and the current used by the motors. This control interface is described in Section 5

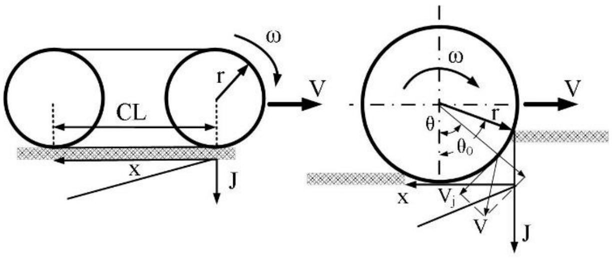


Fig. 4. Development of shear displacement of track and wheel

4. Control Strategy

In this section we propose a preliminary control strategy for switching locomotion mode based on the soil conditions. Decision making process is used to analyze and classify soil conditions. Fuzzy logic [8] and neural network [9] are the most popular decision making process to classify soil and terrain. We can classify the soil based on compactness and moisture level. For that reason, we consider four classes of: firm soil (usual condition), soft soil and soft dry soil (after plowing it), and wet saturated soil (after some heavy rain). For the firm and soft soil, we consider different level of moisture contents, since the moisture affects the soil strength. A soft soil and a soft dry soil have different moisture. A soft soil is more humid, while a soft dry soil being under the sun for long becomes dry. Four features are defined as soil moisture, energy, slip and velocity. Concept of control strategy is shown in Fig. 5. The goal and the expected output of control strategy is described as below.

1) To minimize soil damage. If the robot is sinking too much, robot is damaging too much the soil and causing compaction or distortion. It doesn't matter with a track or a wheel. However, if wheel cannot move, track can help robot to move when wheel stuck. If the robot is not sinking too much, the robot uses wheeled mode, because the track causes higher distortion. Cone penetration test, shear displacement of track as Eq. (1) and wheel as Eq. (3) are used to evaluate soil damages.

2) To minimize energy consumption. The rolling resistance is relative with energy. If robot sink too much, the robot has higher rolling resistance and use higher energy. Track doesn't sink too much if compared with wheel because longer contact area. For this reason, track use lower energy consumption if compared with wheel. However, track use higher energy consumption when robot turn left or right using skid-steering because of longer contact area.

If robot is sinking too much, it will use tracked mode. If robot is not sinking too much, it will use wheeled mode because track cause higher soil distortion. We will test robot on agricultural soil in several conditions and collect data of torque of motor and measure sinkage after robot pass. After we do many times, we can calculate sinkage using torque of motor. The robot uses wheeled mode when it turns left or right because wheel can reduce energy consumption

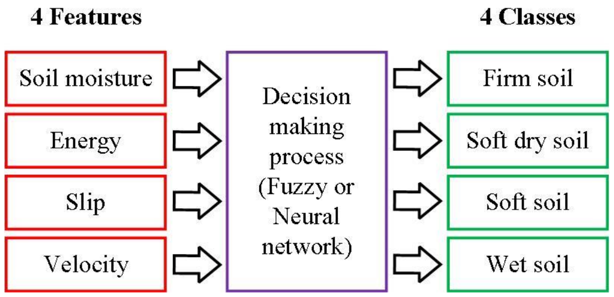


Fig. 5. Classification strategy

5. Control hardware actuator and sensor

After we proposed the control strategy in section 4, we need to consider about sensor, microcontroller and actuator for switching locomotion mode based on the soil conditions. The microcontrollers used in this research are ESP32. ESP 32 is microcontroller that can be transmitter and receiver with wireless. Control diagram of ESP32 Transmitter is shown in Fig. 6. Soil moisture sensor is connected to ESP32 transmitter for measuring the moisture in each soil condition, and send moisture data to ESP32 receiver. Control diagram of ESP32 receiver is shown in Fig. 7. Two Brushless DC Motors for tracks are BLV620K100S (24 VDC, 40 RPM, torque 36.5N.m) and two Brushless DC Motors for wheels are BLV510K100S (24 VDC, 30 RPM, torque 27.4 N.m). Four motors can send data of torque and angular velocity to calculate energy consumption. One electro-hydraulic actuator (EHA) is installed inside robot to switch between tracked and wheeled mode. Relay is used to control direction of EHA. One JRT laser distance sensor (B series) is used to measure distance and calculate velocity of robot. We are planning some simplified experiments and the use of the wall and distance sensor is only a way to easily localize the robot in the field, when this is moving on a prescribed rectilinear path. We can calculate slip using angular velocity from motor and velocity from laser distance sensor as Eq. (2). The robot user interface is developed and tested using Microsoft Visual Basic 2022 as in Fig. 8. The robot user interface allows the user to control the robot for effective operation and get feedback data from the robot. The robot user interface offers users many advantages. Users can easily and quickly operate the robot. Even farmers with little technical knowledge can control the robot with the robot user interface. For the experiment, the robot can be remotely operated, but it is not autonomous path planning. We discuss how to use the robot user interface as below

1) Manual control can be operated independently by pushing arrows ↑forward,↓backward, ←left , →right.

2) Automatic control can be operated by pushing blue button in the top center of screen. The robot will move forward, and use our control strategy to switch between tracked mode and wheeled mode autonomously.

3) Direction of EHA can be shortened and extended by pushing the pull and the push button.

4) Direction of four motors control can be controlled by pushing clockwise and counterclockwise button.

5) Data of angular velocity and load torque of four motors are shown as graphs in the center of screen. Distance, velocity and soil moisture are shown as numerical value on bottom right of screen.

The robot user interface is developed to control and obtain the feedback data easily for control strategy as described in section 4. For the future application, farmers who have little technical knowledge can control the robot remotely and operate the autonomous path planning of robot during the harvesting period with our robot user interface.

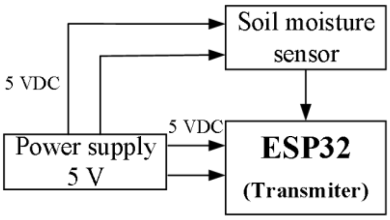

Soil moisture sensor
5 VDC
5 VDC
Power supply 5 V
ESP32 (Transmiter)

Fig. 6. Control diagram of ESP32 Transmitter

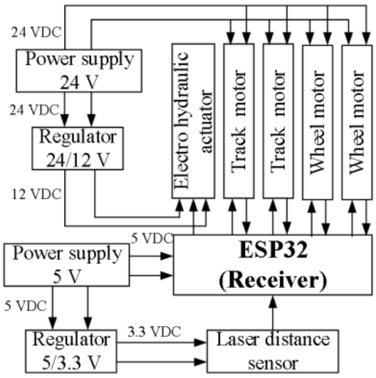
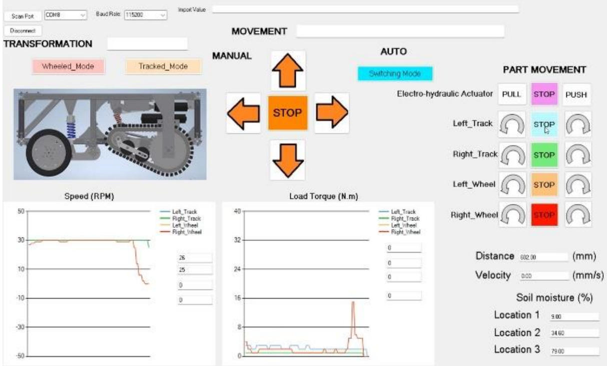

24 VDC
Power supply 24 V
24 VDC
Regulator 24/12 V
12 VDC
Electro hydraulic actuator
Track motor
Track motor
Wheel motor
Wheel motor
5 VDC
Power supply 5 V
ESP32 (Receiver)
5 VDC
Regulator 5/3.3 V
3.3 VDC
Laser distance sensor

Fig. 7. Control diagram of ESP32 Receiver


Scan Port
Baud Rate
Input Value
Disconnect
TRANSFORMATION
Wheeled_Mode
Tracked_Mode
MOVEMENT
MANUAL
STOP
AUTO
Switching Mode
PART MOVEMENT
Electro-hydraulic Actuator
PULL
STOP
PUSH
Left_Track
STOP
Right_Track
STOP
Left_Wheel
STOP
Right_Wheel
STOP
Speed (RPM)
50
30
10
-10
-30
-50
Left_Track
Right_Track
Left_Wheel
Right_Wheel
Load Torque (N.m)
40
32
24
16
8
0
Left_Track
Right_Track
Left_Wheel
Right_Wheel
Soil moisture (%)

Scan Port: COM8
Baud Rate: 115200

26
25
0
0

0
0
0
0

Distance 602.00 (mm)
Velocity 0.00 (mm/s)
Location 1 9.00
Location 2 34.60
Location 3 79.00

Fig. 8. Robot user interface

The robot can switch between tracked and wheeled mode using electro hydraulic actuator (EHA). EHA extends to push track to the ground in tracked mode. Video of switching from wheeled to tracked mode is shown in Fig. 9. EHA shortens to pull and lift the partial track off the ground in wheeled mode. Video of switching mode from tracked to wheeled mode is shown in Fig. 10. After we complete software, we will get data of four features to classify soil conditions for autonomous switching modes. In future work, we will test the robot on various soil conditions as shown in. Fig. 11 experiment setup. Soil moisture sensors and ESP32 transmitters are putted on each soil conditions to measure moisture and send data to ESP32 receiver inside robot. The laser distance sensor is installed on the front of robot to detect wall for estimation position and calculation velocity. GPS navigation will be used to estimate the position in next work. For experiment, the robot will move forward in tracked or wheeled mode on soil condition 1, then change to another mode on soil condition 2 autonomously using our control strategy that is described in section 4.

6. Conclusion and future work

A reconfigurable wheel/track robot is built to minimize soil damage and energy consumption using the advantages of both track and wheel running gear. The electrical system, robot user interface and control strategy to change the locomotion mode based on the soil conditions are presented in this research. In future plan, we will collect data of four features and use fuzzy logic or neural network to classify the four soil conditions for

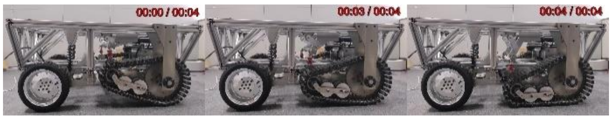


Fig. 9. Switching from wheeled to tracked mode

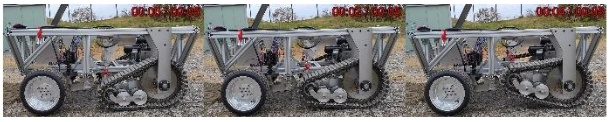


Fig. 10. Switching from tracked to wheeled mode

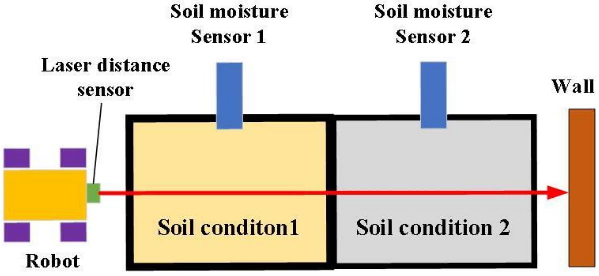


Fig. 11. Experiment setup

autonomous switching modes to minimize soil damage and energy consumption

Acknowledgment

This work was supported by Kyushu Institute of Technology. We thank our friend Raji Alahmad and Tanaka Yoshiki for supporting the robot assembling process.

References

1. W.C.Tim Chamen, A.P. Moxey, W. Towers, B. Balana, P.D. Hallett, "Mitigating arable soil compaction: A review and analysis of available cost and benefit data", Soil Tillage Research 146 (PA), pp. 10–25, 2015.
2. R. Horn, "Stress - Strain effects in structured unsaturated soils on coupled mechanical and hydraulic processes", Geoderma 116(1–2), pp. 77–88, 2003.
3. M. Lamandé et al., "Risk assessment of soil compaction in Europe: Rubber tracks or wheels on machinery", Catena, vol. 167, pp. 353-362, 2018.
4. G. R. Gerhart, "The Bekker model analysis for small robot vehicles", US Army TACOM, 2004.
5. E. di Maria. "Running gear size selection of a wheel/track reconfigurable grape transporting vehicle by FEM analysis", Doctoral dissertation. Kyushu Institute of Technology, 2021.
6. A. Grazioso et al. "Multibody Modeling of a New Wheel/Track Reconfigurable Locomotion System for a Small Farming Vehicle", Machines, 2022.
7. J.Y.Wong, "Theory of ground vehicle", Published simultaneously in Canada, 2008.
8. S. Khaleghian and S. Taheri, " Terrain classification using intelligent tire", Journal of Terramechanics 71, pp.15-24, 2017.
9. E. G. Collins et al., "Vibration-Based Terrain Classification Using Surface Profile Input Frequency Responses", IEEE International Conference on Robotics and Automation Pasadena, CA, USA, May 19-23, 2008

Authors Introduction

Mr. Supaphon Kamon

He received his Master's degree from Department of Information Science and Engineering, Ritsumeikan University, Japan in 2015. He is currently a Doctoral course student in Kyushu Institute of Technology, Japan.

Dr. Enrico di Maria

He received his Master Degree in Mechanical Engineering at University of Salento, Department of Innovation Engineering, in 2016. He completed the Doctor Course at the Kyushu Institute of Technology in 2021, and he is currently a postdoctoral researcher at KIT. His research interests are robotics and terramechanics.

Prof. Kazuo Ishii

He received the PhD degree from the Department of Naval Architecture and Ocean Engineering, University of Tokyo, Tokyo, Japan, in 1996. He is currently a Professor with the Department of Human Intelligence Systems, and the Director of the Center for Socio-Robotic Synthesis, Kyushu Institute of Technology, Kitakyushu, Japan. His research interests include underwater robots, agricultural robots, roboCup soccer robots and intelligent systems.

Sea-floor Image Restoration with Variable Absorbance Coefficient

Irmiya R. Inniyaka
Department of Life Sciences and Systems Engineering, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka 808-0196, Japan

Yuya Nishida
Department of Human Intelligent System, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka 808-0196, Japan

Kazuo Ishii
Department of Human Intelligent System, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka 808-0196, Japan
E-mail: irmiya.inniyaka-reuben350@mail.kyutech.jp, y-nishida@brain.kyutech.ac.jp, ishii@brain.kyutech.ac.jp
https://www.kyutech.ac.jp

Abstract

Light scattering from suspended particles and attenuation over the water column are constant phenomena that reduce underwater image quality due to low contrast and color distortion. Common image restoration techniques such as the image formation model, assumes a constant attenuation coefficient across color channels. This results in a restoration solution with limited application. We propose a method of image restoration that considers the wavelength-dependent attenuation of underwater images by providing concurrent measurements of the coefficient of attenuation for each color channel. In this paper, a description of the design of a turbidity meter is made. It used to extract absorbance of light in the RGB channel. We conduct experiments to validate the operation of the turbidity meter based on Beer-Lamberts law.

Keywords: Underwater image quality, image restoration, Turbidity meter, Coefficient of attenuation

1. Introduction

Underwater imaging is a critical aspect of perception for marine robotic platforms and by extension it influences many underwater scientific missions like Sea life monitoring, population census, exploration, or inspection tasks. Unlike terrestrial imaging, acquiring clear images underwater is a challenge that negatively affects the performance of robotic systems [1].

Images captured underwater suffer from contrast degradation and color distortion mainly because of interaction of light travel in water. When a ray of light travels from the source to the scene, it is affected in 2 ways. First, some of the rays are “absorbed” by the medium thereby reducing its intensity. Secondly, the direction of propagation may be changed when it collides with suspended particles in a phenomenon called scattering. This results in deflections and backscattering that deteriorates the image contrast while the color is distorted due to the wavelength dependent attenuation of light as it travels the water column [2], [3].

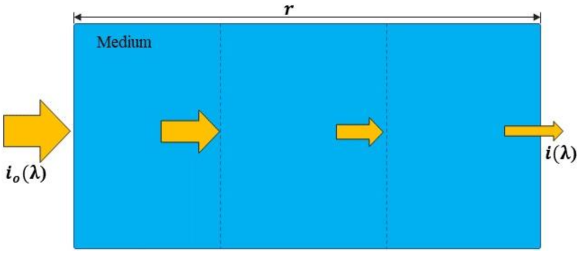


Fig. 1 Attenuation of light in Water medium.

Over the years several methods have been used to correct and enhance distorted underwater images. These methods are generally sensitive to the location of application and often require tuning for different imaging conditions or only provide reliable performance for short range images. The image formation model is also a common method for restoring hazy images. However, its application to underwater images assumes a constant coefficient of attenuation for the image RGB channels. This makes its performance inadequate because attenuation parameters vary with season, climate and geography [4]. In this paper, we describe the design of a simple turbidity meter for real-time measurement of turbidity as a means of applying wavelength-dependent attenuation coefficient to restore degraded images. The paper is organized into the following sections. In Section 2, the design principle and structure of the turbidity meter is described. Section 3 presents the hardware structure and Section 4 describes the experimental setup and result discussion. The conclusion and future task to be conducted are presented in Section 5.

2. Design Principle

2.1. *Underwater Image formation model*

Based on the image formation model, for each color channel $\lambda \in \{\boldsymbol{R}, \boldsymbol{G}, \boldsymbol{B}\}$, the image intensity at each pixel comprises of 2 components, the attenuated signal and background light as shown in Eq. (1):

$$I_\lambda(x) = J_\lambda(x) \cdot t_\lambda(x) + (1 - t_\lambda(x)) \cdot A_\lambda(x) \quad (1)$$

where $\boldsymbol{I}_\lambda$ is distorted image captured by the optical sensor, $\boldsymbol{x}$ is the pixel coordinate, $\boldsymbol{J}_\lambda$ is the object radiance to be restored, $\boldsymbol{t_c}$ is the transmission in the color channel and $\boldsymbol{A}_\lambda$ is the global background light.

The transmission Eq. (2) depends on the distance of objects $\boldsymbol{z}(\boldsymbol{x})$ in the scene and attenuation coefficient $\boldsymbol{\beta}_\lambda$ for each color channel:

$$t_\lambda = exp^{(-\beta_\lambda z(x))}. \quad (2)$$

As opposed to the common assumption of a fixed attenuation coefficient for color channels, we propose to influence the attenuation coefficient $\boldsymbol{\beta}_\lambda$ by varying it with respect to the color channel from measurements of water turbidity. The total attenuation coefficient $\boldsymbol{\beta}_\lambda$ describes how much light of a particular wavelength is attenuated as it travels through a medium as postulated by the Beer-Lambert law which states that "the intensity of light decreases exponentially with distance. The formulation is as expressed below:

$$\boldsymbol{A} = -\log\left(\frac{\boldsymbol{i}}{\boldsymbol{i_0}}\right) \quad (3)$$

$$\boldsymbol{\beta} = -\frac{\boldsymbol{1}}{\boldsymbol{r}}\log\left(\frac{\boldsymbol{i}}{\boldsymbol{i_0}}\right) \quad (4)$$

From Eq. (3), $\boldsymbol{A}$ is the absorbance of light traveling a certain distance from the source $\boldsymbol{i}$ to the scene $\boldsymbol{i_0}$. The equation is rearranged to Eq. (4) where the distance of light travel by $\boldsymbol{r}$ to calculate the attenuation coefficient $\boldsymbol{\beta}$. This is as described in Figure 1.

2.2. *Turbidity meter design principle*

Turbidity is the measure of the degree of suspended particles (turbidity) in water. To achieve this, we design a simple turbidity meter to continually determines the attenuation coefficient of light based on the density of suspended particles in the water. The design solution is based on transmitted light as stipulated by JIS K0101 "Testing Methods for Industrial Water" described in Figure 2 [5]. Basically, the light transmitted from the light source positioned at one side of the water medium is measured at the opposite side by a light sensor.

For our application in the restoration of underwater images, the turbidity meter is designed to satisfy the following requirements.

i. A modular configuration for ease to mount on an underwater vehicle.
ii. Measure turbidity in the 3 color channels (RGB).
iii. Inhibit ambient illumination.
iv. Continuous data acquisition and transfer.
v. Permit water flow to capture variation in turbidity.

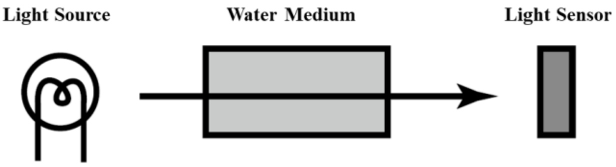


Fig. 2 Turbidity Measurement based on Transmitted Light

3. Hardware structure of turbidity meter

The turbidity meter structure has 3 main parts: LED hull, ambient light inhibitor and sensor hull as shown in Figure 3.

It has a cylindrical structure with dimension of 480x130[mm] and 3 LEDs for Red, Green, and Blue light sources as well as 3 light sensors, for each light source. The ambient light inhibitor is a 3D printed polylactic acid (PLA) model. It has several functions including blocking out external light, restricting each light source to its light sensor, provide a path for water flow, ensuring a constant distance of transmitted light, and provide mounting for the LED hull and Sensor hull. The LED and sensor hulls are 3-inch acrylic hulls facing each other at 100mm. This is the optical path length as recommended by JIS K0101.

The turbidity meter communication is achieved via I^{2}C between the sensors and Arduino-nano while an Xport module transfers the data to PC via an ethernet interface. Table 1 shows the specification of the turbidity meter and Figure 4 shows its cross-sectional view and Figure 5 is an overview of the electrical and communication system of the turbidity meter.

Table 1: Turbidity Meter Component Specification

Category	Specification
Weight in air	2.7kg
Illuminance Sensor	TSL2561 (0.1 – 40,000 lux)
Power LED	RGBLED-OSTCWBTHC1S
Power Supply	5V
Current	18mA
Communication	I^{2}C Arduino Nano XPort

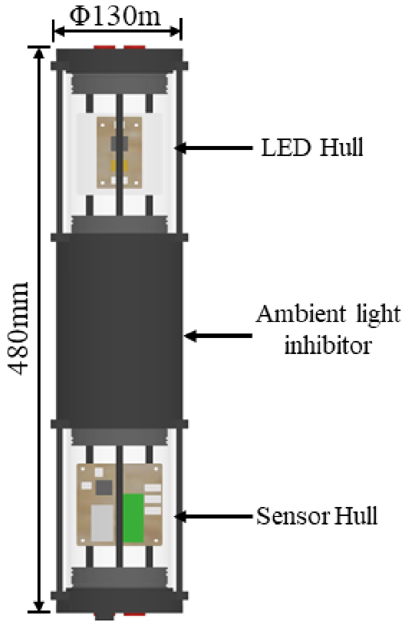

Φ130m
480mm
LED Hull
Ambient light inhibitor
Sensor Hull

Fig. 3 Turbidity Meter

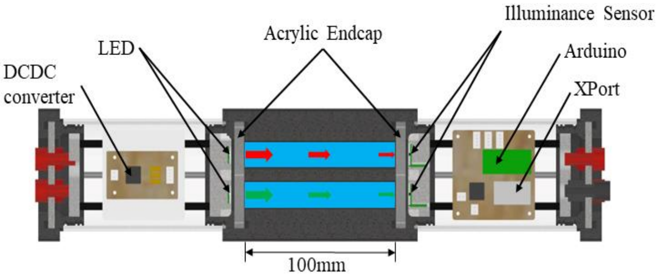

DCDC converter
LED
Acrylic Endcap
Illuminance Sensor
Arduino
XPort
100mm

Fig. 4 Cross-sectional view of Turbidity meter

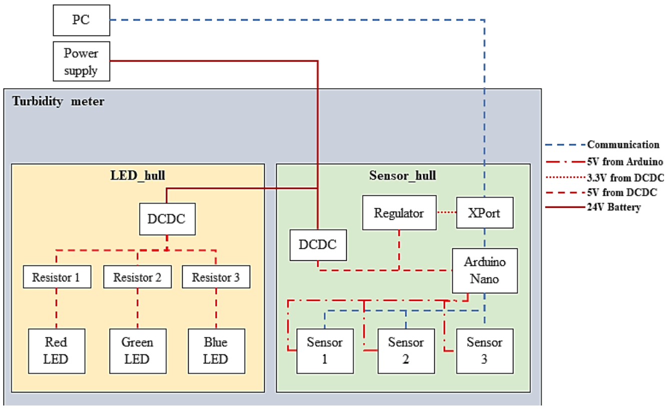

PC
Power supply
Turbidity meter
LED_hull
DCDC
Resistor 1
Resistor 2
Resistor 3
Red LED
Green LED
Blue LED
Sensor_hull
DCDC
Regulator
XPort
Arduino Nano
Sensor 1
Sensor 2
Sensor 3
Communication
5V from Arduino
3.3V from DCDC
5V from DCDC
24V Battery

Fig. 5 Overview of system diagram

4. Experiments and Results

The formazin standard and kaolin standard are the standard solutions used for evaluation of any turbidity measurement device. In this work, the kaolin standard solution is used because it is used in the JIS standard. It is a type of clay that is low-cost, can be obtained easily, and is harmless.

4.1. *Experimental procedure*

We construct an apparatus (Figure 6) to assess how the light is absorbed in turbid water (kaolin solution). The set up consists of a 3-inch cylindrical PVC pipe that prevents ambient light from penetration. The Sensor hull is mounted at the lower part of the pipe using a jig to make the attachment rigid and watertight while the LED hull is mounted at the top side. Using the formulation in Eqn. 5, we calculated volume of water needed to vary the water level to simulate different water depth of 5, 10, 20, 30, 40, 50 and 60 [cm] and measure effects of attenuation.

$$Vol_{water} = Height_{cylinder} * \pi * radius_{cylinder}^{2} \qquad (5)$$

For the kaolin solution, the test solutions are prepared with the following concentration of kaolin in water: 0, 10, 30, 100, 200, and 300 [mg/L]. The absorption coefficients for the 3 light sources are independently measured per concentration of the test solution to determine the wavelength-dependent attenuation coefficient for the various water depth.

4.2. *Results*

In Figure 7, Figure 8, and Figure 9 the result of the relationship between optical path and light intensity for Red, Green and Blue LED are shown respectively. we observe that the minimum coefficient of determination R^2 is 0.95, 0.98, and 0.97 for the Red, Green, and Blue light intensities. These high values validate the Beer-Lambert law which implies that for a distance $\boldsymbol{r}$ (Figure 1) from between the light source and the light sensor, we can obtain the coefficient of absorption which can be calculated from Eq. (4).

Having validated the high correlation between the optical path and light intensity in kaolin test solution, we conduct an experiment to determine the relationship between turbidity and the coefficient of absorption using the turbidity meter.

For this experiment, Kaolin test solution is incrementally dissolved in the water and data log of light reaching each light sensor is recorded. The concentrations of Kaolin are 0, 10, 100, 200, 300, and 400 mg/L.

The result of the data is plotted in Figure 10. The kaolin concentration at 0mg/L is pure tap water. The relationship is linear for the 3 light sources which validates Eq. (4) being that the absorbance coefficient has a proportional relationship to the concentration of kaolin with a strong coefficient of determination R^2 of 0.99 per light source.

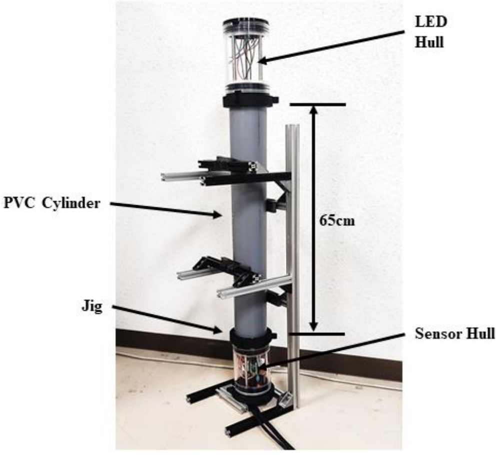


Fig. 6 Experimental Apparatus

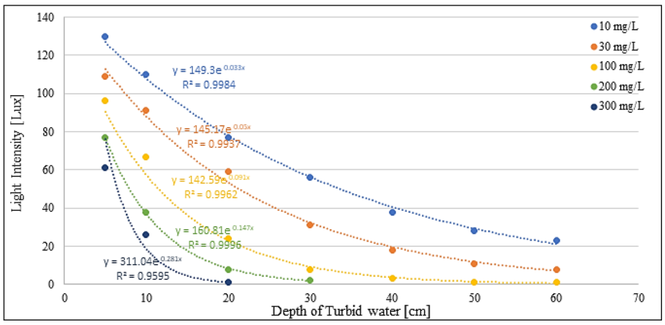


Fig. 7 Relationship between Red LED's optical path and light intensity in kaolin test solution

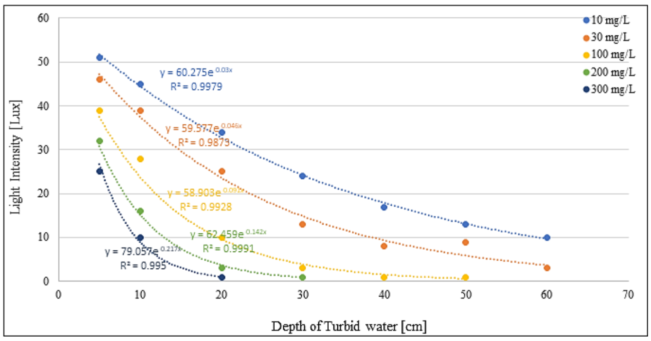


Fig. 8 Relationship between Green LED's optical path and light intensity in kaolin test solution

5. Conclusion

In this paper, we describe the design principle of a turbidity meter based on the transmitted light method and conducted experiments to verify the relationship between the optical path and light intensity in varying concentration of kaolin standard solution. We also validated the Beer-Lambert law by experimenting with the turbidity meter and found a strong correlation between coefficient of absorption and concentration of kaolin.

In the future task, image data will be concurrently collected with turbidity data to provide the attenuation coefficient parameter in Eq. (1) for image restoration. As a final step, we will mount the turbidity meter on an AUV (Tuna Sand2) for data collection in different underwater conditions.

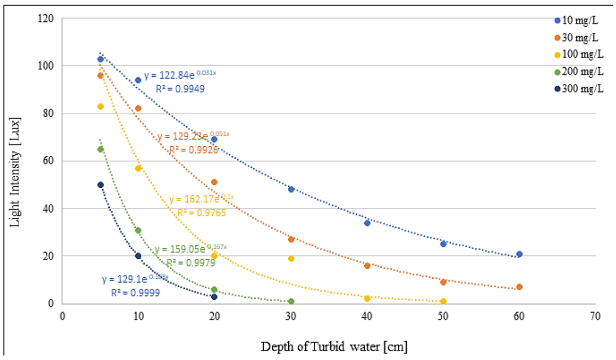


Fig. 9 Relationship between Blue LED's optical path and light intensity in kaolin test solution

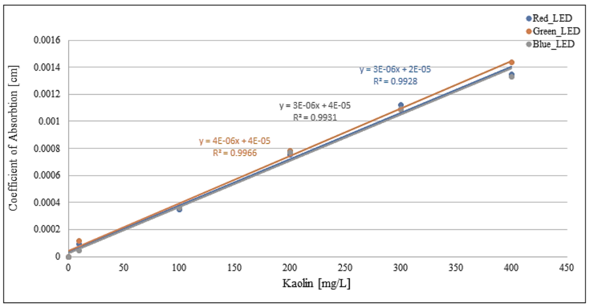


Fig. 10 Relationship between coefficient of light absorption and Kaolin based turbidity

Acknowledgement

This paper is supported by JSPS KAKENHI Grant Number JP18H01643.

References

1. H. Blasinski and J. Farrell, "A three parameter underwater image formation model," *IS T Int. Symp. Electron. Imaging Sci. Technol.*, pp. 1–8, 2016, doi: 10.2352/ISSN.2470-1173.2016.18.DPMI-252
2. K. A. Skinner, J. Zhang, E. A. Olson, and M. Johnson-Roberson, "UWStereoNet: Unsupervised learning for depth estimation and color correction of underwater stereo imagery," *Proc. - IEEE Int. Conf. Robot. Autom.*, vol. 2019-May, pp. 7947–7954, 2019, doi: 10.1109/ICRA.2019.8794272.
3. J. Y. Chiang and Y. C. Chen, "Underwater image enhancement by wavelength compensation and dehazing," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1756–1769, 2012, doi: 10.1109/TIP.2011.2179666.
4. D. Berman, T. Treibitz, and S. Avidan, "Diving into haze-lines: Color restoration of underwater images," *Br. Mach. Vis. Conf. 2017, BMVC 2017*, pp. 1–12, 2017.
5. https://global.ihs.com/standards.cfm?publisher=JSA

Authors Introduction

Ms. Irmiya R. Inniyaka

He obtained his MSc. degree in Automotive Mechatronics at the Department of Mechatronics, Cranfield University, England in 2016. He is currently pursuing his Doctoral program at Kyushu Institute of Technology, Japan. His research interests are in Image processing and Underwater Robotics.

Dr. Yuya Nishida

He is currently an Associate Professor at the Department of Human Intelligence Systems, Kyushu Institute of Technology, Japan. He obtained his Masters degree in Engineering in 2008 and his D. Eng. degree in 2011 at the same university. His research interests are in the field of Underwater Robotics, Field Robotics, and Intelligent Systems.

Prof. Kazuo Ishii

He is currently a Professor at the Department of Human Intelligence Systems of Kyushu Institute of Technology, Japan. He obtained his M. S. degree in 1993 and his D. Eng. degree in 1996 at The University of Tokyo. His research interests are in the fields of Underwater Robotics, Field Robotics, Neural Networks and Intelligent Systems.

Design of A Parameter Update Method of the Database-Driven PID Controller Considering H_∞ Norm of the System

Takumi Ueda
Kyushu institute of technology, 2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka, 808-0135, Japan

Kazuo Ishii
Kyushu institute of technology, 2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka, 808-0135, Japan

Abstract

PID controllers are used in many applications, and they are considered as typical examples of fixed-structure controllers. For each specific system, the PID parameters of the controller must be adjusted appropriately. Therefore, database-driven control, which automatically adjusts the parameters of the PID controller using a large amount of experimental data, has attracted attention. Database-driven control is one of the methods that does not require model information. Therefore, in the Database-driven PID (DD-PID), the Jacobian of the system used for the steepest descent method is unknown. In this paper, we present an optimization method using the Gradient descent method based on the local model of the system. And we proposed a method that can suppress the increase of H_∞ norm more than the conventional DD-PID method.

Keywords: Database-Driven PID, H_∞ norm, Nonlinear system

1. Introduction

PID controllers are widely used in industry [1] and are one of the typical examples of controllers with a fixed structure. To adjust the parameters appropriately, it is effective to identify a mathematical model of the system, and many methods have been reported so far [2]. However, the systems in general use have various nonlinearities, making it difficult to construct a detailed model, and the desired control performance may not be achieved with a conventional fixed PID controller. Therefore, database-driven control, in which PID controller parameters are automatically adjusted according to system characteristics using a large amount of experimental data, has been actively studied. [3],[4],[5],[6]

The DD-PID method, one of the database-driven control methods, constructs a local model based on data accumulated in a database and updates the PID gains using a nonlinear optimization method. The conventional method uses the just-in-time (JIT) method to construct the local model, which calculates the local model by extracting data from the data stored in the database that are similar to the query created using the current operation data as the nearest neighbor data. Appropriate selection of neighboring data and design of the database and query is important to construct a local model using the JIT method, and they also affect the control results.

In this paper, we examine a database design method for the JIT method, which identifies the local model of the system using the Recursive Least Squares (RLS) method and reflects the H_∞ norm of the calculated local model in the database. Furthermore, the Jacobian of the local model identified by RLS is reflected in the gradient method to optimize the PID gain update. Finally, the proposed method is verified by numerical simulations.

2. Proposed method

The DD-PID method proposed in this section updates the PID gains by the following steps.

Step.1 Setting of the reference model and experiment using initial PID gains

The DD-PID method cannot update parameters when there is no data in the database. Therefore, an initial database is constructed using a stable controller. The database used in this paper is defined by equation (1), where $j = 1, \cdots, N_d(0)$ and $N_d(0)$ is the number of databases in the initial state. In addition, $\|M(j)\|_\infty$ is the H_∞ norm of the local model calculated using the RLS method described below, ϕ_d is the information vector, and θ_d is the gain vector of the controller.

$$\boldsymbol{\Phi}(j) \coloneqq [\boldsymbol{\phi}(j), \|M(j)\|_\infty, \boldsymbol{\theta}_d(j)] \tag{1}$$

Define ϕ_d and θ_d by equations (2) and (3). where N_y and N_u are the orders of y and u.

$$\boldsymbol{\phi}_d := [y(j), \cdots, y(j - N_y), u(j), \cdots, u(j - N_u)] \tag{2}$$

$$\boldsymbol{\theta}_d(j) \coloneqq [K_P(j), K_I(j), K_D(j)] \tag{3}$$

The reference model for the DD-PID method is defined by equations (4) and (5). where μ is a parameter related to attenuation, ρ is a parameter related to rising time, and T_s is the sampling time.

$$y_m(t) = \frac{P(1)z^{-1}}{P(z)} = \frac{(1 + p_1 + p_2)z^{-1}}{1 + p_1 z^{-1} + p_2 z^{-2}} r(t) \tag{4}$$

$$\begin{cases} p_1 = -2e^{-(2\mu)^{-1}\rho} \cos\left((2\mu)^{-1}\rho\sqrt{4\mu - 1}\right) \\ p_2 = e^{-\rho\mu^{-1}} \\ \mu \coloneqq 0.25(1 - \delta) + 0.51\delta \\ \rho \coloneqq \sigma^{-1} T_s \end{cases} \tag{5}$$

Step. 2 Getting queries and neighborhood data

Obtain the query q while operating the control target. q is defined by equations (6) and (7).

$$\boldsymbol{q}(t) \coloneqq [\boldsymbol{\phi}(t), \|M(t)\|_\infty] \tag{6}$$

$$\boldsymbol{\phi}(t) := [y(t), \cdots, y(t - N_y), u(t), \cdots, u(t - N_u)] \tag{7}$$

Where $\|M(t)\|_\infty$ is the H_∞ norm of the local system calculated by the RLS method. the system to be estimated by RLS is defined by equation (8).

$$H_{RLS}(z) = \frac{b_1(t)z^{-1} + b_2(t)z^{-2}}{1 + a_1(t)z^{-1} + a_2(t)z^{-2}} \tag{8}$$

Using the retrieved query, the proximity distance of the data to the information vector in the database is calculated, and neighboring data are selected. Although several algorithms have been proposed for selecting nearby data from a large amount of data, this paper uses Kinoshita's[7] method for obtaining nearby data based on similarity. The similarity between the information vector and the query is given by equations (9) and (10).

$$\boldsymbol{S}(\phi_d(j), q(t)) = \prod_{i=1}^{N_y + N_u + 1} \frac{1}{\sqrt{2\pi h_i^2}} e^{-\frac{\alpha^2}{2}} \tag{9}$$

$$\alpha = h_i^{-1}\left(q_i(t) - \phi_{d,i}(j)\right) \tag{10}$$

where h_i is the bandwidth of the probability density function, $q_i(t)$ is the i-th element of the query at time t, and $\phi_{d,i}(j)$ is the i-th element of the j-th information vector in the database. The selection of neighboring data is based on similarity according to equation (11). The designer sets $T_{th} (0 \leq T_{th} \leq 1)$. For example, if $T_{th} = 0.95$, data that match 95% of the queries are selected from the database.

$$\boldsymbol{S}(\boldsymbol{\phi}_d(j), \boldsymbol{q}(t)) \geq T_{th} \prod_{i=1}^{N_y + N_u + 1} \frac{1}{\sqrt{2\pi h_i^2}} \tag{11}$$

The results of neighborhood data selection are also used to update the database.

Step. 3 Local model calculation by JIT method

Using the neighborhood data selected in Step.2, construct a local model using Equations (12) and (13) to obtain the PID gain $\tilde{\theta}$.

$$\tilde{\boldsymbol{\theta}}(t) = \sum_{i=1}^{n_k} \boldsymbol{\omega}_i \boldsymbol{\theta}(i), \quad \sum_{i=1}^{n_k} \boldsymbol{\omega}_i = 1 \tag{12}$$

$$\boldsymbol{\omega}_i = \frac{\boldsymbol{S}(\boldsymbol{\phi}_d(j), \boldsymbol{q}(t))}{\sum_{j=1}^{n_k} \boldsymbol{S}(\boldsymbol{\phi}_d(j), \boldsymbol{q}(t))} \tag{13}$$

The gain $\tilde{\theta}$ calculated by the JIT method may be inappropriate. Therefore, the gain $\tilde{\theta}$ calculated in step 3 is corrected using the gradient method. The gain correction by the gradient method is given by Equation (14). As shown in equation (14), the gain correction by the DD-PID method is delayed by one step.

$$\boldsymbol{\theta}(t) = \boldsymbol{\theta}(t) - \boldsymbol{\eta} \frac{\partial J(t+1)}{\partial \tilde{\boldsymbol{\theta}}(t)} \tag{14}$$

Where η is the learning coefficient for the gradient method, J is the evaluation function, and ε is the difference between the system output and the reference model output, defined by equations (15) - (17), respectively.

$$\boldsymbol{\eta} \coloneqq [\eta_P, \eta_I, \eta_D] \tag{15}$$

$$J(t) \coloneqq 0.5\varepsilon(t)^2 \tag{16}$$

$$\varepsilon(t) = y(t) - y_m(t) \tag{17}$$

From these relationships, the partial derivative in equation (14) can be expanded using the differential chain rule as in equation (18).

$$\frac{\partial J(t+1)}{\partial\, \tilde{\boldsymbol{\theta}}(t)} = \frac{\partial J(t+1)}{\partial \varepsilon(t+1)} \frac{\partial \varepsilon(t+1)}{\partial y(t+1)} \frac{\partial y(t+1)}{\partial u(t)} \frac{\partial u(t)}{\partial\, \tilde{\boldsymbol{\theta}}(t)} \tag{18}$$

The Jacobian of the system is given by equation (19) from equation (8).

$$\frac{\partial y(t+1)}{\partial u(t)} = b_1(t) \tag{19}$$

Step. 4 Update the database

Since the DD-PID method updates gains based on data stored in the database, the amount of data stored in the database should be kept to a minimum. Therefore, the query and PID gains are added to the database only when no neighboring data are selected. This reduces the amount of similar data and thus is expected to reduce the amount of computation.

Step. 5 Evaluation and Convergence Decision

Update the controller parameters by repeating Step.2 through Step.5 until the evaluation function shown in equation (20) becomes sufficiently small or is repeated a specified number of times.

$$\boldsymbol{J}(epoch) = \sqrt{\frac{1}{M}\sum_{i=1}^{M}\left(y_0(i) - y_r(i)\right)^2} \qquad (20)$$

3. *Simulation*

3.1. *System parameter settings*

In this paper, the effectiveness of the proposed method is verified by numerical simulations using Hammerstein's nonlinear model shown in Equations (21) and (22). [8]

$$\begin{aligned} y(t) &= 0.6y(t-1) - 0.1y(t-2) \\ &\quad +1.2x(t-1) - 0.1x(t-2) + \xi(t) \end{aligned} \qquad (21)$$

$$x(t) = \begin{cases} 1.0u - 1.5u^2 + 1.0u^3 (t < 70) \\ 1.0u - 1.0u^2 + 1.0u^3 (t \geq 70) \end{cases} \qquad (22)$$

Target values were set as in equation (23).

$$r(t) = \begin{cases} 0.5(\quad 0 < t \leq \ 50) \\ \quad 1(\ 50 < t \leq 100) \\ \quad 2(100 < t \leq 150) \\ 1.5(150 < t \leq 200) \end{cases} \qquad (23)$$

The reference model is designed as in equation (24), where $\delta = 0$ and $\sigma = 10$.

$$P(z) = 1 - 0.271z^{-1} + 0.0183z^{-2} \qquad (24)$$

Other parameters are shown in Table 1.

Table 1. Setting parameters for DD-PID

Order of the information vector	$N_y = 5$ $N_u = 5$
The constant value for RLS	$\alpha = 1000$ $\boldsymbol{P}_0 = \alpha \boldsymbol{I}$ $\boldsymbol{\zeta}_0 = \boldsymbol{O}$
Similarity threshold	$T_h = 0.9$
Number of forgetting coefficient	$\rho = 0.99$
Number of initial databases	$N(0) = 6$

3.2. *Simulation result*

The initial database is constructed using the initial PID gain shown in equation (25).

$$K_P = 0.486, K_I = 0.227, K_D = 0.112 \qquad (25)$$

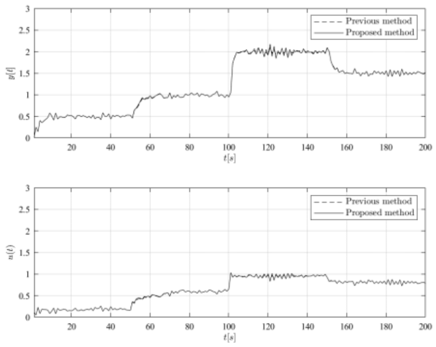


Fig.1 comparison of output between the proposed and previous method

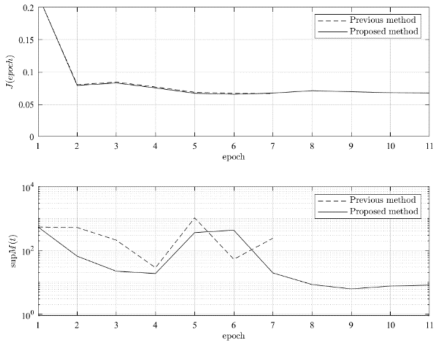


Fig.2 Result of b_1 by RLS

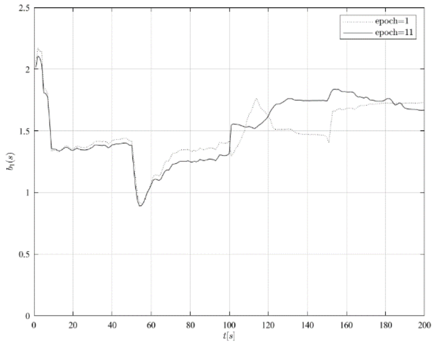


Fig.3 Result of the evaluation function and H_∞ norm

Fig. 1 shows the control results of the conventional and proposed methods, Fig. 2 shows the results of the evaluation function and H_∞ norm, and Fig. 3 shows b_1 estimated by RLS. Fig. 2 shows that the value of the evaluation function of the conventional method is 0.067, while that of the proposed method is 0.068. On the other hand, the H_∞ norm is 240.35 for the conventional method and 8.165 for the proposed method. Fig. 3 shows that the Jacobian b_1 of the system is always positive, but varies widely.

4. *Condition*

In this paper, we propose a method to adjust the learning coefficient of the gradient method appropriately based on the H_∞ norm obtained from the model identified by RLS, and a method to improve the stability of the system by adding information on the local H_∞ norm to the database. The effectiveness of the proposed method was confirmed through numerical simulations by adapting it to the Hammerstein-type nonlinear model discussed in a previous study.

Future work is to investigate a method for directly specifying the H_∞ norm and a method for real-time updating of PID gains when database-driven control is used.

References

1. N. Suda, “PID seigyo (PID control), Asakura syoten, 1992, (In Japanese).
2. S. Adachi, ”Sisutemu seigyo no kiso(Basic system identification)”, Tokyo Denki University Press, 2009,(In Japanese).
3. S. Wakitani, K. Nishida, M. Nakamoto, T. Yamamoto, “Design of a Data-Driven PID Controller using Operating Data,” IFAC Proceedings Volumes, Vol.46, No.11, pp-587-592, 2013.
4. K. Takao, T. Yamamoto, T. Hinamoto, ″A Design of Memory-Based PID Controllers,″ Transactions of the Society of Instrument and Control Engineers, Vol.40, No.9, pp.898-905, 2004,(In Japanese).
5. R. Kurozumi, T. Yamamoto, “A Design of CMAC Based Intelligent PID Controllers”, IEEJ Transactions on Electronics, Information and Systems, Vol.125,No.4, pp.607-615, 2005,(In Japanese).
6. S. Wakitani, Y. Ohnishi, T. Yamamoto, ”Design of a CMAC-Based PID Controller Using FRIT for Nonlinear Systems,” Transactions of the Society of Instrument and Control Engineers,Vol.48,No.12,pp.847-853, 2012,(In Japanese).
7. T. Kinoshita,T. Yamamoto, ”Design of a Database-Driven Control System based on the Similarity with Reducing Calculation Cost”, IEEJ Transactions on Electronics, Information and Systems,Vol.140, No.3, pp.312-319, 2020,(In Japanese).
8. T. Yamamoto, O. Kaneko, Y. Ohnisi, S. Wakitani, T. Kinoshita , H. Kugemoto, ”Data shikougata PID seigyo(Data-Driven PID control)”,Morikita publishing CO.LTD, 2020,(In Japanese).

Authors Introduction

Mr. Takumi Ueda

He received his M.E., in Engineering from Kyusyu Institute of Technology, Japan, in 2017. He is a third-year student in the doctoral program at the Kyushu Institute of Technology. His research interest includes non-linear control by PID using database.

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, Tokyo, Japan, in 1996. His research interests span both ship marine engineering and Intelligent Mechanics. He holds five patents derived from his research.

Occuluded Object Detection by Ultrasonic Sensors

Ryuugo Mochizuki
Center for Socio-Robotic Synthesis, Kyushu Institute of Technology, 2-4 Hibikino Wakamatsuku
Kitakyushu, Fukuoka, 808-0196, Japan

Yuya Nishida, Kazuo Ishii
Department of Human Intelligence Systems, Kyushu Institute of Technology, 2-4 Hibikino Wakamatsuku
Kitakyushu, Fukuoka, 808-0196, Japan

E-mail: Mochizuki.ryuugo126@mail.kyutech.jp, y-nishida@brain.kyutech.ac.jp, ishii@brain.kyutech.ac.jp
www.lsse.kyutech.ac.jp

Abstract

Crops are required to be processed in sanitary manner for lunchbox production in industries. Particularly, foreign object detection is crucial to guarantee consumer's health. Actually, foreign objects are mainly detected by human's eyes. For the automation, image processing is an example, however, is not effective for hidden objects. We propose a method of foreign object detection with ultrasonic reflection analysis. In our experiment, we selected four crops (lettuce, spinach, perilla and *komatsuna*) for masking material, and a coin for occluded objects. The coin were attached on the backside of the masking material. We recorded reflected wave moving the material linearly. As the result, reflection intensity doubled by the existence of occuluded object.

Keywords: Ultrasonic reflection, Occluded object detection

1. Introduction

Under the problem of population ageing and worker shortage in agriculture [1], automation of management, harvesting, and foreign object detection is becoming a challenging task. Above all, foreign object detection and elimination is an essential process to guarantee human health. Foreign objects are generally mixed with crops by attaching crop surfaces, or getting into the gap of leaves. Particularly, the latter case can often be overlooked in visual check. Physical strain of workers may grow up for searching foreign objects while moving leaves with care. Non-destructive manner is proper for the automation of foreign object detection.

Here, we propose to introduce ultrasonic wave analysis for foreign object detection, and analyzed waveform of reflected wave of crop surface.

2. Related Works

For the examples of automatic foreign object detection, Image processing method [2], X-ray analysis [3], Hyper-spectral imaging [4], Ultrasonic wave processing [5] are well applied. For image processing case, stone, metal piece in grain are detected by 94 [%][6]. However, hidden objects cannot be found by image processing. Hyper-spectral image can be valid for detecting objects which are not easily seen by naked eyes, which is enabled by multiple color channels of images. On the other hand, hyper spectral image [7] needs larger memory than usual color image. Such image analysis requires technique for reducing information. For X-ray technique, 100[%] of metal pieces were found in bread or cheeses [8], however, X-ray cannot easily transmit in water [7], and harmful to human beings. Ultrasonic wave can be applied to check inside sample without contact by sensors. P. Pallav, et. al. sandwiched a sample with a receiver and transmitter and analyzed the transmission of ultrasonic wave, then succeeded in finding a 2[mm] size of glass and rubber [9] B-K. Cho et, al. also found a 3[mm] size of glass [10], however, if a crop is large, a transmitter and receiver cannot enclose a thick sample. Furthermore, operation of the transmitter and receiver requires technique to avoid collision on crops.

To eliminate the need of operating transmitter and receiver, we propose to analyze reflected wave from the masking material which includes the reflected wave from foreign objects, And, we aim at expanding area on a crop surface for recording the reflection by multiple receivers.

3. Proposed Method

Figure 1 shows the experimental environment of our proposal. Acoustic reflection moves in various directions depending on the surface roughness of masking object, thus, intensity of the reception varies dramatically with only a receiver. To alleviate the uneven intensity, we attach multiple receivers (Figure 2(a)(b) shows the side view, top view, each other). The table in the figure is fixed with a masking material, which is distant from the transmitter by H_{TS}[mm], and the receiver by H_{RS}[mm]. The polyethylene base under the table is prepared for removing echo. A sample is made up by the base, table, masking material, which is moved manually in the direction of x-axis in Figure 1. Thus, change of reflected wave can be observed depending on the presence or absence of a foreign object. We satisfy the condition of $H_{TS} < H_{RS}$ to focus ultrasonic beam on one spot of the masking. Three receivers enclose a transmitter, and aligned at even interval on the circle whose center matches with the transmitter. The receivers are far from the transmitter by L_B[mm]

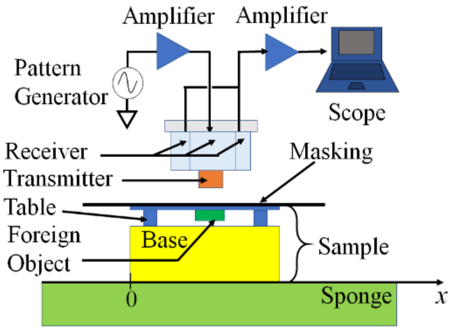
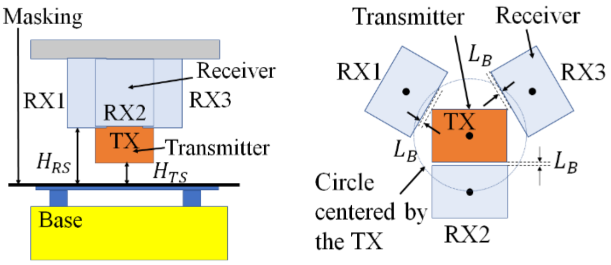


Fig. 1 Proposed Experimental Environment



(a) Side View (b) Top View

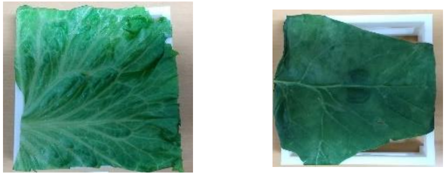
Fig. 2 Arrangement of a Transmitter and Receivers

4. Experiment

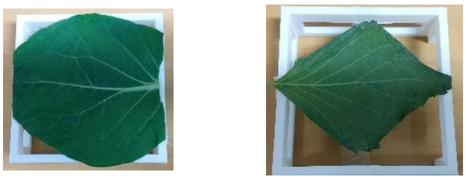
4.1. *Method*

As the experiment, we used four vegetables (Lettuce, Spinach, *Komatsuna*, Perilla, Figure 3) as masking materials, then attached foreign object (10 yen coin) back side of the masking on the table (Figure 4). We moved the sample and recorded change of the reflected waves.

In the environment of Figure 1, we output ten cycles of sign waves from the pattern generator, and input the waves (Amplitude is denoted by A) into the transmitter. Then, we obtained the reflected waves with three receivers and summed them up for the analysis. We set time interval of emission by nearly 5 [sec], and repeated

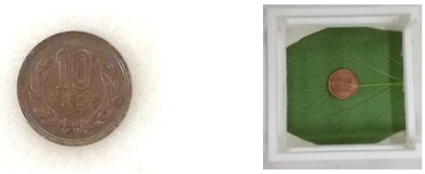


(a) Lettuce (b) Spinach



(c) *Komatsuna* (d) Perilla

Fig. 3 Maskings



(a)Foreign Object (b) Location of the Attachment

Fig. 4 Foreign Object

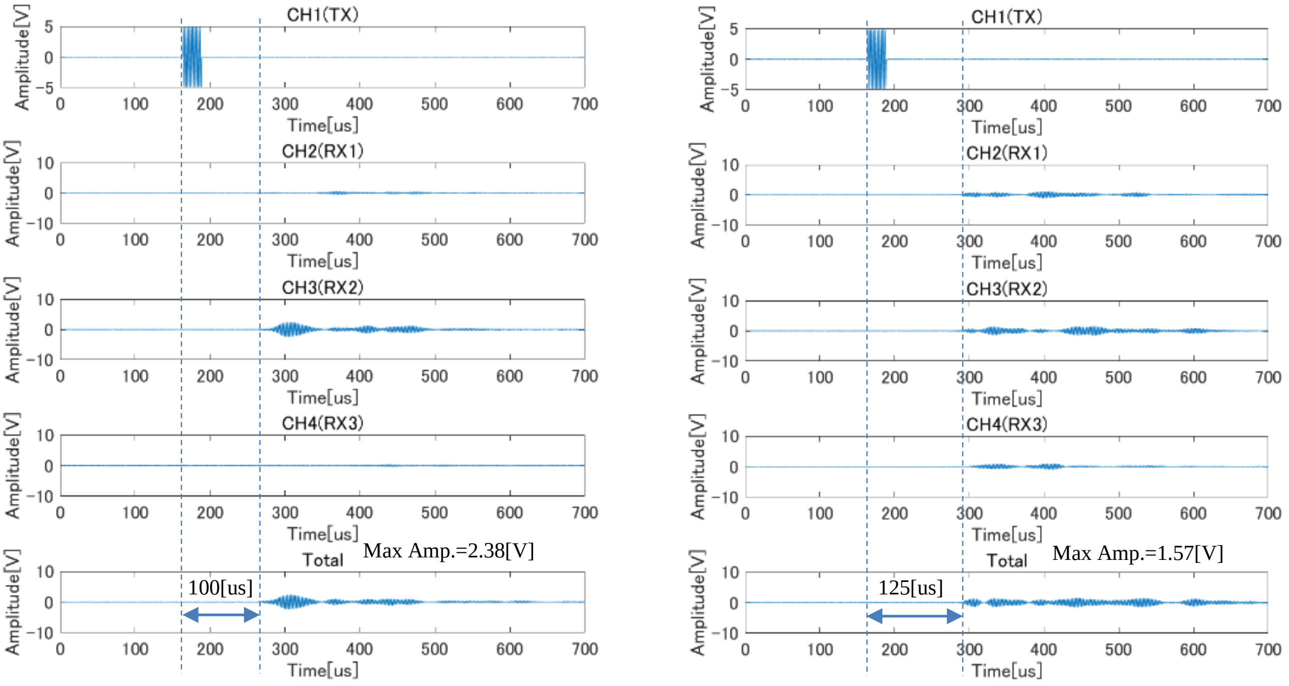
the emission, recording reflection, and sample displacement alternatively. The location of sample is measured by x-axis of the sample's back, which varies in the range of $-50 \leq x \leq 50$ [mm]. At *x*=0[mm], receiver is located over the foreign object. We recorded reflected wave which was amplified with gain *G*. Table 1 shows the parameter setting for the experiment.

Table 1 Parameter Setting

Symbol	Value
H_{TS}	10[mm]
H_{RS}	20[mm]
L_B	1[mm]
A	±50[V]
G	x120

4.2. *Result and Discussion*

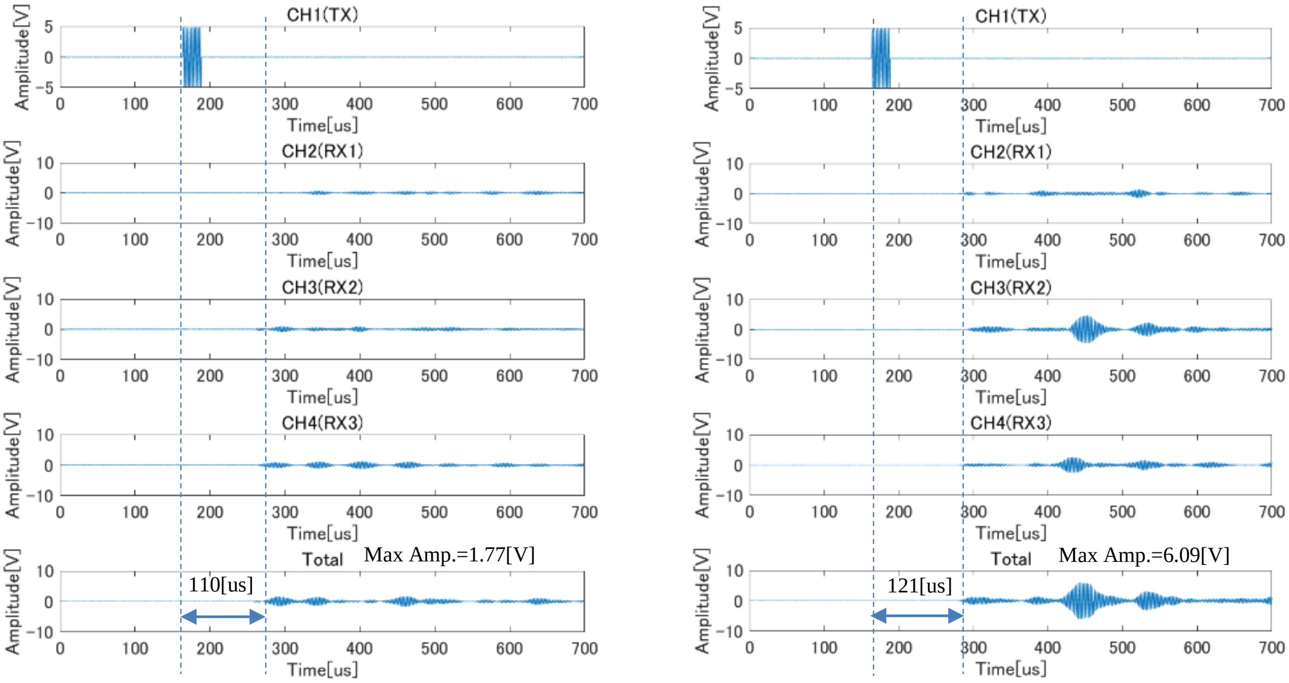
The amplified reflection is shown as Figure 5(Lettuce), Figure 6(Spinach), Figure 7(*Komatsuna*), Figure 8. (Perilla). The amplitude of summed reflection from three receivers, and elapsed time from emission to reception are also shown in the same figures. In case foreign object (10yen coin) is attached backside of the masking, apart from lettuce, the amplitude of the reflection increased to twice or more. However, the increase of the amplitude was as shown in the echo between transmitter and the masking. The time was different from the first arrival of reflection at receiver. Considering sonic velocity is equal to 340 [m/s], nearly 90 [us] is required to reach receivers because $H_{TS} = 10[\mathrm{mm}], H_{RS} = 20[\mathrm{mm}]$. However, the wrinkles on the masking caused of the error (10-20 [us] longer than the calculation). With foreign object attached, the time of arrival delayed more as the result of sinking material by the weight of coin. If H_{TS} is small, transmitted wave easily hits the coin, however, the reflection interfered each other, which made the analysis harder. H_{TS} should be adjusted so that no interference occurs.



(a)No Foreign Object

(b) With Foreign Object

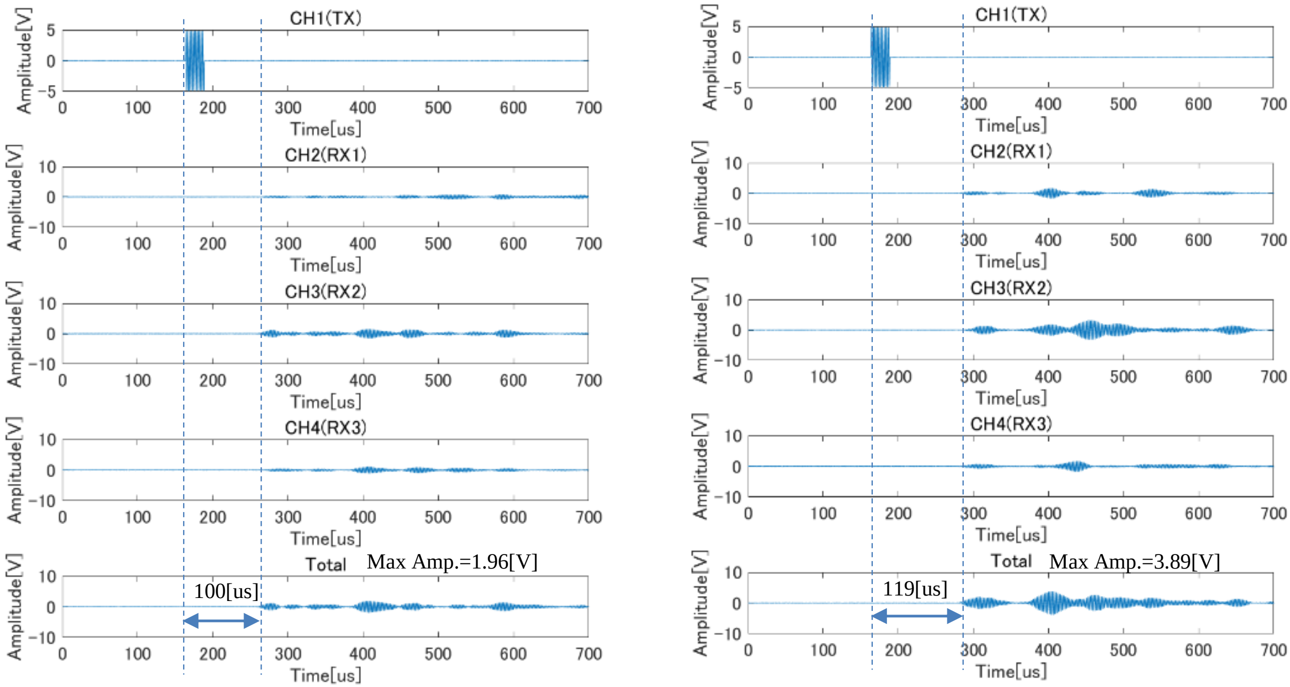
Fig. 5 Reflection from Lettuce (*x*=0)



(a)No Foreign Object

(b) With Foreign Object

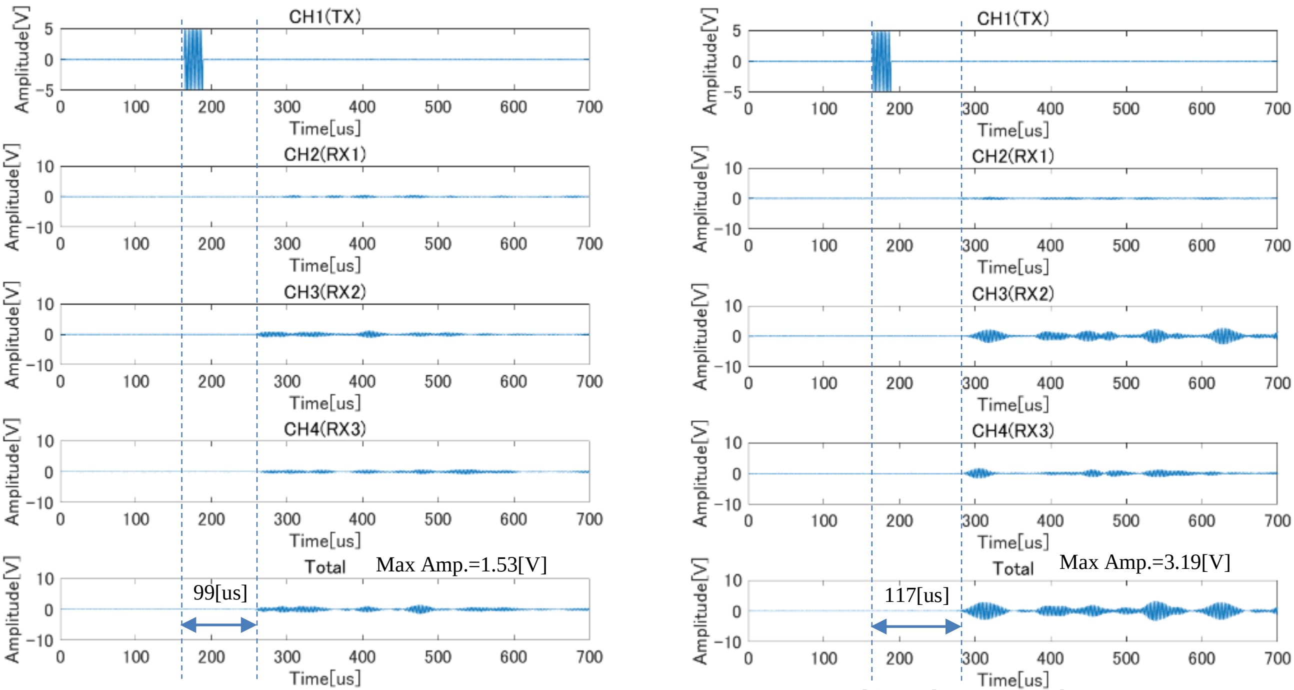
Fig. 6 Reflection from Spinach (*x*=0)



(a)No Foreign Object

(b) With Foreign Object

Fig. 7 Reflection from *Komatsuna* (*x*=0)



(a)No Foreign Object (b) With Foreign Object

Fig. 8 Reflection from Perilla (x=0)

5. Conclusion

We chose four vegetables as masking for the experiment of foreign object detection. As the result, the maximum amplitude of the reflection increased to more than twice if foreign object (coin) is attached backside of the masking (but for lettuce). However, reflection interference was a problem for the analysis if the short distance($H_{TS} = 10$[mm]) was set between the transmitter and masking. The distance should be reconsidered so that the interference is prevented.

References

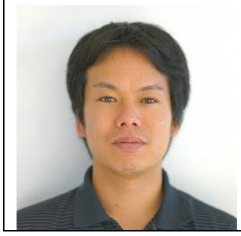
1. N. Noguchi, “Robots for Agriculture and Food Industry (In Japanese)”, Journal of Society of High Technology in Agriculture”, Vol. 33, No. 3, pp. 96-98, 2021.
2. D. Rong,et. al, “Computer Vision Detection of Foreign Objects in Walnuts Using Deep Learning”, Computers and Electronics in Agriculture, Vol. 162, pp. 1001-1010, 2019.
3. R. P. Haff, N. Toyofuku, “X-ray Detection of Detects and Contaminants in the Food Industry”, Sensing and Instrument for Food Quality and Safety, Vol. 2, No. 4, pp. 262-273, 2008.
4. C. Mo, et. al, “Fluorescence Hyperspectral Imaging Technique for Foreign Substance Detection on Fresh-cut Lettuce”, Journal of the Science of Food and Agriculture, Vol. 97, Issue. 12, pp. 3871-4263, 2017.
5. M. Taufiq, et. al, “Detection of Foreign Objects in Milk Using an Ultrasonic System”, Indonesian Journal of Electrical Engineering and Computer Science, vol. 15, No. 3, pp. 1241-1249, 2019.
6. I. E. Elbatawi, G. K. Arafa, “Application of Machine Vision for Detection of Foreign Matter in Wheat Grains”, Arab Universities, Journal of Agricultural Sciences, Vol. 16, No. 2, pp. 275-283. 2008.
7. A. Adedeji, et. al., “Non-Destructive Technology for Detecting Insect Infestation in Fruits and Vegetables under Postharvest Conditions: A Critical Reviews”, Foods, Vol. 9, No. 7, pp. 2-28, 2020.
8. H. Einarsdottir, et. al., “Novelty Detection of Foreign Objects in Food Using Multi-modal X-ray Imaging” , Food Control vol. 67, pp. 39-47, 2016.
9. P. Pallav, et. al., “Air-coupled Ultrasonic Evaluation of Food Materials”, Ultrasonics, Vol. 49, pp. 244-253, 2009.
10. B. –K Cho and J.M.K Irudayaraj:”Foreign Object and Internal Disorder Detection in Food Material Using Noncontact Ultrasound Imaging“, Journal of Food Science, vol. 68, no, 3, pp. 967-974, 2003

Authors Introduction

Dr. Ryuugo Mochizuki

He is a researcher at Graduate School of Life Science and System Engineering, Kyushu Institute of Technology, Japan. His research topic during the PhD course was image processing. Currently, he is involved in signal processing of ultrasonic.

Prof. Yuya Nishida



He is an Associate Professor at Graduate School of Life Science and System Engineering, Kyushu Institute of Technology, Japan. His research area is about filed robotics, its application, and data processing

Prof. Kazuo Ishii

He is a Professor at Graduate School of Life Science and System Engineering, Kyushu Institute of Technology, Japan. His research area is about field robots and intelligent robot system

Development of a Variable Stiffness Function for a New Multifunctional Wire Driven Joint Mechanism

Katsuaki Suzuki
Life Science and Systems Engineering, Kyushu Institute of Technology
2-4 Hibikino,Wakamatsu-ku, Kitakyushu 808-0196, Fukuoka, Japan
E-mail: suzuki.katsuaki448@mail.kyuteck.jp

Yuya Nishida
Life Science and Systems Engineering, Kyushu Institute of Technology
2-4 Hibikino,Wakamatsu-ku, Kitakyushu 808-0196, Fukuoka, Japan
E-mail: ynishida@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

A mechanical variable stiffness can adapt to external forces beyond the control cycle, such as overturning of walking robots, which can help solve problems such as actuator and joint destruction. In this research, we will clarify the structure of a two-input, one-output joint mechanism that can realize three functions: normal motion, instantaneous motion, and variable stiffness function. As part of the development of the variable stiffness function, a mathematical model is derived and simulated for the relationship between the joint angle of the output link and the disturbance torque when a disturbance torque is applied to the output link of the proposed mechanism.

Keywords: Mechanical variable stiffness, Link mechanism

1. Introduction

When robots and humans work together, making the joints of robots flexible is an important issue to prevent robots from harming humans. One way to make robot joints flexible is to achieve variable stiffness. Methods using software, such as impedance control and force control, have been put to practical use in industrial robots as a means of achieving variable stiffness. However, it has some problems such as difficulty in responding to responses exceeding the control cycle. On the other hand, methods to realize variable stiffness mechanically have also been researched [1],[2],[3],[4],[5],[6],[7],[8]. Sonoda et al. proposed a variable-stiffness joint mechanism that mimics living organisms [9]. This mechanism uses two motors, two cams, and wires to achieve variable stiffness and does not use springs. An arbitrary reduction ratio is designed using a cam, and a nonlinear spring element is realized. Since the posture and stiffness can be controlled only by the magnitude of the torque applied to the two motors, the stiffness can be changed at high speed compared to other mechanisms. As a feature of these mechanisms, each mechanism uses two motors, and realizes a variable stiffness function by using an antagonistic structure and a nonlinear spring element. On the other hand, the problem is that it is difficult to use in applications that require rapid motion. The authors have developed a joint mechanism that can perform normal and instantaneous movements in order to expand the range in which the robot can be active [10]. In this paper, as a new function of the joint mechanism that we have developed so far, which is capable of normal motion and rapid motion, we propose a method to realize variable stiffness by devising the driving method of two motors while keeping the existing mechanism. Section 2 describes how the proposed mechanism achieves variable stiffness. In Chapter 3, we derive mathematical formulas and analyze the torque and stiffness with respect to the angle of the output link passively rotated by the disturbance torque. Section 4 presents the conclusions.

2. Variable stiffness function of the proposed mechanism

Fig.1 shows how the proposed mechanism realizes variable stiffness. First, we assume a state in which the initial displacement of the spring is 0, that is, the state in which the two motors are braked with the length of the spring at its natural length. When a disturbance torque T_q is applied to the output link while the brakes are applied to the two motors, tension acts on one of the wires according to the magnitude and direction of the torque, and the slider moves in translation. At this time, slackness occurs in the wire on the side where tension does not occur. At the position where the tension of the wire and the compressive force acting on the spring are balanced, the movement of the slider stops and the passive rotational motion of the output link also stops. This indicates that the output link passively displaced according to the disturbance torque T_q applied to the output link. Here, as shown in Fig.1, we assume a state in which the internal spring is displaced by ε_{sp} and two motors are braked. As in the case where the initial displacement of the spring is 0, applying a disturbance torque to the output link causes tension to act on the wire. The passive motion of the output link then ceases at the point where the tension of the wire and the compression force of the spring are balanced. Comparing the initial displacement of the spring of 0 and ε_{sp}, in the case of ε_{sp}, a compressive force acts on the spring before the load is applied to the output link. Therefore, the passive displacement of the output link does not occur unless the spring contraction direction component of the tension acting on the wire due to the disturbance torque exceeds the spring compression force accumulated by the initial displacement of ε_{sp}. In addition, when the force acting on the spring due to the disturbance torque exceeds the compressive force acting on the spring due to the initial displacement of ε_{sp}, passive displacement of the output link occurs, but compared to the case where the initial displacement is 0, The displacement becomes smaller. In other words, the stiffness of the joint can be changed by adjusting the initial displacement of the spring inside the mechanism. In the proposed mechanism, such a principle is applied to realize a variable stiffness function.

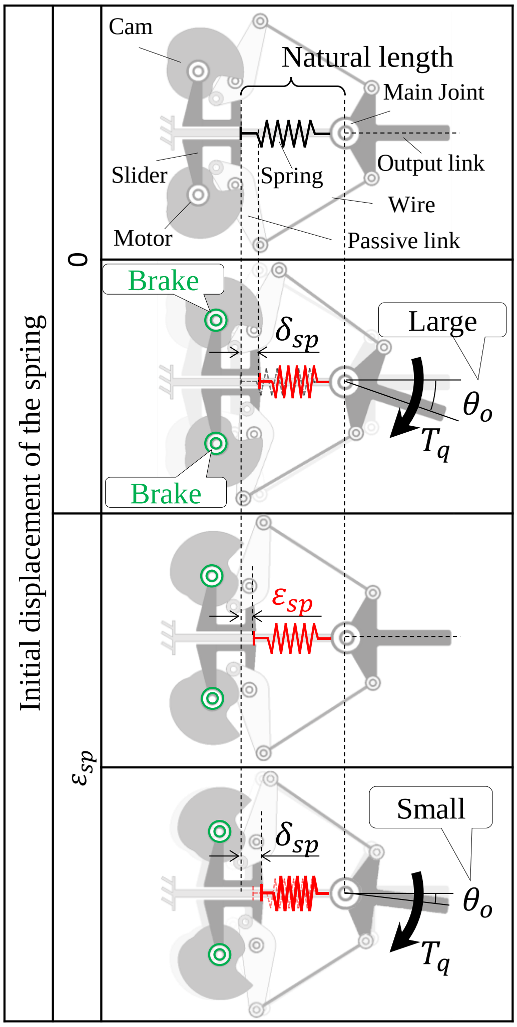


Fig.1 Variable stiffness function of the proposed mechanism

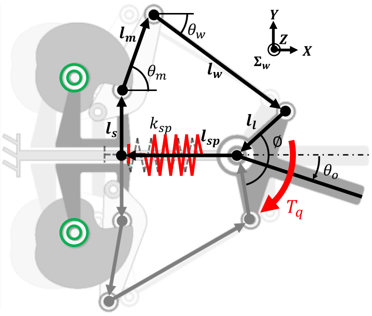


Fig.2 Link vectors of the proposed mechanism

3. Analysis of an output characteristics related to stiffness

In the case of using variable stiffness, we consider the disturbance torque and rotational stiffness of the output link that change with the initial displacement of the spring. Fig.2 shows the link vectors of the proposed mechanism. When a disturbance torque is applied to the output link, the output link rotates. Then, the rotational motion of the output link is converted into the translational motion of the slider through the wire. At this time, the spring inside the mechanism is compressed, and tension acts on the wire due to the compressive force of the spring. Therefore, in order to apply a disturbance torque to rotate the output link, it is necessary to apply a disturbance torque of the same magnitude as the torque generated around the output link due to the internal force such as wire tension. The relationship between the disturbance torque T_q and the internal force acting inside the mechanism is expressed by the following equation from the geometric relationship.

$$T_q = sign(\theta_o) k_{sp} \delta_{sp} l_l (-\sin\beta + \tan\theta_w \cos\beta) \qquad (1)$$

The rotational stiffness K of the output link is expressed by the following equation by differentiating the disturbance torque T_q given to the output link by the angle θ_o of the output link.

$$K = \frac{dT_q}{d\theta_o} = k_{sp} l_l \delta_{sp} \left[\frac{l_l \cos^2\beta\,(1+\tan^2\theta_w)}{l_w \cos\theta_w} - \tan\theta_w \sin\beta - \cos\beta \right] + k_{sp} l_l^2 (\sin\beta - \cos\beta\tan\theta_w)^2 \qquad (2)$$

Here, β used in formulas (1) and (2) is expressed by the following formula.

$$\beta = \begin{cases} \theta_o - \dfrac{\emptyset}{2} \ (\theta_o \geq 0) \\ -\theta_o - \dfrac{\emptyset}{2} \ (\theta_o < 0) \end{cases} \qquad (3)$$

Furthermore, the wire angles θ_w and θ_m are expressed by the following equations from the geometrical relationship.

$$\theta_w = \sin^{-1}\left(\frac{l_s + l_m \sin\theta_m - l_l \sin\beta}{l_w}\right) \qquad (4)$$

$$\theta_m = \cos^{-1}\left(\cos\theta_{mi} - \frac{\varepsilon_{sp}}{l_m}\right) \qquad (5)$$

Fig.3 shows the profile of the disturbance torque T_q of the output link and the angle θ_o of the output link when the initial displacement ε_{sp} of the spring is changed to 0 mm, 10 mm, and 20 mm using equation (1). Fig.3 shows the profile of the disturbance torque T_q of the output link and the angle θ_o of the output link when the initial displacement ε_{sp} of the spring is changed to 0 mm, 10 mm, and 20 mm using equation (1). Similarly, using equation (2), Fig.4 shows the profile of the rotational stiffness K and θ_o of the output link when ε_{sp} is changed to 0 mm, 10 mm, and 20 mm. This analysis was performed using the parameters shown in Table 1. As a result, it was confirmed that the apparent stiffness can be changed by changing the magnitude of the reaction force according to the initial displacement of the spring.

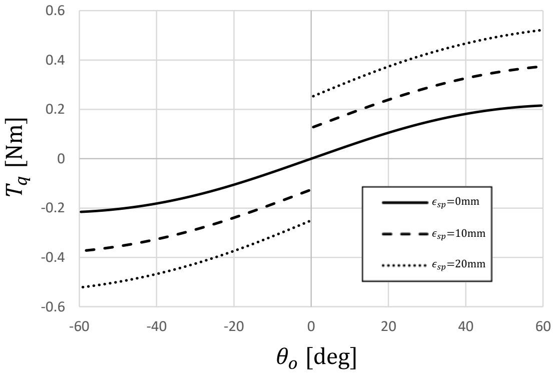


Fig.3 Relationship between rotation angle of output link and disturbance torque

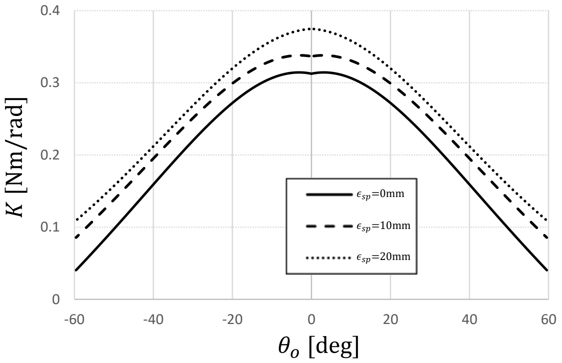


Fig.4 Relationship between output link rotation angle and rotational stiffness

Table 1 Parameters for analysis

k_{sp}	Spring constant	500N/m
l_l	Length of $\boldsymbol{l_l}$	0.025m
l_s	Length of $\boldsymbol{l_s}$	0.012m
l_w	Length of $\boldsymbol{l_w}$	0.068m
l_{spn}	Length of $\boldsymbol{l_{spn}}$	0.092m
l_m	Length of $\boldsymbol{l_m}$	0.03m
l_{cp}	Length of $\boldsymbol{l_{cp}}$	0.0181m
$\emptyset$	Direction of $\boldsymbol{l_l}$	180°
θ_{max}	Maximum movement range	+60°
θ_{min}	Minimum movement range	-60°
θ_{mi}	Absolute angle of $\boldsymbol{l_m}$ when θ_o=0,l_{sp}=l_{spn}	126°

4. Conclusion

In this paper, we presented a method for realizing variable stiffness in the proposed mechanism that can perform normal and rapid motions. In addition, the relational expression between the joint angle of the output link and the disturbance torque and the expression for the rotational stiffness were derived. From the analysis results, it was found that the magnitude of the disturbance torque required to rotate the output link changes according to the initial displacement of the spring, and the apparent stiffness changes. Therefore, it was clarified that the proposed mechanism can achieve variable stiffness by adjusting the initial displacement of the spring.

References

1. Masafumi Okada. Human-Like Compliant Motion and Its Mechanism. Journal of Robotics Society of Japan, 1999, 17. 6: 782-785.
2. LAURIN-KOVITZ, Kirsten F.; COLGATE, James Edward; CARNES, Steven DR, Design of components for programmable passive impedance . In: ICRA. 1991. p. 1476-1481.
3. Jin'ichi YAMAGUCHI, et al. Development of a Bipedal Humanoid Having Antagonistic Driven Joints. Society of Biomechanisms Japan, 1998, 14: 261-271.
4. Toshio Morita, et al. Development of Force-Controlled Robot Arm Using Mechanical Impedance Adjuster. Journal of Robotics Society of Japan, 1998, 16. 7: 1001-1006.
5. Toshiro NORITSUGU, tomu WADA. Applications of Artificial Rubber Muscle to Robot
6. Control. Journal of Robotics Society of Japan, 1991, 9. 4: 502-506.
7. Tomoya Inaba, Toshiki Nakazawa, Koichi Koganezawa. Stiffness and Angle Control of Antagonistially Driven Joint using the Actuator with Non-Linear Elastic System (ANLES) T. Journal of Robotics Society of Japan, 2008, 26. 4: 381-388.
8. Kazuhito HYODO, Hiroaki KOBAYASHI. A Study on Tendon Controlled Wrist Mechanism with Nonlinear Spring Tensioner. Journal of Robotics Society of Japan, 1993, 11.8: 1244-1251.
9. SONODA, Takashi, et al, Development of antagonistic wire-driven joint employing kinematic transmission mechanism. *Journal of Automation Mobile Robotics and Intelligent Systems*, 2010, 4: 62-70.
10. 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.

Authors Introduction

Dr. Katsuaki Suzuki

He received his Ph.D. from the Department of Life Science and Systems Engineering, Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology, Japan in 2021. His research interests include joint mechanisms and their applications.

Dr. Yuya Nishida

He is an Associate Professor at 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 Graduate School of Life Science and System Engineering, Kyushu Institute of Technology, Japan. His research intetrests include field robots and intelligent robot systems.

Analyzing an OFDM System using Cyclic Prefix to Improve the Underwater Communication

Hussam Alraie[*]
Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu , Kitakyushu, Fukuoka 808-0196, Japan[†]

Kazuo ISHII
Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu , Kitakyushu, Fukuoka 808-0196, Japan
E-mail: alraie.hussam973@mail.kyutech.jp, ishii@brain.kyutech.ac.jp
www.lsse.kyutech.ac.jp

Abstract

Underwater Communication is one of the most difficult challenges facing researchers. OFDM technique has been used widely in 4G communication systems, and recently it was approved to be successfully implemented in 5G. Many researchers have improved the Underwater wireless communication system including Acoustic communication using the OFDM technique. We applied the Cyclic Prefix in this study to improve the underwater communication system. We started by studying the effects of Rician fading based on the OFDM system on the wireless channel. Then we analyzed the Bit Error Rate (BER) of the system in several scenarios, by applying the AWGN, Rician fading, and absorption factor to simulate the underwater channel.

Keywords: OFDM, Cyclic Prefix, 5G, Underwater, AWGN, Rician fading.

1. Introduction

The wireless signals underwater are scattered widely which affects the quality of the received data and will add more difficulties to restore the original data.

The propagation of the underwater acoustic signals is limited due to the high absorption of the water [1].

OFDM and MIMO-OFDM are used to improve the underwater communication response which can increase the data rate due to the multi-carrier property.

Tabeshnezhad, A., and Pourmina, M. A. used MIMO-OFDM to increase the data rate with QPSK digital modulation, and they used the Alamouti MIMO-OFDM to improve the underwater optical communication (UWOC) [2].

Thottappilly, A. worked on underwater acoustic communication and analyzed the OFDM system and doppler effectiveness on the transmission signal using MATLAB [3].

In this paper I will show the simulation result of an OFDM system with and without CP by analyzing the system in an Additive White Gaussian noise (AWGN) environment only, then compare the results after adding Rician fading and finally Rician fading with absorption effect which will simulate the underwater communication system.

2. Towards OFDM

A Frequency Division Multiplexing system (FDM) divides the whole bandwidth into non-overlapping channels.

Each channel has its carrier and to prevent the overlapping between each channel there is a guard period. The whole data transfers in parallel through the carriers as shown in Figure 1.

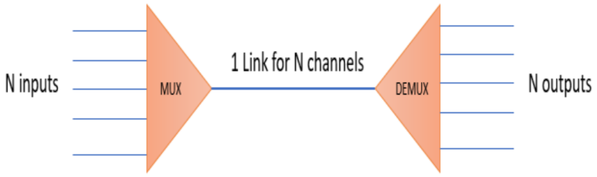


Fig. 1. FDM transfer method

Loss of bandwidth between each carrier is one of the disadvantages of FDM.

Time-divisions multiplexing system divides the channel into time slots that can be used for each user. The disadvantage of this method is that each user should wait until the end of the previous time slot to be able to send its data.

Orthogonal Frequency Division Multiplexing system (OFDM) is a multi-carrier modulation. The data transfers through different subcarriers in one symbol, and each carrier should be modulated using one of the modulation schemes such as QPSK, QAM, and BPSK.

The effects of the Underwater Acoustic (UWA) channel decrease the quality of communication due to the Doppler effect, low propagation speed, and harsh water channel. The cyclic Prefix (CP) adds a guard interval between each OFDM symbol which can help to prevent Inter Symbol Interference (ISI) [4].

OFDM System

In the OFDM system, we start by modulating the input data using one of the digital modulation schemes. The modulated data should be converted to parallel data which will be the input of Inverse Fast Furrier Transform (IFFT). IFFT transfers the signal from the time domain to the frequency domain. The converted data should be transferred as serial data through the channel. Figure 2 illustrates the process of the OFDM system transmission and receiving.

The channel characteristics add additional noise to the signal and decrease its quality.

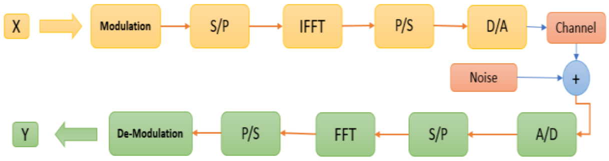


Fig. 2. OFDM system box chart

3. Methodology

We used the MATLAB simulation program to analyze the OFDM system in multiple scenarios with and without CP for 64,128 and 256 subcarriers:

- OFDM with AWGN only.
- OFDM with AWGN and Rician fading.
- OFDM with AWGN and Rayleigh fading.
- OFDM with AWGN, Rician fading, and absorption factor.

Then compare the BER for each scenario.

In the simulation, we assume that the absorption factor for underwater communication affects the signal by multiplying with a random number between [0 - 1].

0 means that the absorption effect suppresses the signal at 100% and 1 means that there is no absorption effect.

4. Results and discussions

(1) Scenario 1: OFDM with AWGN only

In this scenario, we simulate the OFDM system by adding AWGN only and compare the BER with and without CP.

Table 1 shows that BER is decreased by increasing the Signal Noise Ratio (SNR).

Table 1. BER comparison with and without CP

BER in AWGN

Eb/N0 (dB)	BER CP = 0	BER CP = 4
1	0.00417	0.04
2	0.028	0.024
3	0.02164	0.0154
4	0.0078	0.0066
5	0.0027	0.0029
6	0.0011	0.001
7	0.00027	0.00026
8	$7.2 * 10^{-5}$	$3.9 * 10^{-5}$
9	$9 * 10^{-6}$	$4.99 * 10^{-6}$

It can be noticed that the BER decreased by about 50% when the SNR reach 8 dB.

(2) Scenario 2: OFDM with 64,128 and 256 subcarriers

In this scenario, we simulate the OFDM and compare the BER with and without CP for AWGN only then OFDM with AWGN, Rician fading, and absorption factor, using 64,128 and 256 subcarriers as shown in Figure 3.

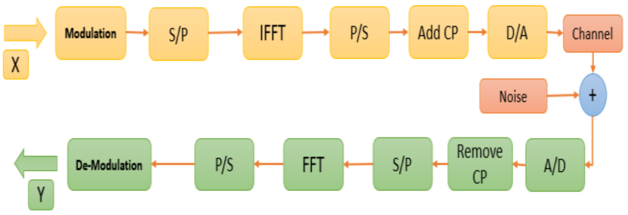


Fig. 3. OFDM simulation flow chart

Figure 4 illustrates the BER comparison in OFDM system CP free using 64 subcarriers. We can notice the effect of the absorption factor which increases the BER of the Rician fading channel.

For 64, 128, and 256 we use the same way of comparison with and without CP, and Tables 2, 3 and 4 show the results.

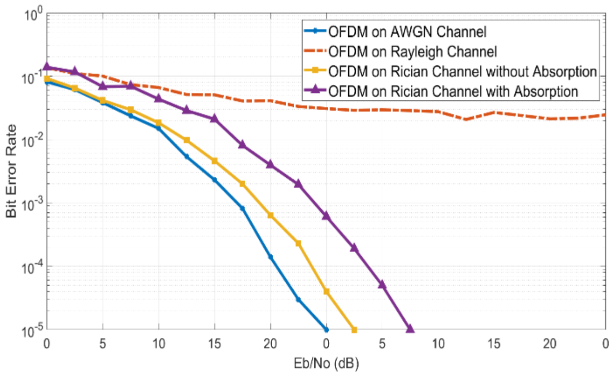


Fig. 4. BER in AWGN environment

Table 2. BER comparison with and without CP for 64 SCs

SNR	BER CP_Free	CP = 1% of SC	CP = 3% of SC	CP = 5% of SC
6 dB	2.1×10^{-2}	1.4×10^{-2}	$1.4 * 10^{-2}$	1.2×10^{-2}
7 dB	8×10^{-3}	7.8×10^{-3}	10^{-2}	7.2×10^{-3}
8 dB	4×10^{-3}	3.3×10^{-3}	3.3×10^{-3}	3×10^{-3}
9 dB	1.95×10^{-3}	1.5×10^{-3}	1.76×10^{-3}	10^{-3}
10 dB	6×10^{-4}	4.5×10^{-4}	5×10^{-4}	4×10^{-4}

Table 3. BER comparison with and without CP for 128 SCs

SNR	BER CP_Free	CP = 1% of SC	CP = 3% of SC	CP = 5% of SC
6 dB	1.6×10^{-2}	1.5×10^{-2}	1.6×10^{-2}	1.6×10^{-2}
7 dB	10^{-2}	7.2×10^{-3}	7.3×10^{-3}	7.8×10^{-3}
8 dB	4.2×10^{-3}	3.9×10^{-3}	3.4×10^{-3}	3.6×10^{-3}
9 dB	1.8×10^{-3}	1.2×10^{-3}	1.5×10^{-3}	1.7×10^{-3}
10 dB	5.9×10^{-4}	3.9×10^{-4}	5.3×10^{-4}	5.9×10^{-4}

Table 4. BER comparison with and without CP for 256 SCs

SNR	BER CP_Free	CP = 1% of SC	CP = 3% of SC	CP = 5% of SC
6 dB	1.3×10^{-3}	1.3×10^{-3}	2.3×10^{-2}	1.8×10^{-2}
7 dB	7.5×10^{-3}	7.5×10^{-3}	1.3×10^{-2}	8.7×10^{-3}
8 dB	2.8×10^{-3}	3.3×10^{-3}	5.7×10^{-3}	3.7×10^{-3}
9 dB	10^{-3}	1.7×10^{-3}	2.2×10^{-3}	2.3×10^{-3}
10 dB	4.9×10^{-4}	5.7×10^{-4}	9.1×10^{-4}	8×10^{-4}

From Table 4 it can be noticed that using CP in an OFDM system with 256 subcarriers increases the BER because increasing the number of subcarriers extends to the OFDM symbol size, as a result, we need to increase the CP length to reduce BER which will reduce the efficiency of bandwidth. Table 2 and 3 show that using CP in an OFDM system with 64 and 128 reduce the BER and as a result, it increases the quality of the received signal.

Referring to Tables 2, 3 and 4 BER enhanced with values as shown in Tables 5, 6 and 7 which will summarize the results of BER improvement in percentage by changing SNR and CP values.

Table 5. BER comparison when CP = 1% of Sc

SNR	BER compared with CP_Free performance CP = 1% of 64 SC	CP = 1% of 128 SC	CP = 1% of 256 SC
6 dB	-33.67%	-6.25%	0%
7 dB	-1.25%	-18%	0%
8 dB	-17.5%	-7.2%	+17.8%
9 dB	-25%	-33.67%	+70%
10 dB	-25%	-34.1%	+16.3%

Table 6. BER percentage comparison when CP = 3% of Sc

	BER compared with CP_Free performance		
SNR	CP = 3% of 64 SC	CP = 3% of 128 SC	CP = 3% of 256 SC
6 dB	-33.67%	0%	+77%
7 dB	+25%	- 27%	+73.33%
8 dB	-17.5%	-19.1%	+103.5%
9 dB	-10%	-16.67%	+120%
10 dB	-8.3%	-10.2%	+85.7%

Table 7. BER percentage comparison when CP = 5% of Sc

	BER compared with CP_Free performance		
SNR	CP = 5% of 64 SC	CP = 5% of 128 SC	CP = 5% of 256 SC
6 dB	-57.15%	0%	Increased more than a tribble
7 dB	-10%	-27%	+16%
8 dB	-25%	-19.1%	+32.14%
9 dB	-51.82%	-16.67%	+130%
10 dB	-33.67%	-10.2%	+63.26%

Tables 5,6 and 7 show that using CP with 1% and 3% of the 64 subcarriers in underwater OFDM system improves the BER by 33.67% and enhanced up to 57% when SNR reaches 6 dB while using 1% of 128 subcarriers improves the BER between 33-34 % but after increasing the SNR to 9 or 10 dB. In contrast, using CP in 256 subcarriers shows a negative response of the OFDM system even if CP length increases up to 5% of 256 subcarriers.

5. Conclusion

In this paper, a proposed relationship between the CP and the number of subcarriers shown in tables 5,6, and 7 showed that using CP improved the OFDM system using 64 and 128 subcarriers compared with 256 subcarriers. In addition, using 1% of CP in both 64 and 128 is the best choice compared with the studied scenarios, because it reduces the reserved bandwidth for CP. To get the same performance while using CP in both suggested systems we should make sure to increase the SNR from 6 dB in the 64-subcarrier system up to 9 dB in the 128-subcarrier system.

Using 3% of subcarriers was efficient for both 64 and 128 subcarrier systems, it improves the BER compared with CP_Free results while using 5% of subcarriers showed a better result in 64 subcarrier systems when the SNR reaches 6 or 8 dB compared with 128 subcarriers.

It can be noticed that using 1% is better than 3% of subcarriers because it reserved less bandwidth.

Using CP in 256 subcarriers simulated system showed the negative response of the BER performance under the comparison percentages.

Acknowledgement

This paper is supported by JSPS KAKENHI Grant Number JP18H01643.

References

1. Chitre, M., Shahabudeen, S., Freitag, L., & Stojanovic, M. (2008). Recent advances in underwater acoustic communications & networking. OCEANS 2008, 1-10.
2. Tabeshnezhad, A., & Pourmina, M. A. (2017). Outage analysis of relay-assisted underwater wireless optical communication systems. Optics Communications, 405, 297-305.
3. Thottappilly, A. (2011). OFDM for underwater acoustic communication (Doctoral dissertation, Virginia Tech).
4. Li, B., Zhou, S., Stojanovic, M., Freitag, L., & Willett, P. (2008). Multicarrier communication over underwater acoustic channels with nonuniform Doppler shifts. IEEE Journal of Oceanic Engineering, 33(2), 198-209

Authors Introduction

Mr. Hussam Alraie

He received his Master's degree from the Department of Communication Engineering, AL-Baath University, Homs, Syria in 2018. He is currently a Doctoral student in Kitakyushu, Life Science and System Engineering, Kyushu Institute of Technology, Japan.

Prof. Kazuo ISHII

He is currently a Professor at the Department of Human Intelligence Systems of Kyushu Institute of Technology, Japan. His research interests are in the fields of Underwater Robotics, Field Robotics, Neural Networks, and Intelligent Systems.

Increasing Selectivity to a Feature Combination Using Inhibitory Synaptic Plasticity in a Spiking Neural Network.

Mahiro Ikeda
Graduate School of Information Science and Technology, Osaka Institute of Technology,
1-79-1 Kitayama, Hirakata, Osaka 573-0196, Japan

Hirotsugu Okuno
Faculty of Information Science and Technology, Osaka Institute of Technology,
1-79-1 Kitayama, Hirakata, Osaka 573-0196, Japan
E-mail: hirotsugu.okuno@oit.ac.jp
www.oit.ac.jp

Abstract

In this study, we designed a spiking neural network that uses synaptic plasticity to increase selectivity to a particular combination of features. We investigated how the time constant of inhibitory presynaptic neurons whose weights were updated by the long-term potentiation of inhibition affects to selectivity of the postsynaptic neurons. The results showed that the selectivity was increased effectively when the time constant of inhibitory neurons was slightly longer than that of postsynaptic neurons.

Keywords: spiking neural network, STDP, LTPi, combination of features

1. Introduction

In multi-layered neural networks for object classification, image features are extracted in each layer and are transferred to the latter layers. It is important for neurons in the intermediate layers to respond selectively to a combination of features because these combinations of features are essential for object classification. In the case of image classification, features are color, contour orientation, and so on.

A large majority of neural networks consist of so called artificial neurons whose inputs and outputs are represented by real numbers. However, actual neurons in the brain use pulsed potential changes called spikes for information transmission, not real numbers. Therefore, spiking neural networks, which mimic the cortical neuronal networks and consist of model neurons that use spikes for information transmission and processing, are expected to be the artificial intelligence of the next generation.

In the actual neuronal circuits, inhibitory neurons play an important role in information processing. Inhibitory neurons decrease the membrane potential of the postsynaptic neurons, and inhibition can increase the selectivity of the postsynaptic neurons to a particular set of stimuli by decreasing sensitivity to the other stimuli. A previous study suggested that long-term potentiation of inhibition (LTPi) is important in fine-tuning cortical circuitry in response to visual experience [1], and another study showed that LTPi can help spiking neural networks to respond to a specific stimulus-pair [2].

In this study, we designed a small network in which inhibitory neurons and their postsynaptic neurons respond to a combination of features and examined the relationship between the time constant of inhibitory neurons and selectivity of the postsynaptic neurons after training based on LTPi.

2. Models of Neurons and Synapses

2.1. *Model of neurons*

In this study, the leaky integrate-and-fire (LIF) model was used as a model of neurons [3]. The change in the membrane potential V_m of the LIF model is expressed as

$$RC\frac{dV_m(t)}{dt} = -(V_m(t) - V_r) + R\big(I_e(t) + I_i(t)\big). \quad (1)$$

The behavior at spike generation is given by

$$V_m(t) = V_R \quad \text{if} \quad V_m(t) > V_t. \quad (2)$$

Here, $I_e(t)$ and $I_i(t)$ represent excitatory postsynaptic currents (EPSCs) and inhibitory postsynaptic currents (IPSCs), respectively. C and R represent the membrane capacitance and resistance. V_r represents the resting membrane potential.

EPSCs increase the membrane potential, and the neuron generates a spike when the potential matches the condition expressed in Eq. (2). Generating a spike is called firing. After firing, the membrane potential decreases to the reset potential V_R , and the neuron enters the refractory period, during which the membrane potential does not change. In this study, we examined the relationship between time constant $\tau(= RC)$ and learning results.

EPSCs and IPSCs are calculated by the following equations:

$$I_e(t) = g_e\big(E_e - V_m(t)\big) \quad (3)$$
$$I_i(t) = g_i\big(E_i - V_m(t)\big) \quad (4)$$

g_e and g_i represent the excitatory and inhibitory synaptic conductance, respectively. E_e and E_i represent excitatory and inhibitory equilibrium potentials, respectively.

The model of input neurons is a Poisson process model. The number of spikes generated by the Poisson process model follows a Poisson distribution. The probability that spikes fire n times in interval Δt is given by

$$P[N(\Delta t) = n] = \frac{(\lambda\Delta t)^n}{n!}e^{-\lambda\Delta t}, \quad (5)$$

where λ represents the average rate at which spikes fire.

2.2. *Model of synapses*

We used a single exponential model for synapses. This model reproduces the exponential decay in the postsynaptic current after a spike firing of the presynaptic neuron. The exponential decay was applied to synaptic conductance g_e and g_i .

2.3. *Learning rule*

Many SNNs use the spike-timing dependent plasticity (STDP) [4], which models the biological synaptic plasticity, as a learning rule. The STDP changes the synaptic weights depending on the time difference between the firing of the presynaptic neuron and the postsynaptic neuron. The time difference causes a long-term potentiation (LTP) or a long-term depression (LTD) of the synaptic weights. This rule can train neurons that are selective to a particular set of stimuli.

In this study, synaptic weights of inhibitory inputs are strengthened by LTPi [2]. This learning rule strengthens inhibitory synaptic weights if the following two conditions are matched. First, the membrane potential of the postsynaptic neuron is higher than a certain value at the firing time of the presynaptic neuron. Second, the postsynaptic neuron does not fire during the time period just before and after the firing of the presynaptic (inhibitory) neuron. Under the conditions described above, inhibition is regarded as effective, and the inhibitory synaptic weight is strengthened. This behavior enables learning of a combination of features [2]. The weights of the inhibition (W) are updated by adding a constant value ΔW as the following equation:

$$W \leftarrow W + \Delta W. \quad (6)$$

3. Spiking Neural Network Configuration

Fig. 1 shows the neuronal network designed in this study. Each cluster with circles represents a group of neurons. The neural network consists of three layers: an input layer, an inhibition layer, and an output layer.

Input neuron groups A, B, and C are assumed to be groups of neurons that respond to three particular types of features. Each neuron in the output and inhibition layers is connected to two of the three groups of neurons in the input layer. Output neuron group α_o and inhibitory neuron group α_i are connected to groups B and C, β_o and β_i are connected to groups A and C, and γ_o and γ_i are connected to groups A and B. Each group in the inhibitory layer is connected to all the groups in the output layer. The output layer receives excitatory inputs from the input layer and inhibitory inputs from the

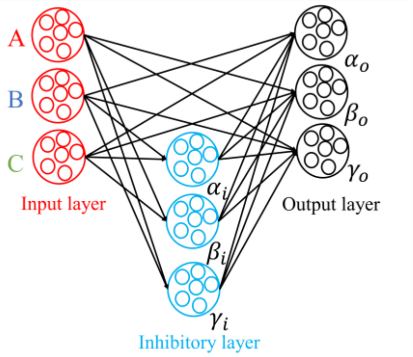


Fig. 1. Network structure. The number of neurons in the input, inhibitory, and output layers are 50, 10, and 30, respectively.

inhibitory layer. Only the weights between the inhibitory and output layers are updated by LTPi.

We set a probability of connection between two neuron groups. The probability between neurons in the input layer and the output layer is 50 %, that between the input layer and the inhibitory layer is 30 %, and that between the inhibitory layer and the output layer is also 30 %.

Because of the network topology, output neurons are selective to a set of input from the beginning. We examined whether the selectivity increases or decreases after learning.

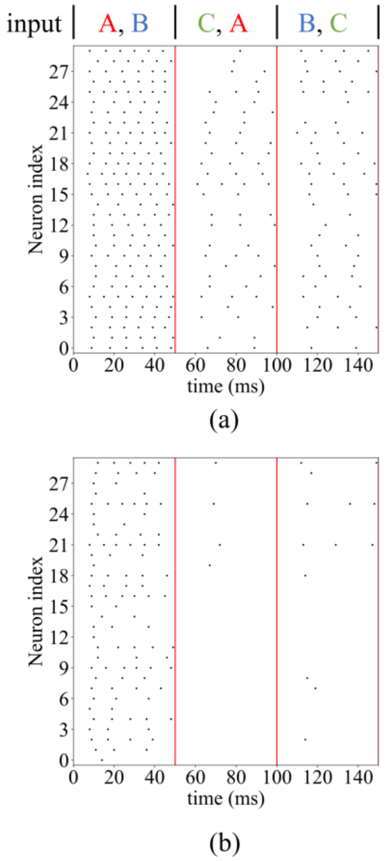
4. Experiments and Results

4.1. *Training of the neural network*

To train the SNN, a pair of neuron groups in the input layer were activated for a certain period (100 ms). The firing rate was approximately 100 Hz. The pair of the groups was selected randomly, and the pair was changed every 100 ms. The membrane potential of each neuron was reset to its resting membrane potential when the pair was changed. The training period of 100 ms was repeated 100 times.

4.2. *Network output*

Fig. 2 shows a spike train of all neurons in the neuron group γ_o before and after training. The activated pair was changed every 50 ms in the test. The firing frequency was high during the period from 0 to 50 ms when the activated pair of the group was (A, B) because the pair is connected to γ_o. After training, the overall firing frequency decreased, but the decrease in the firing frequency during the period from 50 to 150 ms was more significant. Quantitative analysis is provided in the next section.



(a)

(b)

Fig. 2. Spike train of neurons in γ_o. (a) (b) show spikes before and after training, respectively. The time of firing is represented by a dot. The red line indicates the time of the change in the input pair, with an interval of 50 ms. The activated pair of neuron groups at each interval are shown in the upper part of the figure.

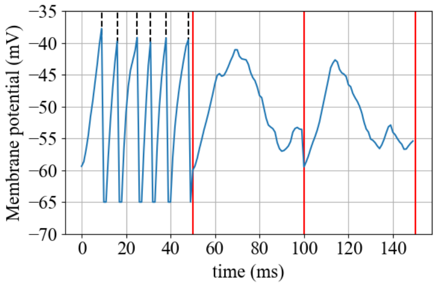


Fig. 3. Membrane potential of the output neuron whose index is nine in Fig. 2(b) after training.

Fig. 3 shows the membrane potential of neuron whose index is nine in Fig. 2(b). The vertical dashed line

Table 1 Average change in the synaptic weights before and after training. $\boldsymbol{W}_{\mathbf{pre}}$ and $\boldsymbol{W}_{\mathbf{post}}$ represent average weights before and after training, respectively.

Connection	$\boldsymbol{W}_{\mathbf{pre}}$	$\boldsymbol{W}_{\mathbf{post}}$	$\boldsymbol{W}_{\mathbf{post}} - \boldsymbol{W}_{\mathbf{pre}}$
$\boldsymbol{\alpha_i \to \gamma_o}$	0.073	0.449	0.376
$\boldsymbol{\beta_i \to \gamma_o}$	0.075	0.511	0.436
$\boldsymbol{\gamma_i \to \gamma_o}$	0.077	0.397	0.320

represents the time of firing. The membrane potential decreased significantly during the period from 50 to 150 ms because of the inhibitory input, resulting in preventing the neuron from firing. This result suggests that the selectivity increased.

Table 1 shows the average change of synaptic weights between neuron group γ_o and three neuron groups in the inhibitory layer before and after training. The change in weights between γ_i and γ_o was the smallest.

4.3. *Selectivity after training*

To evaluate the selectivity of neurons, we introduced a criterion that is given by the following equation:

$$S = \frac{1}{N}\sum_{n=1}^{N}\frac{F_i(n)}{F_{\mathrm{AB}}(n) + F_{\mathrm{BC}}(n) + F_{\mathrm{CA}}(n)}, \qquad (10)$$

where $F_i(n)$ ($i \in \{\mathrm{AB}, \mathrm{BC}, \mathrm{CA}\}$) represents the firing rate of an output neuron when the activated pair of neuron groups is i. N is the number of neurons in the neuron group. n represents the neuron index. The pair i of the numerator is the pair that the output neurons have the direct connection; for example, pair AB for neuron group γ_o.

We calculated the change in the selectivity criterion S for each neuron group. The amount of change is given by

$$\Delta S = S_{\mathrm{post}} - S_{\mathrm{pre}}, \qquad (11)$$

where S_{post} and S_{pre} represent the criterions after and before training, respectively.

Fig. 4 shows the relationship between the time constant of the inhibitory neurons and ΔS. Here, the time constant of the output neuron was set to 20 ms. We examined ΔS 10 times using different initial weights determined randomly. The value shown in Fig. 4 is the average of ΔS obtained in the 10 trials. The amount of increase in selectivity strongly depends on the time constant of the inhibitory neurons, and the time constant around 30 ms induced the largest increase in selectivity.

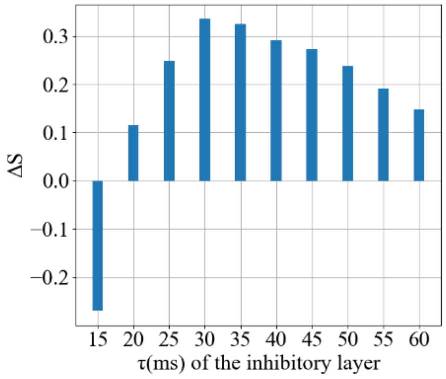


Fig. 4. Mean increase in selectivity of output neurons in γ_o

The increase in selectivity decreased gradually for time constants that is longer than 30ms. This indicates that appropriate parameter settings are necessary for increasing selectivity.

5. Conclusion

In this study, we designed a small spiking neural network that learns a combination of features by using LTPi. The selectivity examination showed that the amount of increase in selectivity strongly depends on the time constant of the inhibitory neurons.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number 19K12916.

References

1. A. Maffei, K. Nataraj, S. B. Nelson, and G. G. Turrigiano, “Potentiation of cortical inhibition by visual deprivation”, Nature, vol. 443, pp. 81-84, September 2006.
2. M. A. Bourjaily, and P. Miller, “Synaptic Plasticity and Connectivity Requirements to Produce Stimulus Pair Specific Responses in Recurrent Networks of Spiking Neurons”, PLOS Computational Biology, vol. 7, no. 2, pp. 1-18, February 2011.
3. C. Koch, “Biophysics of Computation: Information Processing in Single Neurons”, Oxford University Press, 1998
4. G. Bi and M. Poo, “Synaptic Modifications in Cultured Hippocampal Neurons: Dependence on Spike Timing, Synaptic Strength, and Postsynaptic Cell Type”, Journal of Neuroscience, vol. 18, no. 24, pp. 10464-10472, December 1998.

Authors Introduction

Mr. Mahiro Ikeda

He received his B.S. degree from the Department of Information Science and Technology, Osaka Institute of Technology, Japan in 2022. He is currently a Master's course student in Osaka Institute of Technology, Japan.

Dr. Hirotsugu Okuno

He received the Ph.D degree in electrical, electronic and information engineering from Osaka University, in 2008. He is currently an Associate Professor at the Faculty of Information Science and Technology, Osaka Institute of Technology. His research interests include visual information processing in the nervous system and their applications to robotics.

Binocular Disparity Estimation Algorithm Using Multiple Spatial Frequency Information and a Neural Network

Ryoka Sato
Graduate School of Information Science and Technology, Osaka Institute of Technology,
1-79-1 Kitayama, Hirakata, Osaka 573-0196, Japan

Hirotsugu Okuno
Faculty of Information Science and Technology, Osaka Institute of Technology,
1-79-1 Kitayama, Hirakata, Osaka 573-0196, Japan
E-mail: hirotsugu.okuno@oit.ac.jp
www.oit.ac.jp

Abstract

We developed a disparity estimation algorithm that uses a disparity energy, which is computed based on a model of neurons responding selectively to a particular disparity. Disparity energy values depend on the spatial frequency of input images because of the Gabor filters employed. To reduce the frequency dependency, our algorithm uses disparity energy values computed from multiple spatial frequencies for neural-network-based regression. The algorithm was successful in estimating a disparity from images with a certain range of spatial frequencies.

Keywords: neural network, image processing, binocular disparity, disparity energy, visual cortex

1. Introduction

Depth information is one of the most important information for animals, humans, and robots to move in real environments. The visual nervous system can acquire depth information from disparity, which is the difference in the position of images reflected on the left and right retinas. Depth estimation from binocular disparity consists of the following two procedures: searching for corresponding points in the left and right images, and calculation of the gap between the corresponding points. Similar feature points located close to each other can cause incorrect matching between the left and right images, which leads to a failed depth estimation. Therefore, it is important to reduce candidates for the corresponding points with a similar feature.

The visual nervous system reduces candidates with similar features by applying a spatial bandpass filter to images on the retina. The disparity energy model, which is a model of the visual nervous system that estimate a corresponding points utilizing a bandpass filter, has been proposed based on physiological studies [1], and output neurons of the model respond selectively to a particular disparity.

However, the disparity energy model, which uses a single bandpass filter, has a problem with disparity estimation in regions where the amplitude of the spatial band of the bandpass filter is small. To overcome this drawback, previous studies proposed a use of disparity energies computed using multiple bandpass filters [2],[3].

The purpose of this study is to develop a disparity estimation algorithm that uses neural-network-based regression whose input signals are disparity energy values with multiple spatial frequency information.

2. Disparity Energy Model

The disparity energy model is a computational neuronal network model that explains the activity of cortical neurons responding selectively to a particular binocular disparity.

Fig. 1 shows the processing flow of the disparity energy model used in this study. The disparity energy is computed from two sets of images, which are acquired by the left and right eyes and are filtered by four Gabor filters whose phases are different by $\pi/2$ radian. The disparity energy I is given by:

$$\begin{aligned} I(x,y,\phi) = \{&G(x,y,0) * L(x,y) \\ &+G(x,y,\phi) * R(x,y)\}^2 \\ &+\{G\left(x,y,\frac{\pi}{2}\right) * L(x,y) \\ &+G\left(x,y,\phi+\frac{\pi}{2}\right) * R(x,y)\}^2. \end{aligned} \tag{1}$$

$G(x,y,\phi)$ represents a Gabor filter with phase offset ϕ. $L(x,y)$ and $R(x,y)$ represents the left and right input images at coordinates (x,y). The disparity energy is maximized when the phase offset ϕ of the Gabor filter matches the disparity of the input image. Therefore, phase ϕ that maximizes I is the estimated disparity in radians. The estimated disparity d in pixels are given by

$$d = \frac{\phi l}{2\pi}, \tag{2}$$

where l is the wavelength of the sinusoidal function in the Gabor filter. The estimated disparity depends on the frequency, and therefore, the detectable range also depends on the wavelength; the range is $-l/2 < d < l/2$.

Fig. 2 shows an example of the output of the disparity energy model when the input images are grating images and when the right image is shifted by +3 pixels with respect to the left image. The wavelength l of the Gabor filter used is 12 pixels, and the shift of the right image corresponds to $\pi/2$. Therefore, disparity energy $I(\pi/2)$ has the maximum value in a large region among all disparity energy images. This result tells that the estimated disparity in radians is $\pi/2$.

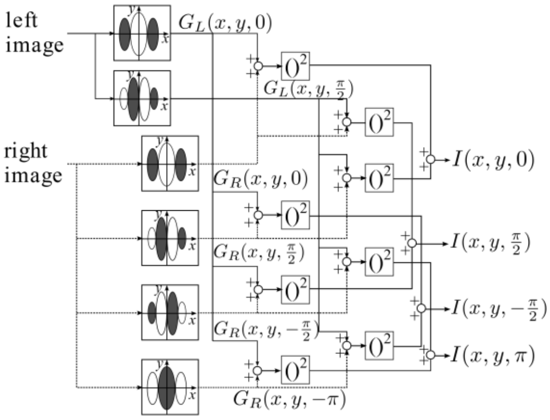


Fig. 1. Processing flow of the disparity energy model used in this study. The black and white ellipses represent the weights of the Gabor filter. The phase φ of the Gabor filter used is (0, $\pi/2$) for the left image and (0, $\pi/2$, $-\pi/2$, $-\pi$) for the right image.

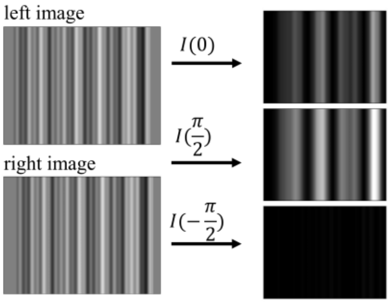


Fig. 2. Examples of disparity energy model output when the input is grating images. The right image is shifted by +3 pixels with respect to the left image. The right column shows disparity energy $I(0)$, $I(\pi/2)$, and $I(-\pi/2)$. Their corresponding disparities in pixels are $d = 0$, +3, and -3, respectively.

3. Disparity Estimation Algorithm

3.1. *Processing flow*

Fig. 3 shows the processing flow of the disparity estimation algorithm proposed in this study. First, the disparity energy is computed for two spatial frequencies using Gabor filters whose wavelength is 16 pixels and 32 pixels. Hereinafter, these energy values are referred to as high-frequency disparity energy I_h and low-frequency disparity energy I_l. These are transferred to the neural network after normalized respectively.

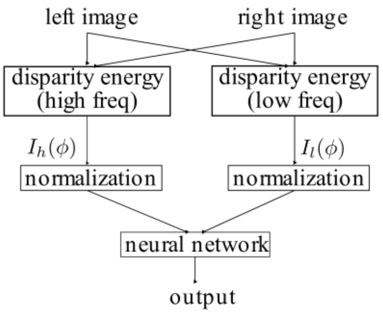


Fig. 3. Processing flow diagram of the proposed disparity estimation model with neural- network-based regression. I_h and I_l represents energy computing using high and low frequencies, respectively.

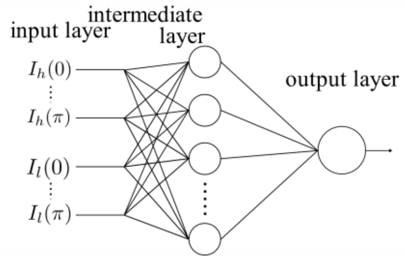


Fig. 4. Neural network used for regression. The number of neurons in the intermediate layer is 20.

3.2. *Neural-network-based regression*

Disparity estimation was performed using neural-network-based regression whose input signals are high-frequency and low-frequency disparity energies. Fig. 4 shows the neural network topology used for regression. The output of the neural network is the estimated disparity.

We used sinusoidal wave grating images as shown in Fig. 2 as training data of the neural network. The function of the wave is given by

$$S(x) = \sum_{i=0}^{N} A_i \sin\left(\frac{2\pi x}{\lambda_i} + \theta_i\right), \tag{3}$$

where N in an integer from 7 to 15. We generated a pair of grating images; one is the wave grating generated by Eq. (3) directory, and the other is the image obtained by shifting the first image along x axis. The second one simulates an image with a certain disparity. We generated 100 images for each disparity ($-6<d<+6$; disparity d is an integer) and extracted 100 points from each image to prepare training data, resulting in 10,000 data points for each disparity.

4. Experiments and Results

4.1. *Experimental environment*

We implemented the model described in the previous section and trained the neural network using Python and Pytorch. The trained models were evaluated in the following two ways. First, we evaluated the model by using wave grating images generated by Eq. (3). The generated data were separated into training data and test data. Second, we evaluated the model by using images acquired by a binocular camera system.

4.2. *Evaluation using grating images*

We used two types of regression models to examine effects of using multiple frequencies: a model using only low-frequency disparity energy and a model using both frequencies. We compared the estimation results at 10,000 points in the grating images.

Fig. 5 shows the estimation results of the regression model that uses only low-frequency disparity energy. Fig. 5(a) shows that most of the estimated disparity is close to the correct disparity. The absolute error is smaller for a small disparity. Fig. 5(b) shows that all the absolute errors were within 0.5 pixels when the correct disparity is 0, but the error increased for a larger correct disparity.

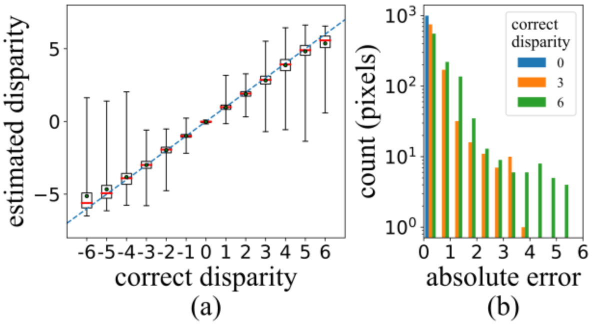


Fig. 5. Estimation results of the regression model that use only low-frequency disparity energy. (a) Box-and-whisker plot of the estimation result. The horizontal and vertical axis represent the correct and the estimated disparity, respectively. The red line and green dots show the median and mean of the estimated disparities, and the dashed line plots the correct disparity. (b) Histogram of the absolute errors. The histogram has 12 bins in the range from 0, to 6 pixels.

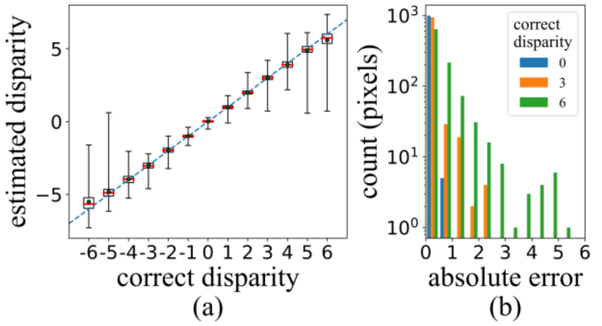


Fig. 6. Estimation results of the regression model using both low and high frequency disparity energies. The method of creating the figures is the same as that in Fig. 5.

Fig. 6 shows the estimation results of the regression model that uses both low and high frequency disparity energies information. Error conditions are similar to Fig. 5, but the error is smaller for a larger correct disparity compared to Fig. 5.

4.3. *Evaluation using a binocular camera system*

Fig. 7 shows the experimental environment used to investigate the disparity estimated from images obtained by a binocular camera system. A sinusoidal wave grating image was placed in front of the binocular system.

Fig. 8 shows the disparity map image estimated by the model that was trained to obtain Fig. 6. Although apparent spatial frequency of the image changes depending on the distance to the object, the results in Fig. 8(a)(b)(c) show that the proposed model estimated the correct disparity irrespective of the input spatial frequency.

5. Conclusion

In this study, we developed an algorithm that uses the disparity energy and neural-network-based regression. The simulation results showed that the use of two spatial frequencies is effective reducing errors. The results obtained by the binocular camera experiments suggest that our method is applicable to robotic vision.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number 19K12916.

References

1. I. Ohzawa, G.C. DeAngelis, and R.D. Freeman, "Stereoscopic depth discrimination in the visual cortex: neural ideally suited as disparity detectors", Science, vol. 249, No.4972, pp.1037-1041, Aug, 1990.
2. D. J. Fleet, H. Wagner, and D. J. Heeger, "Neural encoding of binocular disparity: energy models, position shifts and phase shifts", Vision Research, Vol. 36, pp. 1839-1857, June, 1996.
3. J.J. Tsai, and J.D. Victor, "Reading a population code: a multi-scale neural model for representing binocular disparity", Vision Research, Vol. 43, pp. 445-466, Feb, 2003.

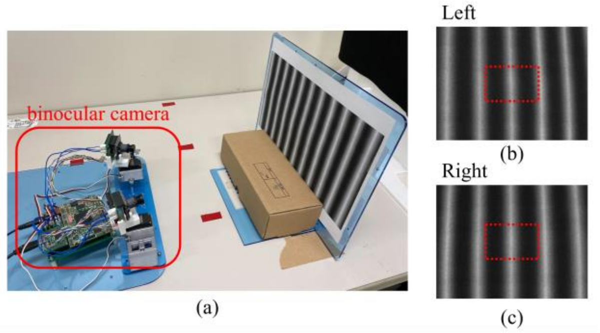


Fig. 7. (a) Experimental environment. A binocular camera system acquires the grating image. (b)(c) Left and right images acquired by the system. The estimated disparity within the red dashed rectangle in (b)(c) was used to obtain the disparity map image in Fig. 8.

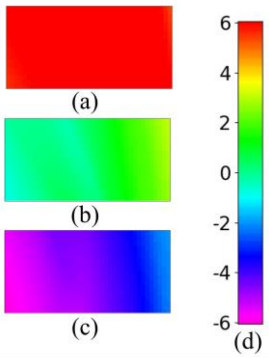


Fig. 8. Disparity map estimated by the proposed model in the dashed red rectangle of Fig. 7(b)(c). (a)(b)(c) Disparity map when the distance to the grating image is far, intermediate, and near, respectively. The color represents estimated disparity, and the correspondence between colors and estimated values of the disparity is shown in (d).

Authors Introduction

Mr. Ryoka Sato

He received his B.S. degree from the Department of Information Science and Technology, Osaka Institute of Technology, Japan in 2022. He is currently a Master's course student in Osaka Institute of Technology, Japan.

Dr. Hirotsugu Okuno

He received the Ph.D degree in electrical, electronic and information engineering from Osaka University, in 2008. He is currently an Associate Professor at the Faculty of Information Science and Technology, Osaka Institute of Technology. His research interests include visual information processing in the nervous system and their applications to robotics.

A Figure-Ground Discrimination Algorithm Inspired by Border-Ownership Selective Cells

Tomoya Kobayashi
Graduate School of Information Science and Technology, Osaka Institute of Technology,
1-79-1 Kitayama, Hirakata, Osaka 573-0196, Japan

Hirotsugu Okuno
Faculty of Information Science and Technology, Osaka Institute of Technology,
1-79-1 Kitayama, Hirakata, Osaka 573-0196, Japan
E-mail: hirotsugu.okuno@oit.ac.jp
www.oit.ac.jp

Abstract

We developed a figure-ground discrimination algorithm that is suitable for robot vision. The algorithm is inspired by the response properties of border-ownership (BO)-selective cells, which possibly support figure-ground discrimination in the visual nervous system. The output of our algorithm is affected by spatial parameters such as the characteristics of the spatial filter used for edge enhancement and the interval of model BO-selective cells. Therefore, we investigated the relationship between the algorithm output with various spatial parameters and object sizes.

Keywords: figure-ground discrimination, border-ownership, bio-inspired.

1. Introduction

Figure-ground discrimination is one of the most important functions for image classification. Although a lot of algorithms for extracting contour has been developed in robotic vision field, it is not straightforward to extract information on which side of the contour the object is located; this information is necessary for figure-ground discrimination.

On the other hand, animals and humans have an ability to discriminate figure and ground as one of their basic visual functions, and can instantly discriminate objects of interest from other backgrounds based on visual information projected on their retina. Border-ownership (BO)-selective cells, which respond selectively to contour only on the object side, could be involved in the function of figure-ground discrimination in the visual nervous system.[1] Several neuronal network models for figure-ground discrimination have been proposed based on the response characteristics of BO-selective cells.[2], [3]

The response characteristics of BO-selective cell can play an important role in image classification. However, the computational cost of the neuronal network models themselves are expensive, and they are not suitable for real-time processing. Algorithms with less computational cost are required for real-world and real-time applications.

The purpose of this study is to develop a figure-ground discrimination algorithm inspired by BO-selective cells that is suitable for robot vision. We also investigated the relationship between the parameters used in this algorithm, such as the spatial filter, and the object's size that can be discriminated against the background by this algorithm.

2. Neuronal Model for Figure-Ground Discrimination

The algorithm developed in this study is based on a neuronal circuit model proposed to explain the figure-ground discrimination function in the visual nervous system.[2],[3] Fig. 1 shows the structure of the neuronal circuit model that consists of three layers. The first layer is the simple cell layer (S_θ in Fig. 1), in which each cell responds selectively to contours of a particular orientation. The second layer is BO-selective cell layer, in which each cell responds selectively to contours and the response of cells depends on which side the object is. The third layer is grouping cell layer, in which cells in object regions respond strongly.

The BO-selective cells receive signals from S_θ and therefore, these cells also respond selectively to contours. In the BO-selective cell layer, a pair of cells located at a position works together and one of the paired cells on a contour respond strongly depending on which side the object is. When the object is on the opposite side, the other cell of the paired cells responds strongly. Therefore, comparing the response of the paired cells tells the side of the object.

The interconnection between BO-selective cells and grouping cells enables the object-side-specific response of BO-selective cells. Each grouping cell receives input from multiple BO-selective cells to determine whether the area is enclosed by a contour. Grouping cells surrounded by a contour show a strong response. Grouping cells also send either excitatory or inhibitory feedback to BO-selective cells. The feedback enhances the response of BO-selective cells that are selective to the correct object direction and inhibits BO-selective cells that are selective to the opposite direction.

3. Figure-Ground Discrimination Algorithm

3.1. *Processing flow*

Fig. 2 shows the processing flow of the figure-ground discrimination algorithm developed in this study.

First, contour features (S_θ) of four orientations are extracted from grayscale input images using a set of Gabor filters.

Next, the contour information is sent to two types of BO-selective cells ($B_{\theta P}$ and $B_{\theta N}$ in Fig. 2 for examples) after half-wave rectification. $B_{\theta P}$ has a facilitative interconnection with grouping cells (G_L), which responds strongly to the contour of objects that are brighter than the background. $B_{\theta N}$ has a facilitative interconnection

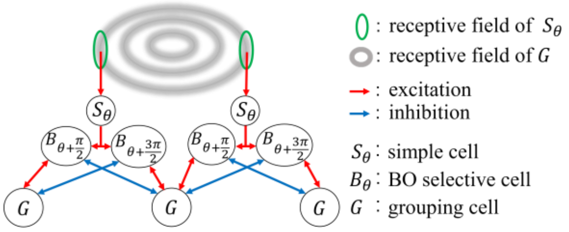

S_θ
$B_{\theta+\frac{\pi}{2}}$
$B_{\theta+\frac{3\pi}{2}}$
G
: receptive field of S_θ
: receptive field of G
: excitation
: inhibition
S_θ : simple cell
B_θ : BO selective cell
G : grouping cell

Fig. 1. Neuronal model for figure-ground discrimination.

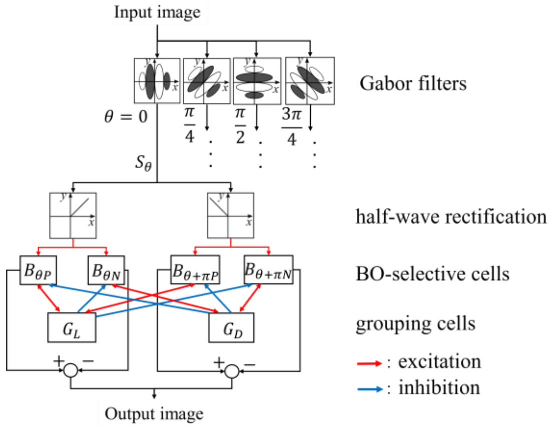

Input image
Gabor filters
$\theta = 0$
$\frac{\pi}{4}$
$\frac{\pi}{2}$
$\frac{3\pi}{4}$
S_θ
half-wave rectification
$B_{\theta P}$
$B_{\theta N}$
$B_{\theta+\pi P}$
$B_{\theta+\pi N}$
BO-selective cells
G_L
G_D
grouping cells
: excitation
: inhibition
Output image

Fig. 2. Processing flow diagram of the proposed algorithm.

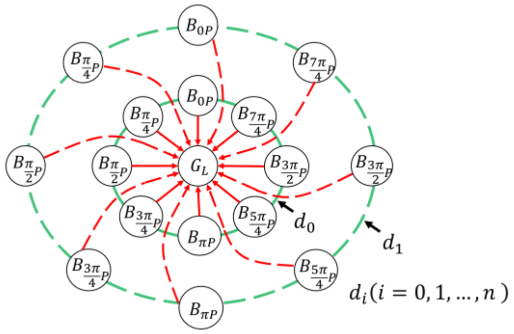

B_{0P}
$B_{\frac{\pi}{4}P}$
$B_{\frac{7\pi}{4}P}$
$B_{\frac{\pi}{2}P}$
G_L
$B_{\frac{3\pi}{2}P}$
$B_{\frac{3\pi}{4}P}$
$B_{\frac{5\pi}{4}P}$
$B_{\pi P}$
d_0
d_1
$d_i(i = 0, 1, \dots, n)$

Fig. 3. Feedforward connection between BO-selective cells and grouping cells.

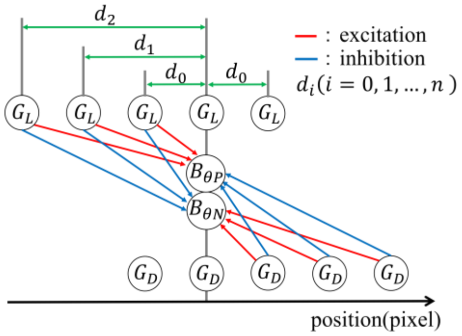

d_2
d_1
d_0
d_0
: excitation
: inhibition
$d_i(i = 0, 1, \dots, n)$
G_L
$B_{\theta P}$
$B_{\theta N}$
G_D
position(pixel)

Fig. 4. Feedback connection between grouping cells and BO-selective cells.

with grouping cells (G_D), which responds strongly to the contour of objects that are darker than the background.

In the neuronal circuit model of the previous study, grouping cells are connected to BO-selective cells on their ring-shaped receptive field. In this study, the connection between grouping cells and BO-selective cells were simplified to reduce computational complexity. Each grouping cell connected to BO-selective cells that are in the eight directions whose angles are different by $\pi/4$, and are at distances of $d_i (i = 0,1, \ldots, n)$ pixels.

Fig. 3 and Fig. 4 show the connections between grouping cells and BO-selective cells. Fig. 3 shows the connections of $B_{\theta P}$ at different distances d_i with respect to G_L. These connections simulate the ring-shaped receptive field of grouping cells in the original neuronal circuit models[2],[3]. The connection between $B_{\theta N}$ and G_D can be illustrated by replacing G_L and each $B_{\theta P}$ in Fig. 3 with G_D and $B_{\theta N}$. Fig. 4 shows the feedback connections between grouping cells and BO-selective cells.

Finally, an output image that indicate the direction of the figure with respect to the background is generated from the response of BO-selective cells.

3.2. *Contour extraction*

The proposed algorithm uses a set of odd-function Gabor filters for contour extraction. The contour orientation (θ) to be extracted is four orientations whose angle differences are π/4. The Gabor function is expressed by

$$g(x, y, \theta, L) = A \exp\left(-\frac{x'^2 + y'^2}{2\sigma^2}\right) \cos\left(\frac{2\pi x'}{L} + \varphi\right), \quad (1)$$

$$x' = x\cos\theta + y\sin\theta \,, \quad (2)$$

$$y' = -x\sin\theta + y\cos\theta \,, \quad (3)$$

where x and y represent the coordinates in the kernel, L represents the wavelength. This algorithm uses three filters whose (σ, L) sets are (2.2, 6), (3.1, 8), and (4.5, 12) pixels. The response of a simple cell (S_θ) is expressed by

$$S_\theta(x, y) = I(x, y) * g(x, y, \theta, L) \,. \quad (4)$$

3.3. *BO-selective cells*

The contour information obtained in the previous section is used to compute the response of two types of BO-selective cells ($B_{\theta P}, B_{\theta N}$). The response of $B_{\theta P}$ is expressed by

$$B_{\theta P}(x, y) = S_\theta(x, y) \times \frac{2}{1 + \exp(-G)} \,, \quad (5)$$

$$G = \sum_{i=0}^{n} \frac{d_0}{d_i} \cdot \left(G_L(x + \Delta x_i, y + \Delta y_i) - G_D(x - \Delta x_i, y - \Delta y_i)\right), \quad (6)$$

$$(\Delta x_i, \Delta y_i) = (d_i \cos\theta, -d_i \sin\theta) \,. \quad (7)$$

The strength of the interaction between BO-selective cells and grouping cells is represented by d_0/d_i, and reduces as distance d_i increases. The response of $B_{\theta N}$ is obtained by just replacing the coordinate of G_D with that of G_L in Eq. (6). The way of computing responses of G_L and G_D is explained in the next section.

3.4. *Grouping cells*

Responses of two types of grouping cells are computed from the responses of BO-selective cells. The response of grouping cells G_L, which respond strongly to objects that are brighter than the background, is obtained by summing $B_{\theta P}$ in the eight directions at a set of distances, as shown in Fig 3. The response of grouping cells G_D, which respond strongly to objects that are darker than background, is obtained by summing $B_{\theta N}$ in the same way as G_L.

Finally, the strengths of the responses of a set of grouping cells (G_L, G_D) at each position are compared, and the smaller one is set to 0 because an object is either darker or brighter than the background.

4. Results

4.1. *Algorithm outputs*

We examined the algorithm outputs using the input image shown in Fig. 5(a). The following parameter was used for this simulation: $d_i \in \{10,14,18,22\}$.

Fig. 5(b) shows the final output whose color indicates the object orientation. The orientation was determined by the response of BO-selective cells computed using a Gabor filter whose $L = 8$ pixels. The color circle on the upper right shows the correspondence between the color and the orientation. On most of the contour, the object orientation was correctly indicated.

Figs. 5(c), (d), and (e) show the sum of the responses of grouping cells G_L and G_D computed using Gabor filters with various values of L. In Fig. 5(d), which was used to obtain Fig. 5(b), cells only in the object region responded strongly.

Our algorithm implemented on a personal computer with an Intel(R) Core(TM) i7-8700 CPU and 16 GB RAM using Python processed one image in about 40 msec.

4.2. *Examination of grouping cell characteristics*

The object's size that can be discriminated by the proposed algorithm is roughly estimated from responses of grouping cells. Responses of grouping cells are strongly affected by the connection between BO-selective cells and grouping cells, and the spatial characteristics of the Gabor filter. Therefore, we examined the relationship between these spatial parameters and the response of grouping cells.

The input image used here has a solid white circle on a black background. The radius of the circle is between 1 to 30 pixels. We examined the response of the grouping cell at the center of the circle.

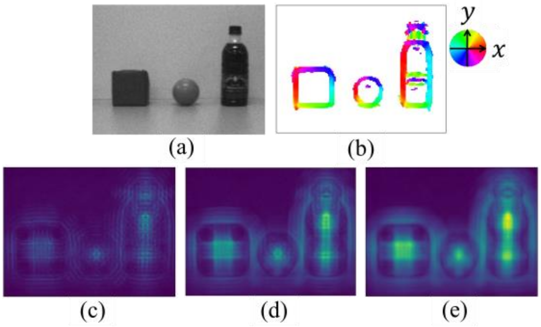
Fig. 6 shows the response of the grouping cell at the center of the circle. The following parameter was used for this examination: $d_i \in \{10,14,18,22,26\}$. Because the circle is brighter than the background, G_L (Fig. 6(a)) should always be larger than G_D (Fig. 6(b)). However, G_D has a larger value in some cases. First, G_D is larger when the radius of the circle is smaller than 8 pixels. This is because d_i was set to values that are larger than 10 pixels and the smallest object to be discriminated can be controlled by changing d_i. Second, G_D is larger when the wavelength L is 6 pixels and the radius is about 13, 17, and 20 pixels. This case is caused by constructive interference of two Gabor waves induced by contour of opposite sides. To avoid this situation, the interval of d_i has to be less than half of the wavelength L.

5. Conclusion

In this study, we developed a figure-ground discrimination algorithm based on the model of BO-selective cells in the visual nervous system. We also investigated the relationship between the spatial parameters used in the algorithm and the object's size that can be discriminated into figures and the background by the algorithm.

6. Acknowledgement

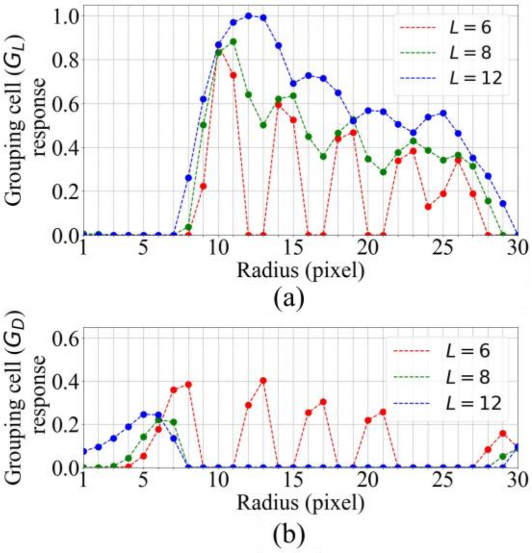
This work was supported by JSPS KAKENHI Grant Number 19K12916.



(a) (b)

(c) (d) (e)

Fig. 5. Examples of input and output images of the algorithm. (a)Input image of 160 × 120 pixels. (b) Response of BO-selective cells computed using a Gabor filter whose $L = 8$ pixels. (c)(d)(e) Response of grouping cells computed using Gabor filters whose $L = 6$, 8, and 12 pixels, respectively.



(a)

(b)

Fig. 6. The relationship between object's size and the response of the grouping cell at the center of the circle. The horizontal axis represents the radius. The red, green, and blue lines plot the response of the grouping cell computed using Gabor filters whose $L = 6$, 8, and 12 pixels, respectively. (a) Response of G_L. (b) Response of G_D.

References

1. H. Zhou, H. S. Friedman, and R. von der Heydt, "Coding of Border Ownership in Monkey Visual Cortex," The Journal of Neuroscience, Vol. 20, No. 17, pp.6594-6611, 2000.
2. E. Craft, H. Schütze, E. Niebur, and R. von der Heydt, "A Neural Model of Figure-Ground Organization," Journal of Neurophysiology, Vol. 97, No. 6, pp.4310-4326, 2007.

3. B. Hu, R. von der Heydt, and E. Niebur, "Figure-Ground Organization in Natural Scenes: Performance of a Recurrent Neural Model Compared with Neurons of Area V2," eNeuro, Vol. 6, No. 3, ENEURO.0479-18, 2019.

Authors Introduction

Mr. Tomoya Kobayashi

He received his B.S. degree from the Department of Information Science and Technology, Osaka Institute of Technology, Japan in 2021. He is currently a Master's course student in Osaka Institute of Technology, Japan.

Dr. Hirotsugu Okuno

He received the Ph.D degree in electrical, electronic and information engineering from Osaka University, in 2008. He is currently an Associate Professor at the Faculty of Information Science and Technology, Osaka Institute of Technology. His research interests include visual information processing in the nervous system and their applications to robotics.

Event-Driven Particle Filter for Tracking Irregularly Moving Objects

Yuki Kawasaki
Department of Human Intelligence Systems, Kyushu Institute of Technology,
2-4, Hibikino, Wakamatsu, Fukuoka, 808-0196, Japan

Masahiro Ohtani
Department of Electrical Engineering, National Institute of Technology (KOSEN), Nara College,
22, Yata-chou, Yamatokooriyama, Nara, 639-1080, Japan

Shinsuke Yasukawa
Department of Human Intelligence Systems, Kyushu Institute of Technology,
2-4, Hibikino, Wakamatsu, Fukuoka, 808-0196, Japan
E-mail: s-yasukawa@brain.kyutech.ac.jp
https://www.kyutech.ac.jp/

Abstract

Conventional object tracking techniques that use general-purpose cameras and particle filters find it difficult to track irregularly and rapidly moving objects. To track an irregularly moving object without losing sight, quickly measuring the position of the object is necessary. In this study, we used a fast-response event-based camera, which is a bioinspired camera that produces a spiking output. We propose an event-driven particle filter that performs processing in response to the input from an event-based camera. Our proposed method was evaluated by presenting an event-based camera with a rectangular motion pattern that combines periodic and constant-velocity motions at various speeds. The experimental results demonstrated that our proposed method could track objects in a test video.

Keywords: Particle filter, Event-based camera, Robot vision

1. Introduction

Visual servo technology is a key technology for controlling a robot by visually detecting changes in the state of a manipulated object. A visual servo requires short sampling intervals of visual information to accurately identify the dynamics of the physical system of the target and track objects. For example, to achieve a pan–tilt system that constantly captures flying objects in its field of view or technology that allows a manipulator to serve spaghetti non-stop on a plate, visually capturing the movement of the object and applying it to the control of the robot is necessary. There is always a delay between the acquisition of visual information and the command of the robot. Therefore, estimating the post-motion state of the target as accurately as possible based on the constraints of the target's physical system is also necessary.

When capturing motion information from the outside world, problems with motion blur are unavoidable with general-purpose cameras having exposure processes. Therefore, event-based cameras [1], which outputs only pixel-by-pixel luminance change information, has attracted attention owing to its high temporal resolution and low-latency response properties. In recent years, considerable research has been performed on computer vision technology using event-based cameras [2].

In this study, we also focused on particle filters, as they can stably predict states using a large number of particles for object tracking. Generally, in image-based object tracking, particle filter calculation cycles are performed in a frame-by-frame manner.

Herein, we propose a new event-driven particle filter that performs an operational cycle for each event.

Next, a method for tracking objects in pulsating motion is proposed by integrating an event-driven camera operating at high speed and low latency with an event-driven particle filter. The proposed method was evaluated by presenting a camera with a test video on a display.

2. Proposed method

High-speed cameras with high temporal resolution and low latency are required to track objects that constantly change their motions. Event-based cameras are sensors that asynchronously detect changes in the brightness of light incident on each pixel and output polarity in combination with coordinates and time information. Event-based cameras are characterized by high dynamic range, low latency, low power consumption, and high temporal resolution.

However, even if a high-performance image sensor is employed, the position and orientation of the object will be different from when the image sensor starts measuring the object, owing to the data acquisition time and calculation time of the sensor. In this study, a particle filter was used as a prediction function to provide the target command value considering the sensing time delays. A particle filter is a type of time-series filter, which is a method of representing the probability distribution of a target state by a set of multiple particles. Particle filters can estimate the state of a target by repeating cycles of state update, likelihood estimation, and resampling. As the input to the particle filter in this study is a set of event data, an event-driven particle filter for this input is proposed and applied to object tracking.

Fig. 1 shows a flowchart of the event-driven particle filter. The process shown in the flowchart in Fig. 1 was performed once for each input event. First, as in a typical particle filter, the state p_{start}, that is, the position and orientation of each particle, is set to a random value at time t_0 as an initiation step. The above initiation step was performed only once at the start-up.

In this filter, given an event input, state pi of each particle determines its next state according to the state transition model M. Referring to Ref. 3, we devised a method for stable tracking against sudden changes in direction by simultaneously using two models with different properties for position estimation—a Gaussian window model and a second-order autoregressive model—to cope with the tracking of complex motions. To assign each particle to the state transition model, a second-order autoregressive model was applied to estimate the position of particles with high likelihood and good tracking and a Gaussian window model for particles with low likelihood. Only the Gaussian window model was used for the orientation estimation.

Next, the likelihood of each particle is evaluated. To determine the likelihood of a particle's position, an exponentially decaying function is defined for the distance from the pixel position where the particle is located to the pixel position where the event occurred, which is added to the current particle's weight. In other words, if an event occurs near the position of a particle, the state of the particle is judged to be likely. To evaluate the likelihood of attitude, a group of Gabor filters with a strong response to five levels of orientation from 0 ° to 90 ° was applied to the input image. The likelihood of a particle is evaluated based on the orientation image with pixel response in the highest filter image, and the particle weights are accordingly calculated.

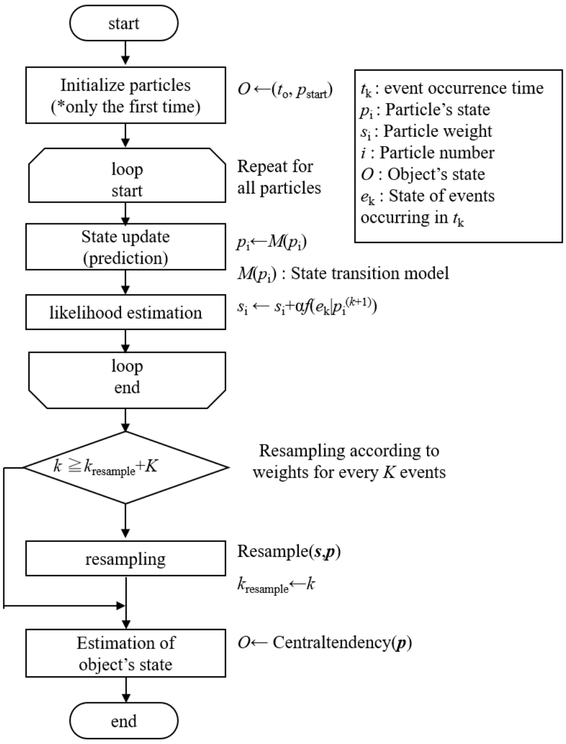

start
Initialize particles (*only the first time)
$O \leftarrow (t_o, p_{\mathrm{start}})$
t_k : event occurrence time
p_i : Particle's state
s_i : Particle weight
i : Particle number
O : Object's state
e_k : State of events occurring in t_k
loop start
Repeat for all particles
State update (prediction)
$p_i \leftarrow M(p_i)$
$M(p_i)$: State transition model
likelihood estimation
$s_i \leftarrow s_i + \alpha f(e_k | p_i^{(k+1)})$
loop end
$k \geqq k_{\mathrm{resample}} + K$
Resampling according to weights for every K events
resampling
Resample($\boldsymbol{s},\boldsymbol{p}$)
$k_{\mathrm{resample}} \leftarrow k$
Estimation of object's state
$O \leftarrow$ Centraltendency($\boldsymbol{p}$)
end

Fig. 1 Flowchart of the event-driven particle filter.

The above state and likelihood estimates were computed for all particles upon receipt of an event input. The state O of the tracked objects is the median of the state p_{start} of all particles.

Finally, at the stage where K event inputs are received, a particle resampling process is performed. In the resampling process, systematic sampling is performed depending on the weight of each particle, with particles with higher weights generating more replicas, and particles with lower weights being deleted. By repeating this operation, event-driven particle filtering can be achieved.

3. Experiments and Results

The proposed algorithm was evaluated in the experimental environment shown in Fig. 2. As the tracking target, a video of a rectangular object with a combination of constant angular velocity and periodic motion along a circular orbit was shown on a 360-Hz display, which was presented to an event-based camera, SilkyEvCam (CenturyArks Co., Ltd). The image resolution of the camera was VGA (640 × 480 pixels). A video was created at 120 °/s as a constant velocity motion, 4-Hz period as a periodic motion, and 30 ° as amplitude. A particle filter was applied to the results of the noise filtering performed on the event information acquired by the camera.

Figs. 3 and 4 show the output results of applying the proposed method every 1 ms in an offline setting. Fig. 3 shows the state of each particle (blue circle in the figure) and the estimated position of the object (red circle in the figure) for the image obtained by integrating the events between 200 and 1 ms at the start of the measurements. The direction of the arrow at the center of each particle and object represents the orientation, and the size of the circle indicates the magnitude of the weight. Fig. 4 shows the transition of the position of the object in the image coordinate system. The distance from the origin (radial direction) is constant, and the circumferential direction shows a tendency of back-and-forth motion characteristics, which is a combination of constant-velocity angular motion and periodic motion. Fig. 3 shows that the particles are distributed such that they cover the object, and the object is well tracked in terms of its probability distribution.

We also confirmed that these operations can be performed within a computation period of 10 ms in the

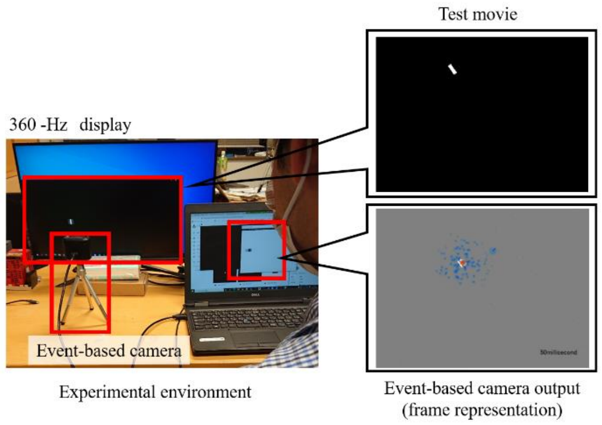


Fig. 2 Experimental environment.

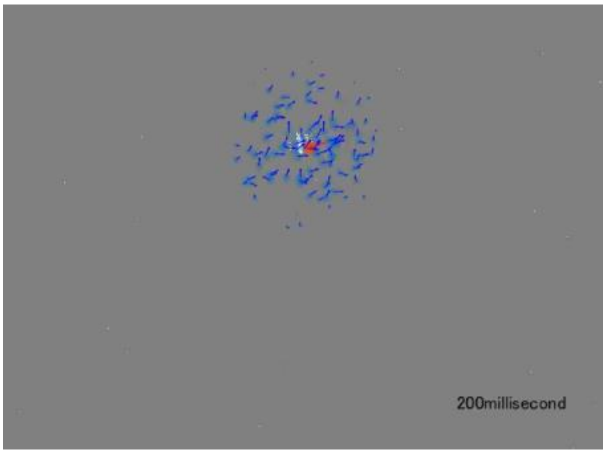


Fig. 3 State of each particle and estimated position and pose of the target at 200 ms.

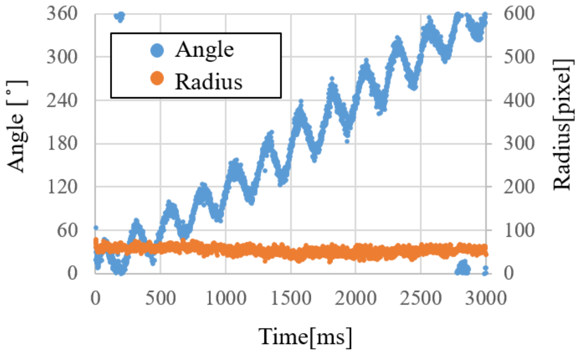


Fig. 4 Changes in target position and orientation predicted by the proposed method.

online mode. As a breakdown of the processing time, noise filtering takes 50 µs, and azimuth selectivity filtering takes 5 ms and the event-driven particle filtering required 100 µs/event for processing a single event, indicating that the azimuth selectivity filter is a bottleneck of the operation. It is apparent that there is an arithmetic bottleneck in the azimuth selectivity filter.

4. Conclusion

In this paper, we proposed an object tracking method using an event-driven particle filter. By presenting a test movie to a camera with a rectangular object to be tracked that combines constant velocity and periodic motion, we demonstrated that the proposed method can track an object offline with a 1-ms period.

The next step is to improve the processing efficiency by making the orientation selectivity filtering process event-driven and to improve the accuracy of the position and orientation estimation.

Acknowledgements

This work was supported by JSPS KAKENHI (grant number: 20K19895).

References

1. C. Posch, D. Matolin, and R. Wohlgenannt, “A QVGA 143 dB dynamic range frame-free PWM image sensor with lossless pixel-level video compression and time-domain CDS”, IEEE Journal of Solid-State Circuits, Vol. 46, No. 1, pp. 259-275, 2010.
2. G. Gallego et al., “Event-based vision: A survey”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 44, No. 1, pp. 154-180, 2020.
3. Y. Shiina, & T. Ikenaga, “Dual model particle filter for irregular moving object tracking in sports scenes”, In: APSIPA Annual Summit and Conference (ASC 2011), 2011.

Authors Introduction

Mr. Yuki Kawasaki

He received his B.S. degree in mechanical information science and technology from Kyushu Institute of Technology, Japan, in 2021. He is currently a Master’s course student in Kyushu Institute of Technology, Japan.

Dr. Masahiro Ohtani

He received his Ph.D. degree in electrical and electronic engineering from Toyohashi University of Technology. He is currently an Associate Professor in the electrical engineering, National Institute of Technology (KOSEN), Nara College, Japan. His research interests include neuromorphic engineering and analog VLSI.

Dr. Shinsuke Yasukawa

He received his Ph.D. degree in electrical, electronic, and information engineering from Osaka University, Japan, in January 2017. Currently, he is an Associate Professor in the Department of Human Intelligence Systems, Kyushu Institute of Technology, Japan. His research interests include information processing in biological systems and their applications in robotics.

A simulation model for analyzing the spatiotemporal receptive field of retinal ganglion cells in the presence of fixational eye movements

Hiroyuki Yokota
Department of Human Intelligence Systems, Kyushu Institute of Technology,
2-4, Hibikino, Wakamatsu, Fukuoka, 808-0196, Japan

Yuki Hayashida
Department of Information Engineering, Mie University,
1577 Kurima-machiya, Tsu, Mie 514-8507, Japan

Shinsuke Yasukawa
Department of Human Intelligence Systems, Kyushu Institute of Technology,
2-4, Hibikino, Wakamatsu, Fukuoka, 808-0196, Japan
E-mail: s-yasukawa@brain.kyutech.ac.jp
https://www.kyutech.ac.jp/

Abstract

Understanding the receptive field dynamics of vertebrate retinal ganglion cells in the presence of fixational eye movements is considered to provide insights into information processing/encoding optimized behaviors in animals and autonomous robots in the future. Previous studies have proposed computational models that account for fixational eye movements, suggesting that the long-range spatial inhibitions provided by wide field amacrine cells to some subtypes of bipolar cells and ganglion cells dynamically shape the responses of ganglion cells. In the present study, we constructed a simulation model of a retinal circuit based on recent physiological findings and computational elements that are feasible for digital hardware implementation. The model validity was tested through computer simulation experiments and by analyzing the time series of spike outputs from the ganglion cell unit. As a preliminary result, we were able to quantify the spatiotemporal receptive field by applying a simple white-noise movie stimulus.

Keywords: Simulation, retina, spike, fixed eye movement, receptive field

1. Introduction

The vertebrate retinal system undertakes energy-saving and robust encoding of visual information in diverse and unpredictable visual environments, and under non-stationary body-eye movement conditions prior to cognitive processing in the brain. Therefore, understanding the principles and mechanisms of this information processing provides useful insights, not only into the visual system of vertebrates but also, in the future, into visual information processing optimized for autonomous robots.

Previous physiological and anatomical experiments have shown the basic response properties of major intraretinal neurons and the synaptic connections between them [1]. To elucidate how real-world visual information is processed, compressed, and encoded in neural circuits, it is necessary to determine the electrical activity of many neurons under natural conditions, where animal body and eye movements are constantly present.

However, it is not easy to conduct this using only ordinary physiological experiments due to technical limitations.

The “in silico” approach simulates the response of neuronal populations on a general-purpose computer that is complementary to physiological experiments. In recent years, hardware tools have been developed to emulate neural images of the neuronal layer in the retina in real time [2]. Robotic systems with gaze control characteristics similar to vertebrate eye movements have also been developed [3], which can simulate the spatiotemporal frequency modulation of the input light to the retina by fixation eye movement.

Recent retinal studies have proposed computational models that consider neuronal responses in the presence of fixational eye movements, suggesting that wide field amacrine cells provide relatively long-range inhibitory inputs to several subtypes of bipolar and ganglion cells, dynamically shaping the receptive field properties of these cells [4][5]. Based on these recent physiological findings, this study proposes a retinal neural circuit model for future hardware implementations. The basic circuit structure was implemented using software, and the receptive field dynamics of the ganglion cell units were quantified by simulation experiments and Spike-Triggered Average (STA) analysis [6] using this model.

2. Proposed Simulation Model

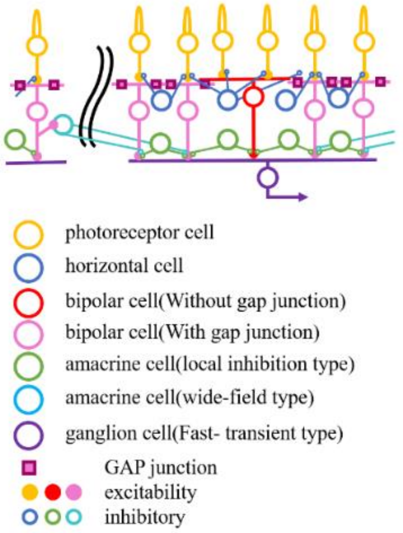

photoreceptor cell
horizontal cell
bipolar cell(Without gap junction)
bipolar cell(With gap junction)
amacrine cell(local inhibition type)
amacrine cell(wide-field type)
ganglion cell(Fast- transient type)
GAP junction
excitability
inhibitory

Fig. 1 Simulation model of retinal neural circuits in the presence of fixating eye movements.

Fig. 1 shows a simulation model of retinal neural circuits taking into account fixational eye movements, proposed with reference to previous studies [4][5]. This model includes photoreceptor cells (PCs), horizontal cells (HCs), bipolar cells (BCs), amacrine cells (ACs), ganglion cells (GCs) and the neural connections between these cells.

First, PCs detect light converting it into an electrical signal. The electrical signals were electrically coupled by gap junctions between the PCs, which propagate the signals faster in the horizontal direction. Electrical signals from PCs also reach the HCs with a delay. Gap junctions also exist between HCs, which propagate signals more widely than PCs. BCs form a center-surround antagonistic receptive field by subtracting the excitatory input from the PCs and inhibitory input from HCs.

A unique feature of this model is the neural circuitry of the inner retina after BCs. The BCs in the model are of the on-transient type, electrically coupled between neighboring bipolar cells (ECBCs) and electrically uncoupled (conventional bipolar cells; CBCs). There are also two types of ACs: local inhibition and wide field.

The output cells in the retinal model were fast-transient-type ganglion cells (FTGCs). From the location of the GCs, CBSs are located near the center, whereas EC BCs are located in the peripheral region. Each BC receives inhibitory signals from local inhibitory ACs. Inhibitory signals from wide field ACs have a larger amplitude and faster response than the local inhibition type and are excited by inputs from ECBCs. ECBCs receive inhibitory signals from distant, wide field ACs. In the presence of eye movements, excitatory signals from ECBCs increase the size of the receptive field of FTGCs, while inhibitory signals from wide field ACs decrease the size of the receptive field of FTGCs. The FTGCs integrate excitatory inputs from these BCs and output spike signals.

The spatial signal propagation properties through gap junctions present in photoreceptor, horizontal, and BCs, which are simulated by spatially convolving the Gaussian kernel in Equation (1).

$$g(x, y, \sigma) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(\frac{-x^2 + y^2}{2\sigma^2}\right) \qquad (1)$$

where σ is the scale parameter that determines the degree of spatial signal transmission spreading, and the scale parameters for PC, HC, and ECBC are σ_{pc}, σ_{hc} and σ_{ecbc} respectively.

The delayed response of HCs to PCs, and the temporal characteristics of bipolar and ACs were simulated by a linear sum of first-order Infinite impulse response (IIR) filters, as expressed in the following equation:

$$J_m[x, \mathrm{y}, \mathrm{n}] = \alpha_m J_m[x, y, n-1] + (1-\alpha_m) I_{in}[x, y, n] \tag{2}$$

$$I_{out} = \sum_{i=1}^{3} k_i J_i \tag{3}$$

where m= {1,2,3}, $\alpha 1 = \alpha hc$ for HCs, k1=1, k2= k3=0, $(\alpha_1, \alpha_2, \alpha_3) = (\alpha_{bc1}, \alpha_{bc2}, \alpha_{bc3})$, $k_1=-1$, $k_2=2$, $k_3=-1$ for BCs and $(\alpha_1, \alpha_2, \alpha_3) = (\alpha_{bc1}, \alpha_{bc2}, \alpha_{bc3})$, $k_1= 2$, $k_2=-1$, $k_3=0$.

The Izhikevich model [7] was employed to simulate the spike generation process using FTGC. The parameters of the model used in the experiments are summarized in Table 1. Note that this model was proposed for future implementations in operational circuits.

Table 1. Parameter list.

	Parameters	Values
PC	σ_{pc}	1
HC	σ_{hc}	3
	α_h	0.953
BC	σ_{ecbc}	1
	$(\alpha_{bc1}, \alpha_{bc2}, \alpha_{bc3})$	(0.875, 0.75, 0.625)
AC	$(\alpha_{ac1}, \alpha_{ac})$	(0.8, 0.6)
GC	(A, B, C, D)	(0.02, 0.2, -60, 8)
	v_{peek}	30

3. Experiments and results

STA was performed to quantify the receptive field dynamics of the FTGC units in the proposed model. The STA of a visual stimulus is widely employed in physiological experiments to localize the receptive field of the retinal ganglion cell[7].

The visual input to the model was a white noise movie stimulus, an example is shown in Fig. 2(a). The number of cells in each layer from the PC layer to the AC cell layer was set to 100 × 100, and the number of cells corresponded one-to-one to the number of pixels. Visual stimulus updates and various filter operations were performed every 5 ms, and numerical calculations of the Izhikevich model were performed with a resolution of 0.5 ms, considering the maximum frequency of spike output by actual GCs, and the speed and accuracy of the calculations. The simulation period was 500 s. The simulation code was implemented using Python programming language.

The spike response of the centrally located ganglion cell unit was recorded, resulting in a spike count of 11532 spikes during the entire simulation period. An STA was obtained by extracting the stimulus image stream in a time window from -150 ms before to 50 ms after the time of each spike and averaging the extracted image stream for all spikes, as shown in Fig. 2. In this spatiotemporal STA, yellow indicates a positive response, and dark blue indicates a negative response.

Fig. 2(c) shows the time profile on the horizontal dotted line in Fig. 2(b), showing a biphasic feature. At the central position, the STA intensity increased from 0 s and reached a positive peak at -10 ms before decaying and reaching a negative peak at -40 ms. The contrast in the time profile decreased further away from the central position. Fig. 2(d) shows the spatial profile of the vertical dotted line in Fig. 2(b). From the spatial profile at -10 ms, which is the positive peak, the ON region was approximately 10 pixels (see ‘4’ in Fig. 2(d)). From the negative peak spatial profile was -40 ms, its OFF region was approximately 10 pixels (see ‘2’ in Fig. 2(d)).

4. Conclustion

The receptive fields of the FTGC units were measured by performing STA analysis on the proposed model. The temporal and spatial structures of the receptive field, as observed physiologically, were visible in the receptive field of the FTGC unit.

A simulation time of approximately 22 h is required to obtain 11532 spikes. Therefore, it is necessary to implement the model in dedicated hardware to reduce the simulation time and the time required for data analysis. In addition, we plan to adjust the parameters of the spatial and temporal scales to obtain physiologically valid spatiotemporal characteristics of the receptive field.

Acknowledgment

This work was supported by JSPS KAKENHI (grant number: 20K19895).

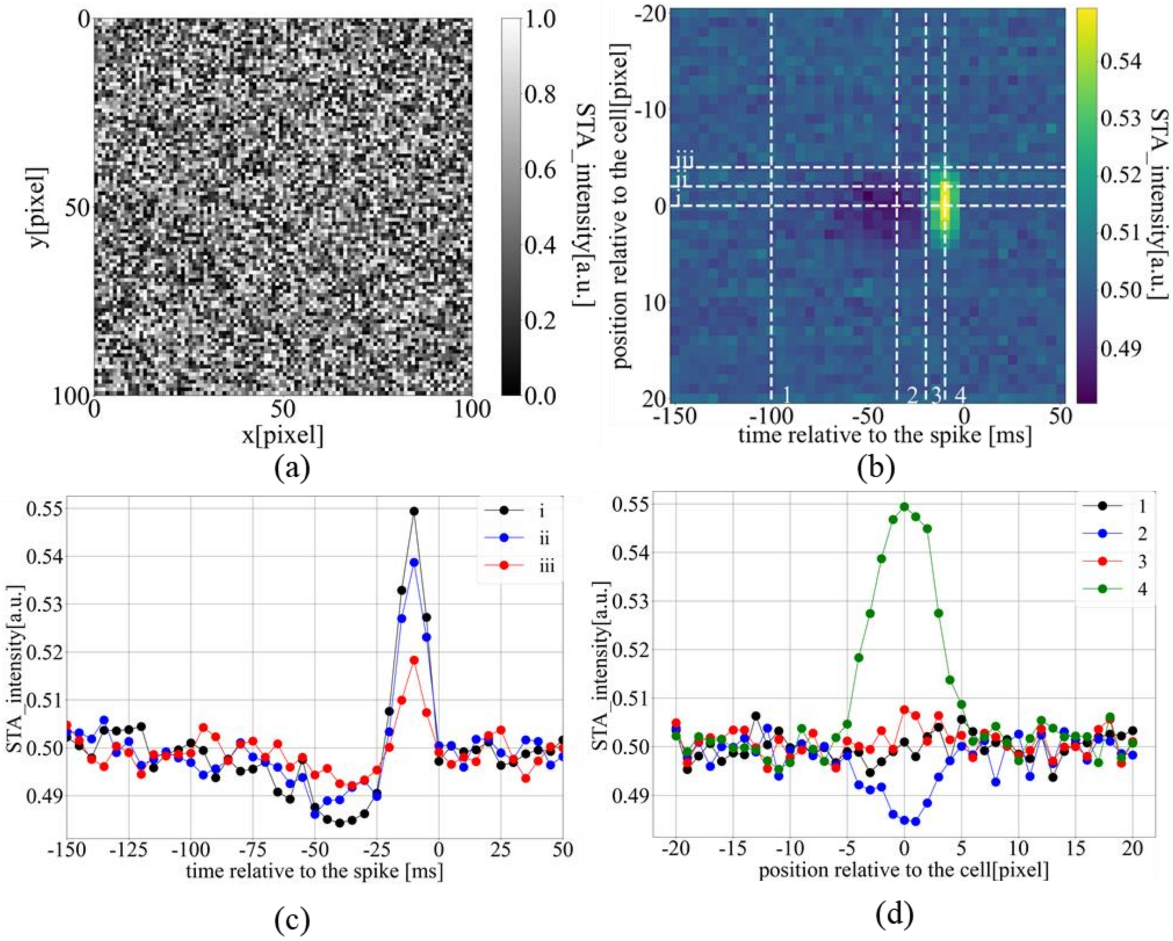


Fig. 2 Simulation experiment to characterize the spatiotemporal receptive field of the ganglion cell using STA analysis. (a) Examples of visual input. (b)Space-time STA plot of the fast-transient type ganglion cell (c) Temporal profile of the STA sliced at three positions labeled “i”, “ii”, “iii” in (b), (d) Spatial profile of the STA sliced at four time points labeled “1”, “2”, “3”, “4”, in (b).

References

1. J. E. Dowling, “The Retina: An Approachable Part of the Brain”, Harvard University Press, 2012.
2. Y. Hayashida, Y. Kudo, R. Ishida, H. Okuno, & T. Yagi, “Retinal Circuit Emulator With Spatiotemporal Spike Outputs at Millisecond Resolution in Response to Visual Events”, IEEE transactions on biomedical circuits and systems, Vol. 11, No. 3, pp. 597-611, 2017.
3. T. Yotsumoto, Y. Hayashida, & S. Yasukawa, “A Robotic vision system emulating fixational eye movements and retinal sampling”, Proceedings of 2022 International Conference on Artificial Life and Robotics, Vol. 27, pp. 519-523, 2022.
4. A. Matsumoto, & M. Tachibana, “Global jitter motion of the retinal image dynamically alters the receptive field properties of retinal ganglion cells”, Frontiers in neuroscience, Vol. 13, Article 979, 2019.
5. S. A Baccus, B. P. Ölveczky, M. Manu, & M. Meister, “A retinal circuit that computes object motion”, Journal of Neuroscience, Vol. 28, No. 27, pp. 6807-6817, 2008.
6. F. Rieke, D. Warland, R. D. R. Van Stevenine, & W. Bialek, “Spikes: exploring the neural code, MIT press, 1999.
7. E. M. Izhikevich, “Dynamical systems in neuroscience”, MIT press, 2007.

Authors Introduction

Mr. Hiroyuki Yokota

He received his B.S. degree in intelligent mechanical engineering from Hiroshima Institute of Technology, Japan, in 2022. He is currently a Master's course student in Kyushu Institute of Technology, Japan.

Dr. Yuki Hayashida

He received his Ph.D. in information engineering from Kyushu Institute of Technology, Kitakyushu, Japan, in 1998. He is currently a Professor with the Graduate School of Engineering, Mie University, Japan. His research interests include physiological and engineering studies on biological nervous and cardiac systems, neural prostheses, and nanomaterial-based neural electrodes.

Dr. Shinsuke Yasukawa

He received his Ph.D. degree in electrical, electronic, and information engineering from Osaka University, Japan, in January 2017. Currently, he is an Associate Professor in the Department of Human Intelligence Systems, Kyushu Institute of Technology, Japan. His research interests include information processing in biological systems and their applications in robotics.

Junior High School Rescue Robot Challenge for Fostering Problem-Solving Skills

Kazuo Kawada, Keita Murai, Yuta Susawa, Hiroyuki Y. Suzuki
Graduate School of Humanities and Social Sciences, Hiroshima University, 7398524 Hiroshima, Japan
E-mail: kawada@hiroshima-u.ac.jp, m212958@hiroshima-u.ac.jp,
m220547@hiroshima-u.ac.jp, hiro-suzuki@hiroshima-u.ac.jp
www.hiroshima-u.ac.jp

Abstract

“Junior High School Student Rescue Robot Challenge” is an annual activity organized by Hiroshima University in cooperation with a construction equipment manufacturer. The Challenge has been continued for 18 years with a given theme of rescuing. The policy of the program was essentially changed, nonetheless, from 2022, in which the both problem-finding and problem-solving skills were emphasized. The theme of “the Challenge in 2022” is totally operated by students, which includes from proposal of the problem to be solved in disaster-stricken areas, to finding resolutions by making prototype rescue robot by modifying remote-controlled excavator model of 1/14 scale.

Keywords: Rescue Robot, Junior High School, Technology Education, Robot Evaluation, Modified Excavator

1. Introduction

Sustainable Development Goals (SDGs) were proclaimed at the UN Summit, in 2015, which will be attained by 2030 according to the statement of **“**The 2030 Agenda for Sustainable Development**”**. Around the same time, “Society 5.0” was proposed in “Fifth Science and Technology Basic Plan” [1] by Japanese Cabinet Office, as an image of upcoming society structure. It will be expected that the Society 5.0 will play a central role to achieve the SDGs in Japan. Recently, the Sixth Plan was also publicized [2] with advanced guidelines for Society 5.0 to realize SDGs more realistically. The plan focused on innovation creation by fostering **“**human resources who can create new value**”.** Besides that, the whole society around the world had affected remarkably, from 2010s, by global environmental changes, political and economic uncertainty, COVID-19 epidemic, and so on. There is a growing need for innovation creation to solve a wide variety of social issues, which is called “social change-type innovation” [3].

According to above mentioned situations, the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) also revised “Courses of Study” from 2008 [4][5][6], focusing more on social issues in school study and aiming to create new values, putting the both "problem finding and solving skills" in the central position for school education.

In such context, our Robot Challenge had been changed as well. “Junior High School Student Rescue Robot Challenge” had started as “the Contest” since 2005, by staffs in teacher training course of technology and information, Hiroshima University. The policy had been changed to “the Challenge” since 2015, and the name as well, with a support by a counterpart company of construction equipment manufacturer. We introduced, from that time, the both problem-finding and problem-solving approaches, but the theme itself was given one from the program side.

Ultimate revision of the program took place in 2022, in succession to tragic cancellation of the Challenge in 2021, because of COVID-19 pandemic. We noticed that the variation of the disasters was going to be wide. Therefore, we had changed from the basic framework of the program. Now, the problem-finding is also involved for students’ task. After finding the problem to be solved, the students will build a prototype rescue robot to resolve the problem.

We also revised the evaluation system as we latterly mentioned.

2. Challenge Theme in 2022

2.1. *Challenge Theme*

Since “the Contest” era, the central theme of the Challenge was “Rescuing Quickly a dummy doll from a large 1/8-scale simulated earthquake site to a safe zone.” [7][8] Large-scale field (1 to 3 meters ling) and the dummy dolls were prepared by the program staff. The students made rescue robots on which the dummy doll was placed, and so on. The evaluation had focused on the time required for rescue, ideas for rescue methods, humanitarian viewpoints, and especially the "gentleness" of the robot to the dummy during the rescue.

The theme of the challenge in 2022 has been significantly changed expecting further growth of junior high school students' ability to find and solve social problems. The project was started from problem-finding by themselves. After finding a problem, they are going to build a prototype rescue robot by modifying a 1/14-scale remote-controlled excavator shown in Fig.1. We did not restrict the problem to be chosen, but had a restriction of the selection of problems, it may be a disaster-stricken area, because of the excavator model.

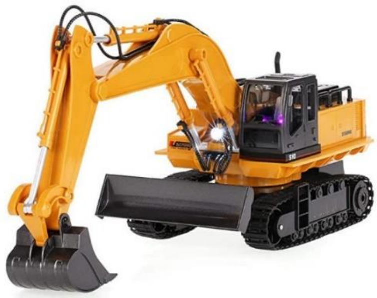


Fig.1. RC excavator to be modified

2.2. *Rescue Robot Idea Evaluation Method*

The evaluation of rescue robot ideas was consisted of two major evaluating sessions.

The first was evaluation of ideas, including from problem-finding to problem-solving, that is, what kind of situation were chosen and how rescue activities could be performed. Each team made a worksheet of conceptual plan with drawings and submit it to us. Before the worksheet making and submission, we instructed them how to incorporate measurement and control techniques. The worksheets were evaluated from following three viewpoints:

(1) Innovation (10 points)
Is it an innovative idea that utilizes shovels?
(2) Feasibility (10 points)
Is it a highly feasible concept or idea?
(3) Functionality (10 points)
Whether the concept/idea is expected to be sufficiently effective for the rescue activities set up.

Based on the above evaluation, all 10 teams were past the first selection.

For the second evaluation, each team made the rescue robot and put it in disastrous area to rescue people or something they assumed. Rescue activity was taken by video and the submitted to us. Documentation on the robot was submitted as well. Those materials were evaluated through an online conferencing system (Teams) in a public session. Three judges were selected to evaluate and announce the results. The following two items were evaluated for the robot:

(1) Feasibility/Improvement (30 points)
How much was the robot fulfilled the planned ideas shown in the first worksheets, and how much was it improved and realized, through real robot making process.
(2) Design (30 points)
Whether the robot's features and performance can be understood by users (or three judges who see it for the first time).

The presentation was evaluated on the following two points:

(1) Objectivity (15 points)
Whether the performances of the rescue robot can be understood and analyzed objectively and accurately by three judges.
(2) Expression (15 points)
Whether the features of the rescue robot was expressed, throughout the presentation, in an easy-to-understand manner.

In addition, the results of the first evaluation were also taken into account.

3. Robot Evaluation Results

The 7th Annual Middle School Rescue Robot Challenge in 2022 (in Japanese fiscal year of 2021) was held on February 12, 2022. A total of 10 teams participated. The results of the first and second rounds of judging are shown in Table 1 and 2, respectively.

The appearances of the robots they built are shown in Fig.2. Their robots had buckets, cabins, rollers, crawlers, traveling parts, and so on, by modifying excavator parts and utilizing its functions with clever modifications.

Table 1. First Evaluation Results

Team	A	B	C	D	E	F	G	H	I	J
Innovation	4.6	5.0	6.8	5.5	8.0	5.8	8.1	7.4	8.4	7.3
Realization	8.1	8.1	7.2	7.6	5.6	7.3	4.9	5.1	5.3	6.4
Functionality	5.8	7.3	6.3	6.5	7.3	5.9	7.4	7.6	7.1	7.1
Total Point	18.5	20.4	20.3	19.6	20.9	18.9	20.4	20.1	20.8	20.8

Table 2. First Evaluation Results

Team	A	B	C	D	E	F	G	H	I	J
Realization / Improvement	21.0	22.0	19.0	17.0	25.0	18.0	22.0	24.0	25.0	23.0
Design	17.0	21.0	24.0	19.0	22.0	20.0	23.0	21.0	23.0	24.0
Robot Point	38.0	43.0	43.0	36.0	47.0	38.0	45.0	45.0	48.0	47.0
Objectivity	10.5	9.5	11.0	9.0	12.0	10.5	10.5	9.0	11.5	12.0
Expression	13.0	11.0	11.5	8.0	11.0	10.0	11.5	11.0	13.0	11.0
Presentation Point	23.5	20.5	22.5	17.0	23.0	20.5	22.0	20.0	24.5	23.0
Total Point	61.5	63.5	65.5	53.0	70.0	58.5	67.0	65.0	72.5	70.0

4. Discussion

This challenge asked middle high school students to propose a rescue robot by modifying an excavator. We hope, by the Challenge, the students will develop their problem finding as well as problem solving skills. Judgements given from the first and second rounds of evaluation are shown in Table 1 and Table 2.

Table 1 shows that the total score for innovation, realization, and functionality. Many team got around 20 points (on a 30-point scale).

It is noteworthy point that, among three evaluation points there can be found somehow trade-off relationship. For example, teams A, B, D, and F have a low innovation rating although a high realization rating. This may show that these teams took more realistic solutions for the problems. In contrast, teams E, G, and I are rated high for innovation but low for realization because they were tried to add further functions that looked hard to be realized than others. These results indicate that these three evaluations were able to accurately assess the balance among concepts, ideas and reality of the robots

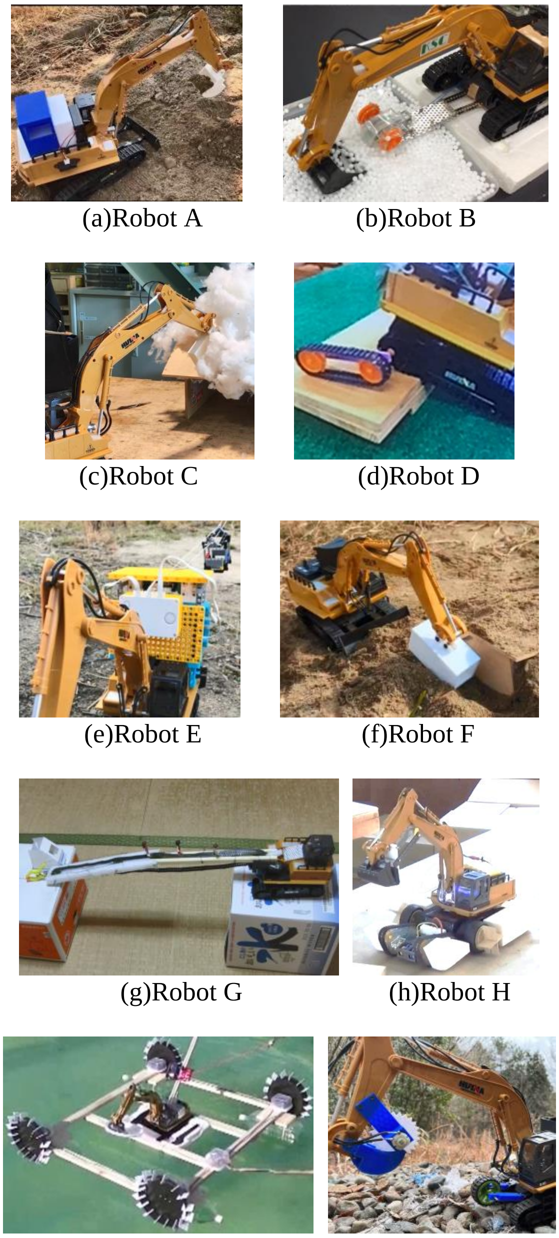

(a)Robot A (b)Robot B
(c)Robot C (d)Robot D
(e)Robot E (f)Robot F
(g)Robot G (h)Robot H
(i)Robot I (j)Robot J

Fig.2. Robots made by junior high school students

they made.

Table 2 shows the evaluation of realization/improvement points of the robot they made. Unfortunately, the teams with poor motion video were rated low evaluations. In design evaluation, the robots which were scarcely modified from those of commercially available power shovels got low scores.
As for the evaluation of presentation, the objectivity was rated by the use of documents and quantitative data. The teams which provided realistic evidence were highly evaluated. In the expression, which is the evaluation of comprehensibility of the presentation, the teams that proposed a clear rescue sequence got high scores. These results indicate that these evaluations were able to accurately assess the robots and presentations made by the junior high school students.

Appearances of robots built by a junior high school student shown in Fig.2 clearly indicate that the students manage to born a brand new rescue robots from their brains. No two robots had the same appearances nor the same function. The robots built by the 10 teams can be classified into the following three categories.

(1) Modification of the bucket part (Fig.2 (a), Fig.2 (b), Fig.2 (c), Fig.2 (f) and Fig.2 (j))
The bucket was modified to make it easier to scoop debris or to have other functions.

(2) Modification of the traveling part (Fig.2 (d), Fig.2 (h) and Fig.2 (i))
New parts were added to make it easier to travel over rubble, and floats were added to enable movement over water.

(3) Adding new functions (Fig.2 (e) and Fig.2 (g))
The excavator had added functions completely different from those of shovels, such as shoveling and leveling.

In total, we can conclude that the junior high school students who participated “the Challenge in 2022” were able to set and solve their own problems, indicating that our revision of theme setting for “the Challenge” was effective.

5. Conclusion

We introduced a new perspective in technology education which is focused on the both "problem finding and solving skills". According to that, theme of “Junior High School Student Rescue Robot Challenge” was changed, which started from problem-finding by students. Then, they were going to make real robots by modifying remote-controlled excavator. Ten teams of junior high school students were participated the Challenge and found unique problems to be solved and made real robots which fulfilled their images. We believe that the Challenge had worked for fostering innovative mind for young generations.

References

1. Japan Cabinet Office: The 5th Science and Technology Basic Plan (2016) (in Japanese)
2. Japan Cabinet Office: The Sixth Science, Technology and Innovation Basic Plan (2021) (in Japanese)
3. Center for Research and Development Strategy, Japan Science and Technology Agency: Trends and challenges of mission-oriented science, technology and innovation policies for solving social problems (2021) (in Japanese)
4. Ministry of Education, Culture, Sports, Science and Technology: Elementary School Curriculum Guideline 2017-03 Notification (2017)
5. Ministry of Education, Culture, Sports, Science and Technology: Lower Secondary School Curriculum Guideline 2017-03 Notification (2017)
6. Ministry of Education, Culture, Sports, Science and Technology: Upper Secondary School Curriculum Guideline 2017-03 Notification (2018)
7. Kazuo Kawada, Masayasu Nagamatsu, Toru Yamamoto: A Practice of Rescue Robot Contest in Junior High Schools, Journal of Japanese Society for Engineering Education, Vol.58-2, pp.33-39 (2010) (in Japanese)
8. Kazuo Kawada, Keisuke Iuchi, Keita Murai, Hiroyuki Y. Suzuki: Junior High School Rescue Robot Challenge using Shock Sensitive Tiny Dummy Robot, Journal of Robotics, Networking and Artificial Life, Vol.8(2), pp. 90–93 (2021)

Authors Introduction

Dr. Kazuo Kawada

He received his B.Eng. degree from Hiroshima University 1995, and his Ph.D. from Hiroshima University in 2005. He is currently an Associate Professor in the Dept. of Technology and Information Education, Graduate School of Humanities and Social Sciences at Hiroshima University. His research interest areas are the development of educational materials related to mechatronics education, data science education and innovation human resource development for K-16.

Mr. Keita Murai

He received his B.Ed. degree in Education in 2021 from the Faculty of Education, Hiroshima University. He is acquiring the M.Ed. in the Graduate School of Humanities and Social Sciences, Hiroshima University. His research interest areas are the development of programming thinking and STEAM materials for K-12.

Mr. Yuta Susawa

He received his B.Ed. degree in Education in 2022 from the Faculty of Education, Hiroshima University. He is acquiring the M.Ed. in the Graduate School of Humanities and Social Sciences, Hiroshima University. His research interest areas are the development of educational materials on programming problem solving in measurement and control and data science problem solving for K-12.

Dr. Hiroyuki Y. Suzuki

He received his B.Eng. degree from Hiroshima University 1992, and his D.E. from Hiroshima University in 2000. He is currently an Associate Professor in the Dept. of Technology and Information Education, Graduate School of Humanities and Social Sciences at Hiroshima University. His research interest areas are education on materials science, engineering and processing and their application on educations, and STEM/STEAM educations.

A Study of Experiential Learning Activities using Model Materials for the Kicking Motion

Teruyuki Tamai
Faculty of Education, Ehime University
3, Bunkyo-cho, Matsuyama, Ehime 790-8577, Japan

Shoki Takeuchi
Faculty of Education, Ehime University
3, Bunkyo-cho, Matsuyama, Ehime 790-8577, Japan

Yoshihiro Ohnishi
Faculty of Education, Ehime University
3, Bunkyo-cho, Matsuyama, Ehime 790-8577, Japan

Kazuo Kawada
Graduate School of Humanities and Social Sciences, Hiroshima University
1-1-1, Kagamiyama, Higashi-hiroshima, Hiroshima 739-8524, Japan
E-mail: tamai.teruyuki.xq@ehime-u.ac.jp, g230072c@mails.cc.ehime-u.ac.jp, ohnishi@ehime-u.ac.jp,
kawada@hiroshima-u.ac.jp
www. ehime-u.ac.jp, www. hiroshima-u.ac.jp

Abstract

In order to develop human resources for the realization of Society 5.0, it is necessary for students to acquire the ability to solve problems that transcend the frameworks of subjects. For this purpose, cross-curricular learning is attracting attention. In this study, we examine teaching materials that connect the content of measurement and control in technology education with studies in science, health and physical education. A simple mechanism model teaching material is proposed to allow students to learn about ball operation through experiential learning by conducting motion experiments on a model that simulates the kicking motion.

Keywords: a simple mechanism model, technology education, teaching material, cross-curricular learning

1. Introduction

In 2018, the Ministry of Education, Culture, Sports, Science and Technology (MEXT) released "Human Resource Development for Society 5.0~Changes to Society, Changes to Learning~ (Summary)". [1] In the summary, the MEXT indicated that the direction to be taken in the future is to transcend the humanities/sciences divide. Then, the Cabinet Office issued the "Science, Technology, and Innovation Basic Plan" for 2021.[2] In this plan, the realization of Society 5.0 is indicated. In order to develop human resources for this purpose, it is said that it is important to give them direct experience of the real thing through STEAM (Science, Technology, Engineering, Art(s), Mathematics) education. In these ways, it can be said that learning is required for students to acquire the ability to solve problems that transcend the frameworks of subjects.

In Japan, technology education is mainly conducted in junior high school technology and home economics (technology field). The Lower Secondary School National Curriculum Standard, which sets standards for organizing the curriculum, calls for cross-curricular learning. On the other hand, it also calls for the development of information literacy. In addition, many teaching materials and classes have been developed for learning related to measurement and control. [3], [4] Therefore, we decided to examine teaching materials that allow students to learn about measurement and control as part of their technology education and that are connected to the content of their studies in science, health, and physical education.

In this study, a simple mechanism model teaching material was proposed to enable experiential learning of ball operation by conducting motion experiments on a model that simulates the kicking motion.

2. Composition of Teaching material

The proposed teaching materials include the learning contents of technology, science, and health and physical education shown in Table.1. The proposed teaching materials are shown in Fig. 1. The mechanism model was created with the hip joint motion in the kicking motion as sarvomotor_A and the knee joint motion as sarvomotor_B. This motion has connections to the study of football. The hardware was fabricated using a servomotor (SG90, Made by Tower Pro Pte Ltd), a stainless-steel frame (Made by Yamazaki), and plywood as materials. The fixture for attaching the servomotor to the plywood was made by myself using a 3D printer. To operate the servomotor, Studuino was used as an educational microcontroller board.[5] Three AA batteries, 4.5 V in total, were used to power the servomotor.

Table 1. Learning content of each subject.

Subject	Learning contents
Technology	Mechanism, Model
Science	Force, Point of action
Health and physical education	Football, Ball operation

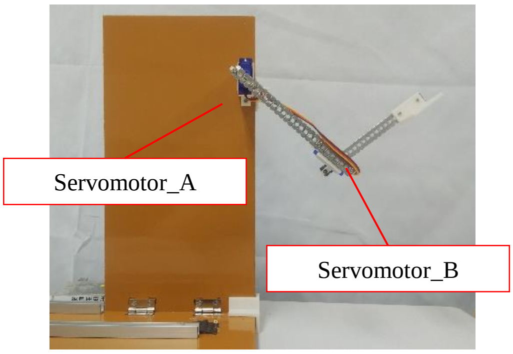


Fig. 1. Mechanism model teaching material

3. Motion of Teaching material

The proposed teaching material can be operated in two patterns. The first pattern (kick_A) is to operate the sarvomotor_A. The force to poke the ball in kick_A is obtained from the torque relation in Eq. (1) below. Where $T[N \cdot m]$, $F[N]$, $r[m]$.

$$T = F \cdot r \tag{1}$$

Stall torque listed in the data sheet of the sarvomotor used is $1.8kgf.cm$. The turning radius r_1 shown in Fig. 2 is $240mm$.

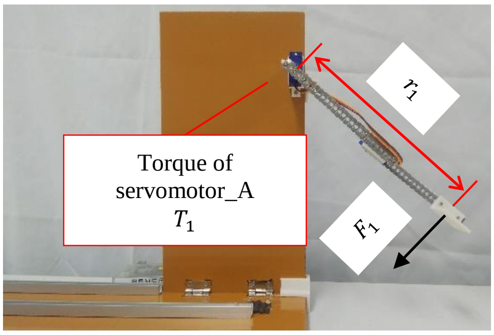


Fig. 2. Radius of gyration of servomotor_A in kick_A

From Eq. (1), the value of F_1 is Eq. (2).

$$F_1 = \frac{T_1}{r_1} \doteqdot 0.735[N] \tag{2}$$

The second pattern (kick_AB) operates sarvomotor_A, sarvomotor_B. sarvomotor_A can apply force F_2, shown in Fig. 3. However, due to the operation of sarvomotor_B, the turning radius r_2 is less than $240mm$. sarvomotor_B can apply a force of F_3, as shown in Fig. 4. However, the radius of rotation r_3 is $140mm$.

From Eq. (1), the value of F_2 is Eq. (3).

$$F_2 = \frac{T_2}{r_2} \geq 0.735[N] \tag{3}$$

From Eq. (1), the value of F_3 is Eq. (4).

$$F_3 = \frac{T_3}{r_3} \fallingdotseq 1.26[N] \tag{4}$$

From Eq. (2)-(4), the forces on kick_A and kick_AB are related in Eq. (5).

$$F_1 < F_2 + F_3 \tag{5}$$

As shown above, kick_AB can apply more force to poke the ball than kick_A. Thus, there is a connection with the learning of the force applied to the point of action.

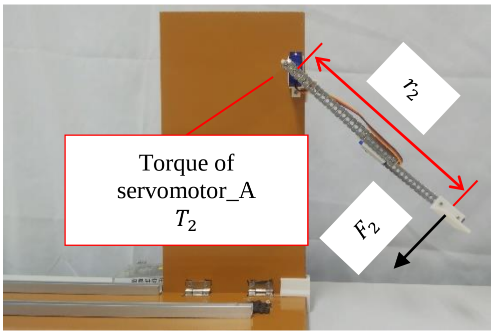


Fig. 3. Radius of gyration of servomotor_A in kick_AB

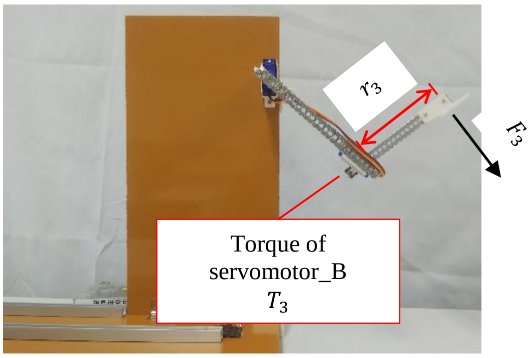


Fig. 4. Radius of gyration of servomotor B in kick_AB

4. Experimental method

Learners will experiment with the proposed teaching material in action in order to learn experientially the relationship in Eq. (5). The positional relationship between the proposed teaching material and a plastic ball ($\phi = 51.5mm$) is shown in Fig. 5. In kick_A and kick_AB, the angle θ_A of servomotor_A was set at 45° as shown in Fig. 6. As shown in Fig. 7, the angle θ_B of servomotor_B was set at 110°. In order to measure the velocity of the ball poked by kick_A and kick_AB, a video was recorded using an iPad. Each was operated ten times and the average velocities were compared.

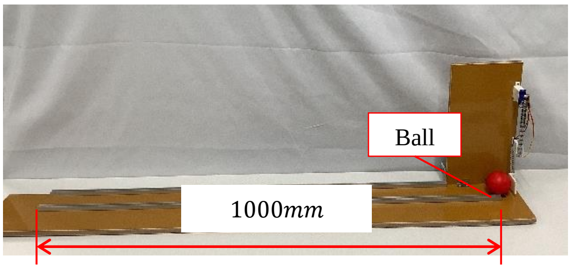


Fig. 5. How to measure ball speed.

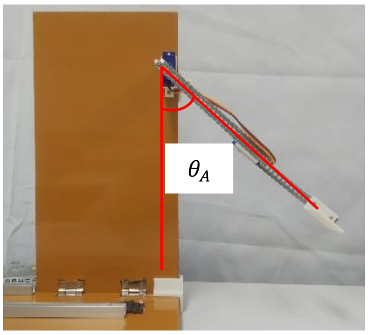


Fig. 6. Servomotor_A angle.

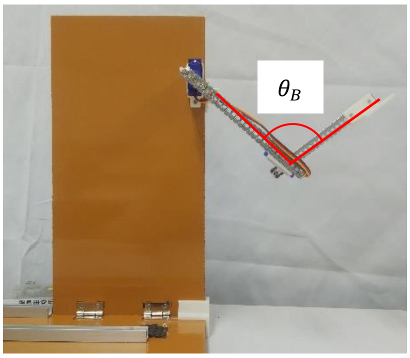


Fig. 7. Servomotor_B angle.

5. Experimental result

The measured velocity of the ball for kick_A and kick_AB are shown in Table.2. The average velocities were $1.09m/s$ for kick_A and $1.22m/s$ for kick_AB. The results showed that the ball velocity was higher when two sarvomotor_A and sarvomotor_B were used. Thus, similar results to Eq. (5) obtained from the torque relation equation are obtained from the motion experiments of the proposed teaching material. These results suggest the possibility of having students learn that the ball velocity can be increased by linking the hip joint motion with the knee joint motion when kicking a ball.

Table 2. Measurement result.

	kick_A		kick_AB	
	time[s]	velocity[m/s]	time[s]	velocity[m/s]
1th	0.90	1.11	0.82	1.22
2th	0.92	1.09	0.82	1.19
3th	0.92	1.09	0.82	1.22
4th	0.93	1.08	0.82	1.22
5th	0.92	1.09	0.84	1.19
6th	0.93	1.08	0.83	1.20
7th	0.90	1.11	0.81	1.23
8th	0.93	1.08	0.81	1.23
9th	0.93	1.08	0.80	1.25
10th	0.92	1.09	0.82	1.22
average	0.92	1.09	0.82	1.22

6. Conclusion

In this study, a simple mechanism model was proposed that can be made to learn about the kicking motion from the motion experiments of the model. Then, it was shown that the difference of kicking motion can be discussed from the relational equation and the motion experiment of the mechanism model. In the future, it will be considered to be practiced by learners as cross-curricular learning. This work was supported by JSPS KAKENHI Grant Number JP21K02951.

References

1. MEXT(Ministry of Education, Culture, Sports, Science and Technology), "Human Resource Development for Society 5.0〜Changes to Society, Changes to Learning〜(Summary)", 2018
2. Cabinet Office, " Science, Technology, and Innovation Basic Plan", 2021 https://www8.cao.go.jp/cstp/english/sti_basic_plan.pdf [Accessed December 10,2022]
3. Y. Ohnishi, K. Honda, R. Nishioka, S. Mori and K. Kawada, "Robotics Programming Learning for Elementary and Junior High School Student", Journal of Robotics and Mechatronics, 29(6) 2017, pp. 992-998
4. S. kabemune, S. Shirai and S. Tani, "Informatics and Programming Education at Primary and Secondary Schools in Japan", Olympiads in Informatics, Vol.11, pp.143-150, 2017
5. K. Omata and S.Imai, "Practice of Programming Education using Finger Robot", Journal of Robotics,Networking and Artificial Life, Vol.6, No.4, pp.262-264, March,2020

Authors Introduction

Dr. Teruyuki Tamai

He received his Ph.D. at Graduate School of Education in Hiroshima University. He is an Assistant Professor in the Faculty of Education, Ehime University.

Mr. Shoki Takeuchi

He is acquiring the B.Ed. in the Faculty of Education in Ehime University.

Dr. Yoshihiro Ohnishi

He received his Dr. Eng. from Osaka Prefecture University in 2002. He is a professor at Faculty of Education in Ehime University.

Dr. Kazuo Kawada

He received his Ph.D. degree at Graduate School of Education in Hiroshima University. He is an Associate Professor in the Graduate School of Humanities and Social Sciences, Hiroshima University.

Making High Precision Single Balance in Active Learning Seminar for Hiroshima Univ. Monozukuri Junior Doctor Special Educational Program

Hiroyuki Y. Suzuki, Kazuo Kawada, Masayasu Nagamatsu
Graduate School of Humanities and Social Sciences, Department of Educational Science, Hiroshima University,
1-1-1-Kagamiyama, Higashi-Hiroshima City, Hiroshima 739-8524, Japan
E-mail: hiro-suzuki@hiroshima-u.ac.jp, kawada@hiroshima-u.ac.jp, nagamatu@hiroshima-u.ac.jp

Abstract

New learning materials which can cultivate competency of young people are expected. STEM/STEAM educational materials in technology field are the candidate of them, so that we proposed Active Learning seminar of “Making High Precision Single Balance”. The seminar took place in “Hiroshima Univ. Monozukuri Junior Doctor” with around 40 participants of young (11 - 15 in age) students. We prepared a number of pre-fabricated parts of deferent sizes, giving them to young participants teams (about 5 people each) to assemble single balance, putting their own ideas in the design by selecting favorite parts by themselves. Majority of the students could understand key concept of the balance (rule of moment) and assemble single balance with accuracy of milligrams. We believe that the seminar acts as a competency nurturing content.

Keywords: Technology education, STEM, Quantitative evaluation, Accuracy

1. Introduction

A new paradigm is emerging in school education. It puts emphasis on “competency” [1], in contrast to “contents” based current educational curriculums. The word “competency” has many facets in its image, but one simple image related to technology field is “capacity for problem solving”. In such context, Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) starts to introduce “Tankyu (Inquiry, in Japanese)” subjects in senior high-school curriculums from 2022 [2], in which students try to solve problems found around their daily life, which contain, on the other hand, some level of complexity.

By the way, giving solutions to real problems are somehow essential aspect of technology. Every engineer around the world struggles to solve problems, by thinking, manipulating, re-combining his/her knowledges and skills in daily job. Actually, problem solving curriculums in technology or engineering subjects have longer history than other subjects, known as Project (or Problem) Based Learning (PBL) [3], STEM/STEAM educations [4], and so on. It is worthy to propose contents from technology field, since it is filled with problems must be solved. However, introduction of such contents to school education is still limited.

One of the major obstacles must be the level of problems that engineers tackle with. Many of them are too much specialized in one aspect, or too much complicated to untangle by young students. This is partly because of dramatic development of technology in 21th century. The technology nowadays becomes very sophisticated huge systems, but at the same time, it also becomes very difficult to grab the whole picture, especially for non-specialized people. We have to manage to extract fragmented problems which are matched to different ages of students to be solved, from such huge entangled situation.

In the present study, we introduced a new theme of "Making High Precision Single Balance" as an Active Learning (AL) material. We tried to deduct the complexity of problem, but maintaining the essential part of the technology in it. This material was given to several special seminars from elementary to high school students, and was positively accepted by the students as good problem (theme) to be solved.

2. Basic Policy on Problem Selection

We proposed several STEM/STEAM educational contents in resent studies [5][6] and found, through those experiences, what are the essentials for problem selection, which are,

1. the problem will be solved by making things using knowledge and skills in STEM fields,
2. but the STEM elements to be used must be reachable for respective students,
3. it is favorable to have room for free design in making process,
4. but evaluating point must be quantitative,
5. precise assembling, exact and smooth movement will result in high evaluation,
6. and it will be better that the problem will be solved by group activity.

In our perspective, the theme of "Making Balance" is one of potency material for STEM education, since the basics of the balance is quite simple, that lever rule is taught from elementary school and the main mechanism is only a pivot, but also there will be a number of variations in solutions, as we see many kinds of balance products in market.

We already proposed a STEM educational material of "Making Electric Balance" in previous study [5], in which ample STEM knowledge and skills are incorporated. The target of the "Making Electric Balance" is junior to senior high school students, and estimated duration of the project will be several months.

The seminar in the present study, on the other hand, was only a half day one, and the participant included elementary school students. Therefore, we have to trim the content to appropriate size, extracting essential part from previously developed material. Eventually, we changed the key of the problem from making "Electric" balance to making "Single" balance.

A STEM program of Making Single Balance was placed in a half day seminar in a weekend. The seminar was originally planned for Hiroshima University Monozukuri Junior Doctor, a special educational program for young (11-15 in age) volunteered students. The program was founded by Japan Science and Technology Agency (JST) with a set of their policy, of only few excellent students will be selected "individually" and nurtured to higher level than current school education. Nonetheless, we introduced distinctive characteristics of "group activity" to Hiroshima University's program, because of the reasons as we mentioned above. The idea of group-based activity was approved and highly evaluated by JST.

3. Active Learning Seminar on Making High Precision Single Balance

3.1. *Preparation*

First idea of material development was preparation of pre-fabricated parts (Fig. 1). In technology classes, processing of each part is usually time consuming. For the seminar, we prepared a number of pre-fabricated parts to omit cutting and finishing process.

The pre-fabricated parts included a stand made by laser cut plywood, plate assembly made by metals, pivot assembly, thread (M6) rods of distinct sizes which will be used beams of balance, and two balancing weights of brass. Students will select favorite parts by themselves,

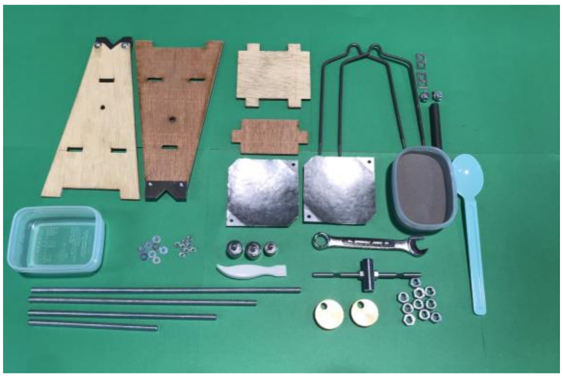


Fig. 1 Pre-fabricated parts.

putting their own ideas in the design, and directly can assemble precise single balance (Fig. 2).

Iron sand, weight (of, 10, 20, 30, 40, 50g) and a high accuracy scale for calibration were also prepared.

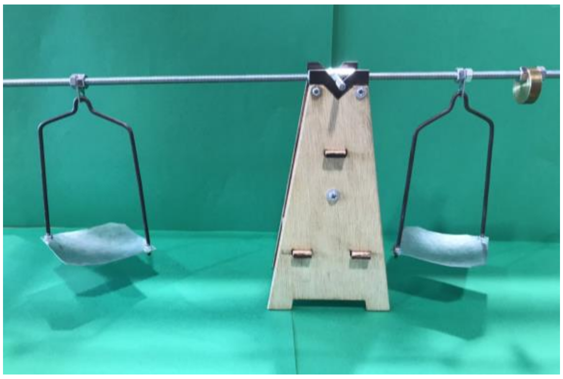


Fig. 2 Example of assembled single balance.

3.2. *Sequence of the seminar*

Sequence of the seminar was as follows. Lectures and exercises were combined. Breaktime was set in each hour since elementary school students are included.

13:10 - 13:30 A single balance and lever rule (lecture)
13:30 - 14:00 Assembling basic parts (exercise)
14:00 - 14:10 Break
14:10 - 14:30 Giving task and designing (lecture)
14:30 - 15:00 Assembling Balance (exercise)
15:00 - 15:10 Break
15:10 - 15:30 Calibrating weight (exercise)
15:30 - 16:00 Trial to improve performance (exercise)
16:00 - 16:10 Break
16:10 - 16:30 Presentation by each group
16:30 - 16:50 Summary

The most important key of the seminar was the "Task". We gave the task of,

"Measuring 80g using only one weight of 10, 20, 30, 40, 50g".

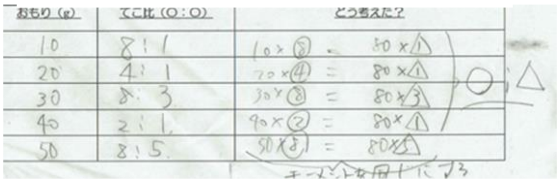
Note that, with this task the students had to choose single balance, that is, the length of left and right levers must be different. We here put a level of complexity, expecting more exercises in their brain, hand and group activities.

4. Results and Discussions

4.1. *Understanding basic concept*

Six groups of junior high school (JHS) students and three of elementary school (ES) students were participated to the seminar.

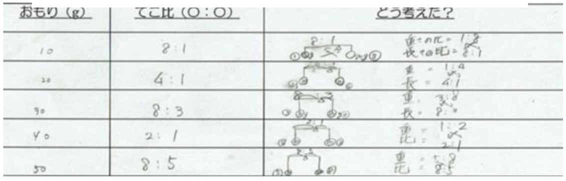
After giving preliminary lecture, the students found leverage could be used. Typical result from ES and JHS groups are shown in Fig. 3. Surprisingly enough, majority of the students could understand, including ES students, key concept of the balance (rule of moment) and found that every weight (10 to 50g) had a counterpart leverage. It looked that the group discussion helped to lead deeper understanding of the concept.



Weight Leverage How did you think?

おもり (g)	てこ比 (○:○)	どう考えた？
10	8:1	10×⑧ = 80×1
20	4:1	20×④ = 80×1
30	8:3	30×⑧ = 80×3
40	2:1	40×② = 80×1
50	8:5	50×⑧ = 80×5

(a) Elementary school group



Weight Leverage How did you think?

おもり (g)	てこ比 (○:○)	どう考えた？
10	8:1	重さの比 = 1:8 長さの比 = 8:1
20	4:1	重 = 1:4 長 = 4:1
30	8:3	重 : 3:8 長 : 8:3
40	2:1	重 : 1:2 比 : 2:1
50	8:5	重 : 5:8 比 : 8:5

(b) Junior high school group

Fig. 3 Calculating leverages with each weight.

There were differences in the way of leading leverages between ES and JHS groups. For ES students, leverages were led by counting numbers of weight, as seen in the figure. In contrast, many JHS groups found that the weight ratio was the inverse of leverages, using fractions.

4.2. *Assembling "Real" balance*

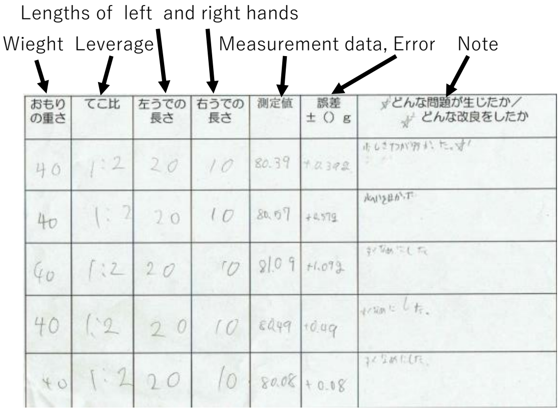
Understanding the key concept was not so difficult for them. However, projection of the concept onto the real single balance was much more difficult. First obstacle was the selection of leverage. The concept showed every weight could be used, but there was no hint which leverage was better than others. Moreover, they had to decide the concrete length of levers. Many groups used somehow linger time to decide them.

After deciding lever lengths, they start to assemble the balance and met the biggest hurdle, that was, problem of "dead weight".

Problems of the balance they met in daily science or mathematics classes always disregarded the dead weight, but real balance had. That is, only understanding main concept was not enough when one wanted to resolve the real problem. Majority of students did not clearly understand the concept of dead weight and role of counter weight. We had to teach dogmatically that the balance must be horizontal before the measurement, putting and moving counter weight in shorter, namely lighter lever. It was somehow disappointed result but was good chance to know difficulty of real problem resolution.

4.3. *Accuracy of the balance*

An example of calibration result for ES group are shown in Fig. 4. It indicates that they were gradually improve the balance and reduces the errors. This group finally reached an error of +0.08g, that is, 0.1% error. Final average error for three ES group was 0.48g, although every group uses the leverage of 1:2 and no dramatic improvements were performed.


Lengths of left and right hands
Wieght
Leverage
Measurement data, Error
Note

おもりの重さ	てこ比	左うでの長さ	右うでの長さ	測定値	誤差 ± ( ) g	☆どんな問題が生じたか／☆どんな改良をしたか
40	1:2	20	10	80.39	+0.39g	[illegible]
40	1:2	20	10	80.57	+0.57g	[illegible]
40	1:2	20	10	81.09	+1.09g	[illegible]
40	1:2	20	10	80.49	+0.49	[illegible]
40	1:2	20	10	80.08	+0.08	[illegible]

Fig. 4 An example of calibration result for ES group.

On the other hand, improvement of JHS groups were more dynamic. Many groups changed the leverage and lengths of the levers. Final leverages were 1:2 for two groups, 3:8 for 1group and 5:8 for three groups. Final average error in JHS was 1.71g, that was larger than ES result but still kept high accuracy.

5. Summary

An Active Learning seminar of "Making High Precision Single Balance" took place as a seminar in "Hiroshima Univ. Monozukuri Junior Doctor". Around 40 participants of young (11 - 15 in age) people were participated. Preparing pre-fabricated parts made easy to assemble single balance in short time. The group activity also helps to their affirmative cooperation. Although some concepts such as dead weight were difficult to take into accounts, they could assemble single balance of milligram accuracy. We believe that the seminar acts as a competency nurturing content.

Acknowledgements

This program was fully supported by JST fund of Junior Doctor Educational program.

References

1. Definition and Selection of Competencies: Theoretical and Conceptual Foundations, DeSeCo Annual Report 2001/spring 2002.
2. Ministry of Education, Culture, Sports, Science and Technology: Upper Secondary School Curriculum Guideline 2017-03 Notification (2018)
3. S. Bell : Project-Based Learning for the 21st Century: Skills for the Future, The Clearing House, vol. 883 (2010) pp.39-43.
4. Y. Li, K Wang, Y. Xiao and J. E. Froyd: Research and Trends in STEM Education: a systematic review of journal publications, Int. J. STEM Edu., (2020) pp.7-11.
5. H. Y. Suzuki and S. Nakamura: Development of Spring Scale as STEM Educational Materials using Step-by-Step Conceptual Designing 1 - Development of Electric Readout Circuit -, Technology Education Vol. 11 (2021) pp.1-7 [in Japanese].
6. H. Y. Suzuki and N. Masuda: Diaphragm Pump for STEM Educational Materials - Fabrication and Evaluation of Pump Made with Everyday Available Parts -, Technology Education Vol. 11 (2021) pp.8-16 [in Japanese].

Authors Introduction

Dr. Hiroyuki Y. Suzuki

He received his B.Eng. degree from Hiroshima University 1992, and his D.E. from Hiroshima University in 2000. He is currently an Associate Professor in the Dept. of Technology and Information Education, Graduate School of Humanities and Social Sciences at Hiroshima University. His research interest areas are education on materials science, engineering and processing and their application on educations, and STEM/STEAM educations.

Dr. Kazuo Kawada

He received his B.Eng. degree from Hiroshima University 1995, and his Ph.D. from Hiroshima University in 2005. He is currently an Associate Professor in the Dept. of Technology and Information Education, Graduate School of Humanities and Social Sciences at Hiroshima University. His research interest areas are the development of educational materials related to mechatronics education, data science education, and innovation human resource development for K-16.

Prof. Masayasu Nagamatsu

He received his B.Edu.degree from Hiroshima University in 1983, his M.Edu.from Naruto University of Education 1989 and his D.Eng. from Hiroshima University in 2015. He is a Professor in the Dept. of Technology and Information Education, Graduate School of Humanities and Social Sciences at Hiroshima University. His research interest areas are information education and technology education.

Capstone Class of “Mechatronics Innovation Project” as STEM Educational Curriculum for Teacher Training Course

Hiroyuki Y. Suzuki, Masayasu Nagamatsu, Kazuo Kawada
Graduate School of Humanities and Social Sciences, Department of Educational Science, Hiroshima University,
1-1-1-Kagamiyama, Higashi-Hiroshima City, Hiroshima 739-8524, Japan
E-mail: hiro-suzuki@hiroshima-u.ac.jp, kawada@hiroshima-u.ac.jp, nagamatu@hiroshima-u.ac.jp

Abstract

We have a set of STEM oriented special curriculum with name of “Mechatronics” in our teacher training course of “Program in Technology and Information Education, in School of Education, Hiroshima University”. The mechatronics classes start from first grade, and they are eventually integrated in a capstone class of “Mechatronics Innovation Project” in third grade. We set a different theme in each year, in which the mechanical, electronic and programming elements are integrated, and students are going to solve it by making automatic “mechatronics” machines. The theme in 2021 was wood climbing robots.

Keywords: Technology education, STEM, Practical work, Robot making

1. Introduction

Importance of practical work class is well known in educational curriculums of engineering programs in universities and colleges. After graduation, engineers are expected to be a mastery of their knowledges in their daily job, that is, only knowing “passively” respective rules and formulae is almost useless, one has to utilize “actively” and “practically” their knowledges and skills to resolve real problems. Therefore, it becomes ordinal to involve several practical work classes in curriculums of specialized engineering courses [1] [2].

Our course in Hiroshima university is teacher training course of technology and information subjects, not a merely for engineering, but we believe that the same kind of practical works are quite effective to develop teaching skills of technology and information as well. Practical work classes must not be restricted to specialized educations. Actually, it is much easier to grab technological notions and/or mechanisms by touching and manipulating real machines for every people regardless of ages and study levels.

The remaining issue is how to adopt currently introduced practical woks, although many of which are specialized for engineering course, to our curriculum. Number of methods were proposed heretofore, On the Job Training (OJT), Project (or Problem) Based Learning (PBL) [3], Active Learning (AL), STEM/STEAM Educations [4], and so on, indicating that there had not be found a standardized method applicable for all curriculums. From other point of view, nonetheless, we get the chance to develop original teaching program, adopting, mixing and modifying above mentioned methods to tailor to our student needs. We are going to explain, in this work, STEM oriented practical work classes introduced in our course, focusing on the terminal capstone class of “Mechatronics Innovation Project”.

2. Mechatronics Curriculum

Advantage of our course is that the students are going to study a wide variety of subjects in technology field, that is, all the mechanical, electronic, agricultural and information technologies are studied. In other words, it is involved many of STEM character in the curriculum. Therefore, for the purpose of integrating those elements, we introduced a set of STEM oriented special classes with name of "Mechatronics" in our teacher training course since 2010. There are in total of 4 mechatronics classes. First class starts from 1st semester just after the admission, with a simple theme of repairing bicycle, knowing forces act on the bicycle flame assembling simple models, and so on.

The theme of mechatronics classes becomes gradually complex ones, the 2nd one is focused on CAD/CAM, making small staff with free design. In the 3rd class, students design and fabricate hill-climbing remote-controlled cars in which the both mechanical and electrical elements, namely, all the STEM elements are contained and integrated into a real machine. After experiencing 3 classes of Mechatronics, they are eventually integrated into 4th capstone class of "Mechatronics Innovation Project (MIP)" in third grade.

3. Mechatronics Innovation Project (MIP)

For the MIP class, we set a different theme in each year which requires mechanical, electronics and programming elements with higher levels than previous classes, and students are going to resolve it by making automatic "mechatronics" machines. The theme in recent years are shown in Table 1. Recent trend in theme is that it contains Z axis movements which requires more mechanical challenges.

Table 1 Theme of Mechatronics Innovation Project in recent years.

Year	theme
2017	Wood climbing robot
2018	Anti-virus apparatus
2019	Step climbing rescue robot
2020	Automatic transfer robot
2021	Automatic cleaning robot

We have to carefully choose the theme and set the level of it. The theme must not be too difficult nor too easy. Actually, the level of difficulty can be controlled the size and shape of the field, as we will see later.

We place the MIP class in the first semester of the third year, taking four periods of time per week. 18 students (including 3 transfer students from technology colleges) were participated in 2021th class, and they were divided into 6 groups, (3 students per group). In addition, a group of graduated students was also participated to the MIP. Each student was assigned to one of a mechanical, electrical, or information part according to his or her favorite, and was responsible for designing and manufacturing of designated part.

3.1. *Overall flow of the MIP class*

The overall flow of the class was as follows, (1) presentation of the assignment and overview of the class, (2) group discussion, (3) project presentation, (4) submission of design documents, (5) ordering of parts, (6) processing and assembly, (7) performance evaluation, (8) final presentation, and (9) submission of a final report. The flow was unchanged in every year.

3.2. *Theme in 2021*

We announced the theme for 2021 MIP class as "Automatic Wood Climbing Robot", getting idea from commercially available wood climbing and branch cutting machines. The robots to be made were actually mimic of such machines but, after considering the level of the robot mechanisms, we excluded the branch cutting apparatus.

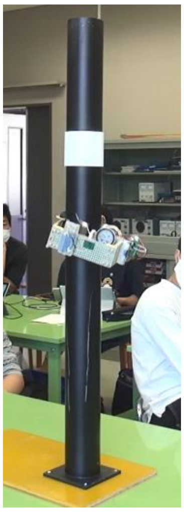


Fig. 1 Appearance of field for the wood climbing robot.

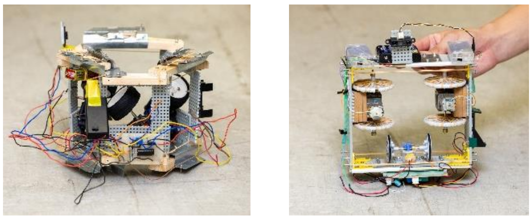
Basic technique for adjusting the level of the theme we used was changing the size and shape of the field. We usually called it "the Field", although 2021th field was merely vertically standing pipe (Figure 1). The size and complexity of the fields were most considerable point for the theme setting in every year.

We eventually made the field using vinyl chloride pipe (100 in diameter). In the midst of the pipe, a step was inserted as obstacles. The pipe was painted to black color. On the other hand, step and curve parts were remained original gray color to be noticed by sensors.

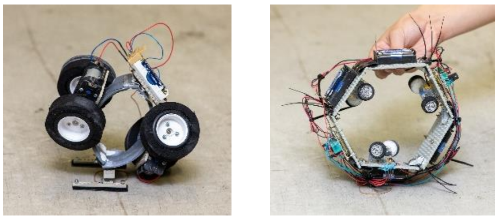
4. Results and Discussions

4.1. *Design plan until project presentation*

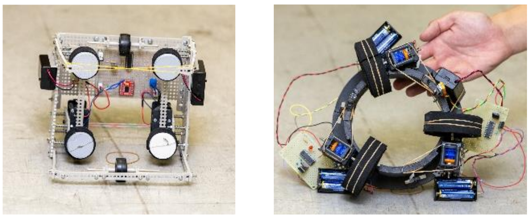
Until the idea presentation day, several unique mechanisms, for example a mimic of scale insect mechanism that would climb up a tree by extending and retracting the distance between two rings, were proposed. After the presentation and coming into concrete design process, in contrast, all groups eventually chose climbing mechanisms with tires or crawlers. The bodies of the robots were divided into two- or three-parts tied with springs or rubber bands. The divided bodies were expected to be enable to expand and contract when the robot pass through at the step.



(a)Robot A (b)Robot B



(c)Robot C (d)Robot D



(e)Robot E (f)Robot F

Fig. 2 Robots made by undergraduate student group (main participants of the MIP class).

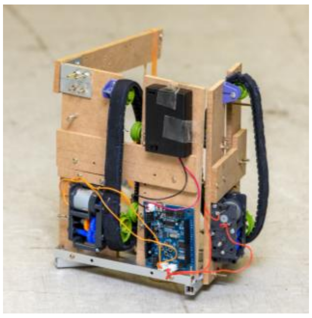


Fig. 3 Appearance of robot made by graduated student group.

4.2. *Designing and fabrication of robots*

Appearances the robots made by students are shown in Figure 2. Another robot was also made (Figure 3) made by a group of graduated students (1st grade in master course, experienced the MIP class two years ago). Design specifications of the robots are summarized in Table 2.

For designing process, each group considered the climbing mechanism and calculated traction power needed. Then, they were going to select body materials, actuators, gear boxes, and so on. Furthermore, control system with PIC, electric circuits and programming were designed and made. Each group could to put some of originality so that the robots got different appearances.

Table 2 Design specifications of robots.

Robot	Number of actuators	Traction parts	Setting angle of traction parts [deg]
A	3	tires	45
B	2	tires	0
C	2	tires	45
D	3	tires	45
E	2	tires	0
F	3	tires	0
G	2	crawler	0

Five robots used plastic, and other two used lumber for body construction, both of which were easy to be cut, shaped and assembled. For traction parts, undergraduate student groups chose rubber tires, whereas master students selected crawler. Size and number of tires were different among them.

Installation angle of tire or crawler were varied from 0deg. (4 group) to 45deg. (3 group). With 0deg. the robot climbed vertically straight, which requires more power in actuator so that those robots equipped stronger actuators of brushed DC motor with miniature gear box. In contrast, spiral climbing was realized with 45deg. robots, with relatively smaller motors, as we see in the model D, which equipped the smallest motors. The model D also equipped with the smallest tires which compensated the traction, at the expense of climbing speed.

One of disappointed point was the robots of undergraduate students could not to open, therefore they had to be put from top of the pipe, which was definitely useless in real situation. Only graduated students resolved that problem. Their robot could put the robot at the bottom of the pipe from the side.

5. Summary

STEM oriented practical work class suitable for teacher training course of technology and information was performed. The theme of the class was making wood climbing robot which contained all the mechanical, electrical and information elements. By controlling the level of the field, students could fabricate unique robots putting their own ideas in them. We believe that this kind of “active experiences” will be quite fruitful for their future job.

References

1. Kaneda T, Fujisawa S., Yoshida T, Yoshitani Y, Nishi T, Shidama Y, and Wasaki K: “Ensemble by Seven Musical Performance Robots,” Journal of Robotics and Mechatronics, (2001) Vol.13, No.1, pp. 50-55.
2. Kawazoe Y, Mitsuoka M and Masada S: “Practical Education Curriculum for Autonomous Mobile Robot (Project Learning Program for School Basedon Subsumption Architecture),” Journal of Robotics and Mechatronics, (2011) vol.23, no.5, pp.684–700.
3. S. Bell : Project-Based Learning for the 21st Century: Skills for the Future, The Clearing House, vol. 883 (2010) pp.39-43.
4. Y. Li, K Wang, Y. Xiao and J. E. Froyd: Research and Trends in STEM Education: a systematic review of journal publications, Int. J. STEM Edu., (2020) pp.7-11.

Authors Introduction

Dr. Hiroyuki Y. Suzuki

He received his B.Eng. degree from Hiroshima University 1992, and his D.E. from Hiroshima University in 2000. He is currently an Associate Professor in the Dept. of Technology and Information Education, Graduate School of Humanities and Social Sciences at Hiroshima University. His research interest areas are education on materials science, engineering and processing and their application on educations, and STEM/STEAM educations.

Prof. Masayasu Nagamatsu

He received his B.Edu. degree from Hiroshima University in 1983, his M.Edu. from Naruto University of Education 1989 and his D.Eng. from Hiroshima University in 2015. He is a Professor in the Dept. of Technology and Information Education, Graduate School of Humanities and Social Sciences at Hiroshima University. His research interest areas are information education and technology education.

Dr. Kazuo Kawada

He received his B.Eng. degree from Hiroshima University 1995, and his Ph.D. from Hiroshima University in 2005. He is currently an Associate Professor in the Dept. of Technology and Information Education, Graduate School of Humanities and Social Sciences at Hiroshima University. His research interest areas are the development of educational materials related to mechatronics education, data science education, and innovation human resource development for K-16.

Effect of Spraying Dispersion Using UAV Spraying System with Different Height at Paddy Field

Hazry Desa, Muhammad Azizi Azizan, Nik Noriman Zulkepli, Nurfadzillah Ishak, Teh Xi Hang, Siti Syuhaidah Yahya, Aisyah Arina Mohammad Shahrazel, Fakhrul Mukmin Mansor
Centre of Excellence for Unmanned Aerial Systems (COE-UAS), Universiti Malaysia Perlis (UniMAP), Malaysia

Siti Zaleha Abdul Aziz
Department of Robotic & Automation, Mara Japan Industrial Institute (MJII), Malaysia

Abadal-Salam T. Hussain
Department of Medical Instrumentation Engineering Techniques, Al-Kitab University, Altun Kupri, Kirkuk, Iraq

E-mail: hazry@unimap.edu.my, aziziazizan@unimap.edu.my, niknoriman@unimap.edu.my,
nurfadzillah@unimap.edu.my, tehxh97@gmail.com, syuhaidah323@gmail.com, aisyahaa2300@gmail.com,
keranamumalaya@gmail.com, szaleha.aziz@mara.gov.my, asth2@uoalkitab.edu.iq
www.unimap.edu.my

Abstract

This study investigated the UAV spraying system height in relation to spraying uniformity and dispersion. The operating heights of the UAV spraying system at heights of 1 m, 1.5 m, and 2 m from the hollow cone nozzles were investigated within a wind speed of 2.8 m/s. The tests were to determine the spray uniformity and dispersion on the water sensitive paper that was placed on the paddy plant. The results of water droplet samples were evaluated using ImageJ software. The results show the droplet distribution at 1.5 m height has high values for average droplet density, which is 162.7 deposits/cm² at the top area and 161.8 deposits/cm² at the bottom area. The percentage of coverage was also high, at 55.21% at the top area and 51.4% at the bottom area.

Keywords: UAV, Spraying System, Heights, Droplet Density, Average Coverage

1. Introduction

Throughout the Eleventh Malaysian Plan (2016–2020) and the National Agro-Food Policy (2011–2020), Malaysia continues its proactive and progressive measures to promote paddy and rice sector development. Agro-chemicals are being used on crops to control pests and weeds through conventional spraying systems with little consideration of substantial variation in plant populations and canopies [1]. It is estimated that about 3 million metric tons of pesticides are used annually worldwide to control the disease [2]. These chemicals not only control weeds but also help to control insect pest attacks that cause a reduction in crop yield and affect the crop quality as well [3]. However, the World Health Organization reported that around one million illness cases resulted from the manual spraying of pesticides in the fields [4].

Alternatively, UAV spraying systems are being used for crop monitoring and management for spraying pesticides. However, the proper spray height and nozzle opening need to be investigated. Investigations are needed as there are many problems in the area, such as non-uniform spray or late spraying [5], poor penetrability and poor distribution to the crop. A study conducted by [6] revealed that UAV spraying systems could increase efficiency by more than 60% with a 20–30% decrease in pesticide dose. Furthermore, the self-adjustment routing mechanism of UAVs can significantly reduce pesticide waste and fertilizer application [7]. [8] suggested that UAV spraying systems should be carefully tested to ensure application accuracy. It was also suggested to fly at a low altitude of 1 to 5 m. That will help in avoiding spray drift.

However, there is no study available on different flight configurations for altitude and spray nozzle openings that might be helpful for farmers to optimize the UAV spraying system. Such investigations will reduce the risks of overuse of pesticides due to their outside drift as well as overlap of areas under spray [1].

In the current study, an effort has been made to evaluate spraying dispersion through the UAV spraying system at different heights and to estimate the droplet dispersion, especially on paddy fields, to overcome these difficulties.

The findings will be helpful in recommending efficient spraying for paddy fields through UAV spraying systems as well as an alternative way to come up with standard guidelines for the incorporation of drones in agricultural operations.

2. Materials and methods

2.1. *Testing method and data collection*

This study was conducted at the open field near to the Centre of Excellence for Unmanned Aerial Systems (COEUAS), Universiti Malaysia Perlis (UniMAP), Malaysia (Latitude: 6.43744 N, Longitude: 100.18868 E). The total area of the field was 140 m x 70 m.

The UAV spraying system sprayer was tested at three different altitudes of 1 m, 1.5 m, and 2 m in height: wind speed of 2.8 m/sec with an average temperature of 33 °C and humidity of 84%. A hollow cone nozzle with 100% openings was used in this test. Water-sensitive papers (7.6 cm x 2.6 cm) were placed at the top and bottom areas of the paddy plant. The top area is the upmost part of the plant, whereas the bottom area is within 20 cm of the ground. The goal is to evaluate the water droplet sample obtained from the spraying nozzle, as shown in Figure 1.

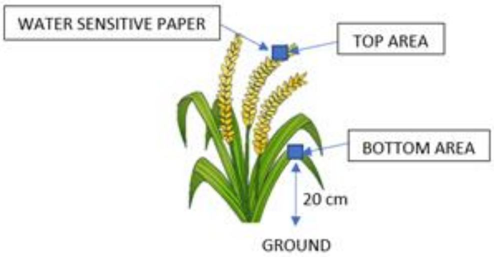


Fig. 1. Water sensitive paper placement

Attention was given to where the water-sensitive papers were placed at a 30 cm distance between the left and right areas of the main point paddy plant to get drone spray impressions (dots/marks) as shown in Fig. 2 and Fig. 3 below. A similar testing method to that in Fig. 2 was suggested by [9].

The UAV spraying system was flown at a speed of 12 m/s from point A to B and was controlled by IFLY application software with 10 repetitions for the hollow cone nozzle. Then, the samples were collected, and the paper was checked to be dried before the next procedure. The same methods were applied to the different heights of nozzle. This testing involved a total of 300 samples, with 100 samples for each type of nozzle.

The spray droplet impressions on water-sensitive paper were observed by using ImageJ software (DepositScan). The number of spray dots per cm² was counted. If the number of fine spray dots were equal to or greater than 100, then the spray was considered uniform. The number of spray dots per cm² should be equal to or greater than 100, depending on the level of height, wind speed and spray nozzle opening. This is the recommended value for uniformed spraying. On the other hand, if the number of spray counts is less than 100, the height level, wind speed, and spray opening nozzle are not the recommended values for uniformed spraying. As such, this would lead to non-uniformed spray and a waste of resources [9].

The spray dispersion was evaluated by analyzing three parameters: (i) average droplet density on water sensitive paper by counting the number of droplets per unit area (Deposits/cm²), (ii) average coverage, which represents the percentage of water sensitive surface covered by the droplet (%), and (iii) total deposit counted, which represents the total of droplet deposition distribution in the target area.

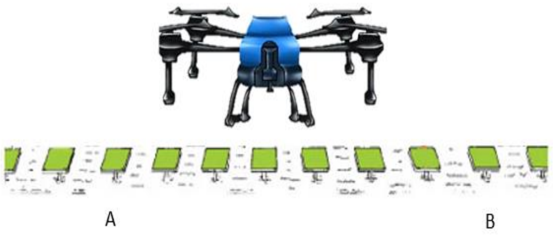


Fig. 2. Layout of WSP sampling

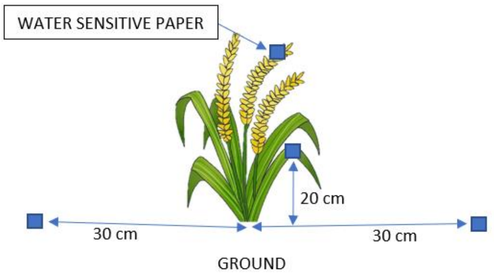


Fig. 3. WSP distances from the main plant

2.2. *Unmanned Aerial Vehicle (UAV) Spraying System Specifications*

In this study, the tested UAV spraying system is the HSSB10L-606 Sprayer Drone, as shown in Fig. 4, which has six arms, six motors, and a hollow cone type of spray nozzle. The complete specifications of this system are provided in Fig. 5 below. The spraying tank capacity was 10 liters with a UAV take-off payload capacity of 28 kg. This system had smooth take-offs and landings in its autonomous mode with a flying speed of 0–12 m/s. This UAV spraying system software interface used the IFLY application for instructions of flight height, spray span for overlapping or precise spray, flight speed, and turn time.

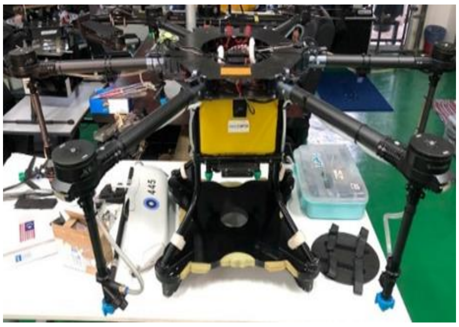


Fig. 4. HSSB10L-606 sprayer drone

UAV part	Description
Number of arms	6
Tank capacity	10 Litres
Maximum take-off capacity	28 kg
Flying time	10-15 minutes
Flying height	0-30m
Flying speed	0-12 m/s
Spray speed	0-8 m/s
Spray width	>4-6m
Spray flow	1-1.15 L/min
UAV size	2.0m*1.3m*0.45m

Fig. 5. UAV spraying system specification

3. Results and discussions

Spray uniformity and dispersion were tested with hollow cone nozzles at different heights. The spraying dots imprinted on the WSP were scanned using a scanner and uploaded to DepositScan software to determine the average number of droplets in the sample data. Tests were performed at 1 m, 1.5 m, and 2 m altitudes with a 100% nozzle opening and a wind speed of 2.8 m/sec. As shown in Fig. 6 below, the spray dots imprinted on the WSP have different working heights. The results of deposition analysis for this observation are provided in Fig. 7 and Fig. 8 below.

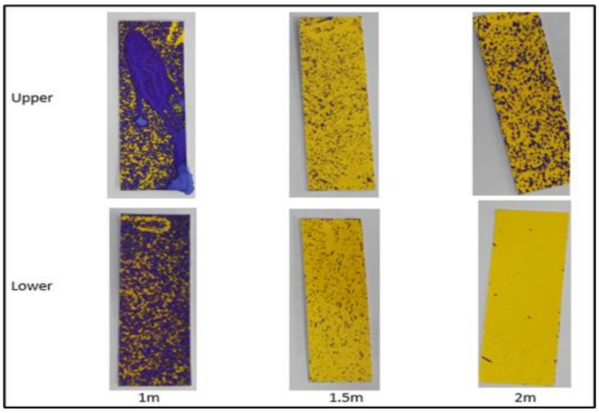


Fig. 6. Spraying pattern with three different working heights

Nozzles	Height (meter)	Sampling Site	Average Droplet Density (Deposits/ cm²)	Average Coverage (%)	Total Deposit Counted	Average coverage differences (%)
Hollow Cone Nozzle	1	Top Area	79.2	29.71	369	15.5
		Bottom Area	63.7	24.53	320	
	1.5	Top Area	162.7	55.21	1003	0.9
		Bottom Area	161.8	51.4	885	
	2	Top Area	45.7	12.58	214	6.9
		Bottom Area	38.8	3.46	171	

Fig. 7. Dispersion of a UAV spraying system at 100% nozzle opening and 2.8m/sec wind speed

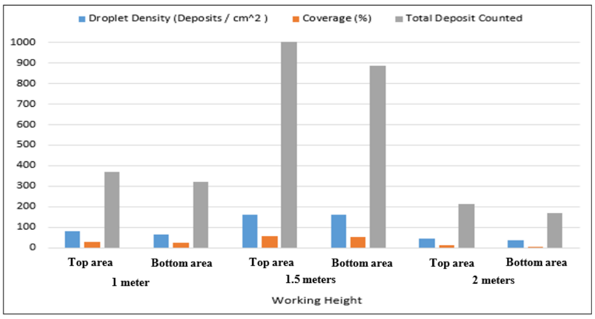


Fig. 8. Droplet deposition distribution

Fig. 8 shows the bar graph of spraying distribution for droplet density against working height from the nozzle. Based on Fig. 7, the hollow cone nozzle at 1.5 m height had the highest average droplet density on WSP, which was 162.7 deposit/cm² at the top and 161.8 deposit/cm² at the bottom area, with an average coverage of 55.2% at the top and 51.4% at the bottom area. Then, the total deposit count, which represents the total of droplet deposition distribution in the target area, was the highest, at 1003 at the top area and 885 at the bottom area. The percentage difference between the average coverage was the lowest value amongst working heights, at 0.9%, which found that the droplet successfully penetrated from the top area to the bottom area. In 1914, [10] proposed the general use of hollow-cone nozzles to apply pesticides or fungicides to field crops when foliage penetration and complete coverage of the leaf surface are required. The spray drift potential is higher from hollow-cone nozzles than from other nozzles due to the small droplets produced. Generally, this type of nozzle should not be used to apply herbicides. It is also stated by [11] that the spray can be effectively done with 100% nozzle openings at wind speeds less than 6 m/sec with a flying height of 1.5 m.

Then, the tested hollow cone nozzle at 2 m demonstrated a non-uniform pattern or less on spray distribution of droplet density, which is 79.0 deposit/cm² at the top area and 63.7 deposit/cm² at the bottom area, and the total deposit is counted by the lowest number between 214 at the top area and 171 for the bottom area. Because of the faulty height, the percentage difference between the average coverage was 6.9%, indicating that the droplet could not be distributed evenly or penetrated to the ground area. As stated by [11], a working height of 2 m is efficient at working nozzle opening at 75% with a wind speed of less than 5 m/sec. To avoid spray losses, the working height should be less than 2 m if the wind speed is greater than 5 m/sec.

The result also showed the highest differences for average coverage between the top area and the bottom area, at 15.5%, which found that the top area reached more droplets density compared to the bottom area. The effect of spraying dispersion at a working height of 1 m is poor. The WSP was drenched by too much drip. This is fine for water application but not suitable for pesticide application. Overdosing of the pesticide's chemical may also damage the paddy and waste the resource. This process is also similarly stated by [12] in their research paper that the droplet distribution in a 1.1 m canopy area is poor for pesticide spraying purposes, but it can be improved by increasing the spray volume in the future.

4. Conclusion

In this study, three flight heights of spraying nozzles were carried out by using a hollow cone nozzle with a 100% flow rate opening under wind conditions in a paddy field. The average droplet density, average coverage, and total deposit distribution of droplets in the target area between the top area and the bottom area were compared and analyzed in this research. The conclusions are shown as follows:

1. The average droplet density of droplet distribution in the target area was influenced by the height of the nozzle, windspeed, and flow rate. There were significant differences in the droplet distribution rate in the target area between the three-flight heights. The average droplet density at 1.5 m flight height had the highest and excellent result because the deposit density was more than 100 deposits/cm². This finding is in line with the study by [9]. The percentage of average droplet difference between the top area and the bottom area was also the lowest at 0.9%, which concluded that the droplet was successfully penetrated from the top area to the bottom area.
2. The average coverage results of the droplets were influenced by the total droplet distribution, which also depends on the flight height. The droplet dispersion at 1.5 m flight height had the highest percentage of average coverage at 55.21% for the top area and 51.5% for the bottom area.
3. The highest total deposit counted for the droplets in the target area of dispersion at a height of 1.5 m had the highest number of 1003 for the top area and 885 for the bottom area, which indicated a suitable working height of the nozzle for better droplet distribution.

The experiment demonstrated that the flight height of the nozzle is one of the most important factors that affect the droplet distribution and drift for pesticide spraying of UAVs. Experimental results revealed that, at 1.5 m flight height, with 100% nozzle opening at 2.8 m/sec wind speed, results in uniformed spraying results. Awareness and consideration for the use of a hollow cone nozzle with an appropriate flying height is critical for better spraying dispersion and droplet distribution while avoiding pesticide drift harm to the environment and humans. Consequently, the type of nozzle from the different flying heights, nozzle angle, and orientation should be studied for better precision in agricultural aerial spraying.

Acknowledgment

The authors would like to thank the Universiti Malaysia Perlis (UniMAP) for providing the facilities and financial support under Pre-Commercialisation Grant, Grant No: 9001-00639.

References

1. Faical, B.S., H. Freitas, P.H. Gomes, L.Y. Mano, G. Pessin, A.C. deCarvalho, B. Krishnamachari and J. Ueyama., "An adaptive approach for UAV-based pesticide spraying in dynamic environments", pp. 210-223, 2017.
2. Pimentel, D., "Pesticides and pest control. In: R. Peshin and A.K. Dhawan (eds.), Integrated Pest Management: Innovation-development process", 2005.
3. Tudi, M., Daniel Ruan, H. Wang. L., Lyu. J., Sadler. R., Connell. D., Chu. C., Phung. D.T., "Agriculture Development, Pesticide Application and Its Impact on the Environment", Int. J. Environ. Res. Public Health, pp. 2-23, 2021.
4. Mogili, R.U.M., B.B.V.L. Deepak, "Review on application of drone system in precision agriculture", International Conference on Robotics and Smart Manufacturing (RoSMa2018), Procedia Comput. Sci., pp. 502-509, 2018.

5. Hussain, S., Cheema, M. J. M., Arshad, M., Ahmad, A., Latif, M. A., Ashraf, S., & Ahmad, S, "Spray uniformity testing of unmanned aerial spraying system for precise agro-chemical applications", Pakistan Journal of Agricultural Sciences, Vol 56, No 4, pp. 897–903, 2019.
6. Qin, W., X. Xue, S. Zhang, W. Gu and B. Wang., "Droplet deposition and efficiency of fungicides sprayed with small UAV against wheat powdery mildew", Int. J. Agric. Biol. Eng. pp. 27-32, 2018.
7. Faical, B.S., F.G. Costa, G. Pessin, J. Ueyuma, H. Freitas, A. Colombo, P.H. Fini, L. Villas, F.S. Osorio, P.A. Vargar, T. Braun, "The use of unmanned aerial vehicles and wireless sensor networks for spraying pesticides", J. Syst. Architect, pp. 393-404, 2014.
8. Gao, Y., "Study on distribution of pesticide droplets in gramineous crop canopy and control effect sprayed by unmanned aerial vehicle, Northeast Agric. Univ. Harbin, China", 2013.
9. Yanliang, Z., L. Qi, Z. Wei. "Design and test of a six-rotor unmanned aerial vehicle (UAV) electrostatic spraying system for crop protection", Int. J. Agric. Biol. Eng, pp. 68-76, 2017.
10. Johnson, M. P., Swetnam, L. D., & Engineering, A. (n.d.), "Sprayer nozzles: selection and calibration", Agricultural Engineering.
11. Hussain, S., Cheema, M. J. M., Arshad, M., Ahmad, A., Latif, M. A., Ashraf, S., & Ahmad, S, "Spray uniformity testing of unmanned aerial spraying system for precise agro-chemical applications", Pakistan Journal of Agricultural Sciences, Vol 56, No 4, pp. 897–903, 2019.
12. Chen, P., Lan, Y., douzals, J.-P., Ouyang, F., Wang, J., & Xu, W., "Droplet distribution of Unmanned Aerial Vehicle under several spray volumes and canopy heights in the cotton canopy", International Journal of Precision Agricultural Aviation, pp. 74–79, 2018.

Authors Introduction

Prof. Dr. Hazry Desa

He received his Bachelor of Mechanical Engineering from Tokushima University, Japan. He pursued his Ph.D. at the Artificial Life and Robotics Laboratory, Oita University, Japan in 2003. He is also the Director at Center of Excellence for Unmanned Aerial Systems (COEUAS), Universiti Malaysia Perlis, Malaysia.

Dr. Muhammad Azizi Azizan

He received his PhD degree in Civil Engineering from Universiti Malaysia Perlis and is currently working as a Senior Lecturer at the same institution. He is also co-founder and director of Haas Solutions Sdn Bhd, a technology startup company of Universiti Malaysia Perlis.

Dr. Nik Noriman Zulkepli,

He received his PhD in polymer engineering from Universiti Sains Malaysia in 2011. He is currently working as a Senior Lecturer in the Polymer Engineering Division, School of Materials Engineering, Universiti Malaysia Perlis. He is also Head of Cluster (Materials) at the Centre of Excellence for Unmanned Aerial Systems.

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 Ecellence for Unmanned Aerial System (COEUAS), Universiti Malaysia Perlis.

Mr. Teh Xi Hang

He received his Bachelor's Degree in Mechatronics Engineering in 2022 from the Faculty of Electrical Engineering Technology, Universiti Malaysia Perlis (UniMAP) in Malaysia.

Ms. Siti Syuhaidah Yahya

She received her Bachelor of Engineering (Honours) in Electronics Engineering from the Universiti Malaysia Perlis (UniMAP), Malaysia.

Ms. Aisyah Arina Mohammad Shahrazel

She received her Bachelor of Engineering (Honours) in Manufacturing Engineering from the Universiti Malaysia Perlis (UniMAP), Malaysia.

Mr. Fakhrul Mukmin Mansor

He received his Bachelor of Electronic Engineering Technology in Electronic Network Design with Honours from the Universiti Malaysia Perlis (UniMAP), Malaysia.

Mrs. Siti Zaleha Abdul Aziz

She received her Master's degree in Green & Energy Efficient Building in 2014, a joint award between UniKL, Malaysia and the University of Applied Sciences Rosenheim, Germany. She is currently a Vocational Training Officer at Mara Japan Industrial Institute (MJII), Beranang.

Mr. 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.

Spraying Dispersion Analysis with Different Nozzle Types Using a UAV Spraying System in a Paddy Field

Hazry Desa, Muhammad Azizi Azizan, Nik Noriman Zulkepli, Nurfadzillah Ishak, Tan Yew Tian, Siti Syuhaidah Yahya, Aisyah Arina Mohammad Shahrazel, Fakhrul Mukmin Mansor
Centre of Excellence for Unmanned Aerial Systems (COE-UAS), Universiti Malaysia Perlis (UniMAP), Malaysia

Siti Zaleha Abdul Aziz
Department of Robotic & Automation, Mara Japan Industrial Institute (MJII), Malaysia

Abadal-Salam T. Hussain
Department of Medical Instrumentation Engineering Techniques, Al-Kitab University, Altun Kupri, Kirkuk, Iraq

E-mail: hazry@unimap.edu.my, aziziazizan@unimap.edu.my, niknoriman@unimap.edu.my,
nurfadzillah@unimap.edu.my, yewtian0414@gmail.com, syuhaidah323@gmail.com, aisyahaa2300@gmail.com,
keranamumalaya@gmail.com, szaleha.aziz@mara.gov.my, asth2@uoalkitab.edu.iq
www.unimap.edu.my

Abstract

This study investigates the ability of Unmanned Aerial Vehicle (UAV) spraying systems to be used as an agriculture spraying method in Malaysia. The operating height of the UAV was 1.5 m with three different nozzles were investigated within a wind speed of 1.15 m/s to determine spray uniformity and dispersion in the paddy field conditions. The results from these samples were evaluated by using ImageJ software. The results show that the droplet distribution by using an electrostatic centrifugal nozzle has a high average droplet density, which is 134.03 deposits/cm2 for the top area and 153.93 deposits/cm2 for the bottom area. The electrostatic centrifugal nozzle also testified to the high value of total droplet deposit at 3478 for the top area and 3255 for the bottom area.

Keywords: UAV, Spraying System, Nozzles, Droplet Density, Average Coverage

1. Introduction

The Malaysian government has spent millions to develop the agricultural sector, mainly to enhance productivity, increase farmers' income, and provide employment [1]. Farmers have been using conventional techniques for seed planting, composting, pesticide application, etc. The traditional techniques used for pesticide and fertilizer spraying not only require more time but also are less effective, so there is a need for technological advancement in this field [2]. Manual spraying operations can be very difficult because of the crop's height.

Therefore, smart farms use drones for spraying, which reduces human contact with fertilizers, pesticides, and other harmful chemicals [3]. Alongside, [1] also reported that farmers' complaints of breathing difficulties during and after spraying of pesticides (51.5%), itchiness and soreness (26%), as well as rashes and peeling of skin on their hands (13.7%). To a more severe extent, there were incidences of farmers collapsing, experiencing stomach aches, vomiting, and being admitted to the hospital. Pesticide exposure was associated with respiratory symptoms such as coughing, wheezing, and airway inflammation [4].

Nowadays, the application of UAV spraying systems is growing at a very fast rate in agribusiness [2]. In this situation, the UAV spraying system is frequently utilized on farms to help farmers as a part of "Precision Agriculture" to modernize farming in developed countries [5]. Using the UAV spraying system for applying pesticides and fertilizer is an exuberant contraption. The UAV spraying system adequately reduces the rate of health dilemmas and the number of workers, which is quite an impressive landmark [6]. The proper selection of a nozzle type and size is essential for proper pesticide application. However, there have been issues reported with nozzle problems in determining the amount of spray applied to an area, the uniformity of application, the coverage obtained on the target surface, and the amount of potential drift. Nozzles are important not only to break the disposed liquid into droplets but also to form the spray pattern and to propel the droplets in proper directions [7].

In relation to the various problems of the nozzle, this study is aimed at determining the three different types of nozzles' application in the paddy area and estimating the droplet dispersion towards the paddy field to overcome these difficulties.

2. Materials and methods

2.1. *Testing method and data collection*

This study was conducted at the open field near the Centre of Excellence for Unmanned Aerial Systems (COEUAS), Universiti Malaysia Perlis (UniMAP), Malaysia (Latitude: 6.43744 N, Longitude: 100.18868 E) in June 2022. The UAV spraying system was tested at 1.5 m height with a wind speed of 1.15 m/sec, an average temperature of 33° C, and a humidity of 84%. The total area of the field was 140 m x 70 m.

The duration of this test was 2 weeks. The three different types of nozzles that were tested were flat fan, hollow cone, and electrostatic centrifugal with 100% openings. Water sensitive papers (WSP) (7.6 cm x 2.6 cm) were placed with a 1 m gap at the top area and bottom area of the paddy plant to evaluate the water droplets sampled received from the spraying nozzle, as shown in Figure 1.

The water-sensitive papers were placed in the middle of the paddy plant between the upper and lower leaf surfaces to see how far the droplets could penetrate.

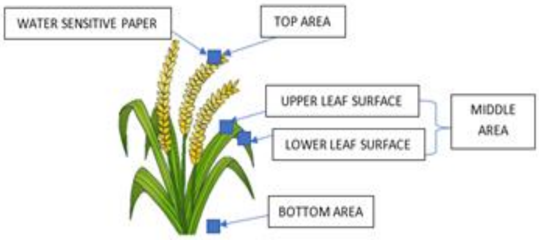


Fig. 1. Water sensitive paper placement

2.2. *Unmanned Aerial Vehicle (UAV) Spraying System Specifications*

In this study, the tested UAV spraying system is the HSSB10L-606 Sprayer Drone, as shown in Fig. 2, which has six arms, six motors, and three types of spray nozzles for this study. The complete specifications of this system are provided in Fig. 3. The spraying tank capacity was 10 liters, with a drone take-off payload capacity of 28 kg. This system had smooth take-offs and landings in its autonomous mode with a flying speed of 0–12 m/s. This UAV spraying system software interface used the IFLY application for instructions of flight height, spray span for overlapping or precise spraying, flight speed, and turn time.

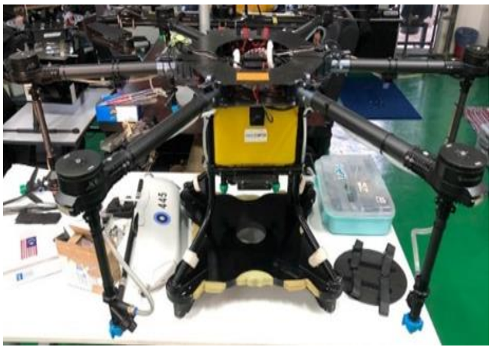


Fig. 2. HSSB10L-606 sprayer drone

UAV part	Description
Number of arms	6
Tank capacity	10 Litres
Maximum take-off capacity	28 kg
Flying time	10-15 minutes
Flying height	0-30m
Flying speed	0-12 m/s
Spray speed	0-8 m/s
Spray width	>4-6m
Spray flow	1-1.15 L/min
UAV size	2.0m*1.3m*0.45m

Fig. 3. UAV spraying system specification

3. Results and discussions

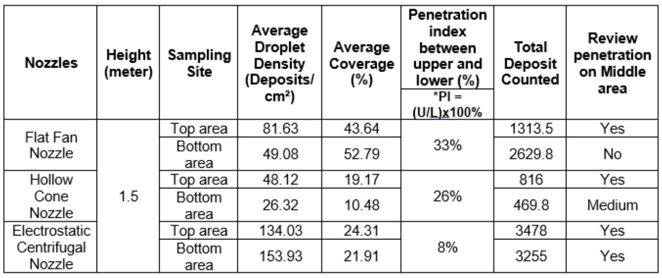
Spray uniformity and dispersion were tested with different nozzles at the same height. The spraying dots imprinted on the water-sensitive paper were scanned using a scanner and uploaded to DepositScan software to determine the average number of droplets on the sample data. Uniformity at 1.5 m altitude with 100% nozzle opening was observed. The results of the deposition analysis for the flat fan nozzle, hollow cone nozzle, and electrostatic centrifugal nozzle are provided in Fig.4 and Fig. 5 below.

Based on Fig. 4, the electrostatic centrifugal nozzle had the highest average droplet density on water-sensitive paper, which was 134.03 deposit/cm² at the top area and 153.93 deposit/cm² at the bottom area, with an average coverage of 24.31 % at the top area and 21.91 % at the bottom area. Then, the total deposit count, which represents the total droplet deposition distribution in the target area, is the highest, which is 3478 at the top area and 3255 at the bottom area.

The penetration index was the lowest value amongst nozzles for droplets to reach from the top to bottom area, at 8%, which was considered that the droplet could penetrate easily to the ground and lower leaf surface that had difficult areas to reach. According to Fig. 4, the electrostatic centrifugal nozzle showed excellent droplet distribution from the top to bottom area for paddy plants because of the smallest mist. [8] proposed that electrostatic forces on tiny droplets are more prominent than gravitational forces; thus, the electrostatic charge of mist droplets can lead to better deposition with less drift.

Next, the test on the flat fan nozzle demonstrated less spray distribution on top but received lots of droplet density on the bottom, which is 1313.5 at the top area and 2629.8 at the bottom area. This nozzle can be rated as defective at all nozzle heights compared to the hollow cone nozzle. The result also revealed the highest penetration index between the top area and bottom areas at 33%, indicating that the bottom area reached more droplets during the spraying process. This process is also similarly stated in [9] research paper that a flat fan nozzle needed improvement in penetration into the crop canopy at an 80-degree nozzle angle.

However, the hollow cone nozzle showed the middle percentage index of penetration between the top area and bottom area, which is 26% with an average droplet density of 48.12 deposit/cm² at the top area and 26.32 deposit/cm² at the bottom area. The total deposit is determined by the lowest number between the nozzles, which is 816 at the top area and 469.8 at the bottom area, indicating that the droplet cannot be distributed evenly or penetrated to the ground area due to the faulty height. [9] also stated that the hollow cone nozzle is suitable for spraying the plant at the early growing stage from the emergence till tillering stage as the plant is still not too high and the density is medium, but after this stage, the hollow cone nozzle is not suitable for use.



Nozzles	Height (meter)	Sampling Site	Average Droplet Density (Deposits/ cm²)	Average Coverage (%)	Penetration index between upper and lower (%) *PI = (U/L)x100%	Total Deposit Counted	Review penetration on Middle area
Flat Fan Nozzle	1.5	Top area	81.63	43.64	33%	1313.5	Yes
		Bottom area	49.08	52.79		2629.8	No
Hollow Cone Nozzle		Top area	48.12	19.17	26%	816	Yes
		Bottom area	26.32	10.48		469.8	Medium
Electrostatic Centrifugal Nozzle		Top area	134.03	24.31	8%	3478	Yes
		Bottom area	153.93	21.91		3255	Yes

Fig. 4. UAV spraying system dispersion by using three different nozzles

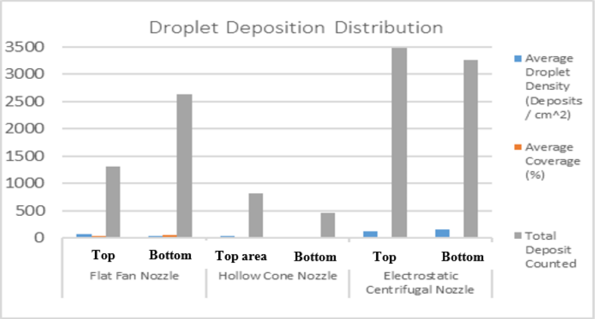


Fig. 5. Droplet deposition distribution

4. Conclusion

In this study, three types of spraying nozzle tests with the same height and same flow rate in a paddy field were carried out. The average droplet density, average coverage, total deposit distribution, and penetration index of droplets in the target area between top and bottom were compared and analyzed in this research. The conclusions are shown as follows:

1. The average droplet density and penetration results of droplet distribution in the target area were influenced by the type of nozzle, height, and flow rate. There were significant differences in the droplet distribution rate in the target area between the three nozzle tests. The average droplet density rate with an electrostatic centrifugal nozzle of 134.03 deposits/cm² at the top area and 153.93 deposits/cm² at the bottom area was the lowest and excellent penetration of the droplet, which is 8%. The droplet distribution rate from the top to bottom of the rice canopy was improved in the target area by using the better type of nozzle.
2. The average coverage results of droplets did not increase the total droplet distribution due to an increase in droplet density, and that of the droplets by using an electrostatic centrifugal nozzle with an average coverage of 24.31% for the top area and 21.91% for the bottom area was the lowest difference and greatest for droplet distribution starting from the rice canopy until to the ground area.
3. The highest total deposit counted of the droplets in the target area by using an electrostatic centrifugal nozzle was 3478 for the top area and 3255 for the bottom area, which indicated that the better droplet distribution was successfully distributed to the target area because of the smallest mist.

The experiment demonstrated that droplet size is one of the most important factors affecting droplet distribution and drift for pesticide spraying. For the application of UAV spraying systems with better spraying, to reduce droplet drift and improve the droplet distribution, the use of nozzles at an appropriate height should be considered downwind of the spraying field to avoid harm to the plant caused by pesticide drift. Furthermore, the shape of the spray pattern profile of a nozzle depends on the type and capacity of the nozzle utilized, which in turn is influenced by the pressure at the nozzle, the height of the nozzle from the spray surfaces, and the angle at which the nozzle is oriented. These should be studied to achieve precision agricultural aerial spraying in the future.

Acknowledgment

The authors would like to thank the Universiti Malaysia Perlis (UniMAP) for providing the facilities and financial support under Pre-Commercialisation Grant, Grant No: 9001-00639.

References

1. Fuad, M. J. M., Junaidi, A. B., Habibah, A., Hamzah, J., Toriman, M. E., Lyndon, N., Er, A. C., Selvadurai, S., & Azima, A. M.,"The Impact of Pesticides on Paddy Farmers And Ecosystem", Advances in Natural and Applied Sciences, 6(1), pp. 65–70, 2012.
2. Hafeez, A., Husain, M. A., Singh, S. P., Chauhan, A., Khan, M. T., Kumar, N., Chauhan, A., & Soni, S. K.," Implementation of drone technology for farm monitoring & pesticide spraying", A review. In Information Processing in Agriculture. China Agricultural University., 2022.
3. Pathak, H., Kumar, A.K., Mohapatra, S.D., Gaikwad, B.B., "Use of Drones in Agriculture: Potentials, Problems and Policy Needs", ICAR-NIASM 12(2), pp. 12-23, 2020.
4. Rudzi, S. K., Ho, Y. bin, Tan, E. S. S., Jalaludin, J., & Ismail, P., "Exposure to Airborne Pesticides and Its Residue in Blood Serum of Paddy Farmers in Malaysia", International Journal of Environmental Research and Public Health, 19(11), 2022.
5. Debangshi, U., "Drone-Applications in Agriculture Drones-Applications in Agriculture", 2021.
6. Rahman, M. F. F., Fan, S., Zhang, Y., & Chen, L.; "A comparative study on application of unmanned aerial vehicle systems in agriculture", In Agriculture (Switzerland), Vol. 11, Issue 1, pp. 1–26, 2021.

7. Johnson, M. P., Swetnam, L. D., & Engineering, A. (n.d.), "Sprayer nozzles: selection and calibration", Agricultural Engineering.
8. M. Ahmed and M. S. Youssef, "Characteristics of mean droplet size produced by spinning disk atomizers," J. Fluids Eng. Trans. ASME, Vol. 134, No. 7, 2012.
9. Mairghany, M., Yahya, A., Adam, M., Suhaizi, A., Su, M., & Elsoragaby, S., "Evaluation of Pesticide Spraying Quality in Wetland Rice Cultivation in Malaysia", 2019.

Authors Introduction

Prof. Dr. Hazry Desa

He received his Bachelor of Mechanical Engineering from Tokushima University, Japan. He pursued his Ph.D. at the Artificial Life and Robotics Laboratory, Oita University, Japan in 2003. He is also the Director at Center of Excellence for Unmanned Aerial Systems (COEUAS), Universiti Malaysia Perlis, Malaysia.

Dr. Muhammad Azizi Azizan

He received his PhD degree in Civil Engineering from Universiti Malaysia Perlis and is currently working as a Senior Lecturer at the same institution. He is also co-founder and director of Haas Solutions Sdn Bhd, a technology startup company of Universiti Malaysia Perlis

Dr. Nik Noriman Zulkepli,

He received his PhD in polymer engineering from Universiti Sains Malaysia in 2011. He is currently working as a Senior Lecturer in the Polymer Engineering Division, School of Materials Engineering, Universiti Malaysia Perlis. He is also Head of Cluster (Materials) at the Centre of Excellence for Unmanned Aerial Systems.

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 Ecellence for Unmanned Aerial System (COEUAS), Universiti Malaysia Perlis.

Ms. Tan Yew Tian

She received her Bachelor's Degree in Mechatronic Engineering in 2022 from the Faculty of Electrical Engineering Technology from Universiti Malaysia Perlis (UniMAP), Malaysia.

Ms. Siti Syuhaidah Yahya

She received her Bachelor of Engineering (Honours) in Electronics Engineering from the Universiti Malaysia Perlis (UniMAP), Malaysia.

Ms. Aisyah Arina Mohammad Shahrazel,

She received her Bachelor of Engineering (Honours) in Manufacturing Engineering from the Universiti Malaysia Perlis (UniMAP), Malaysia.

Mr. Fakhrul Mukmin Mansor

He received his Bachelor of Electronic Engineering Technology in Electronic Network Design with Honours from the Universiti Malaysia Perlis (UniMAP), Malaysia.

Mrs. Siti Zaleha Abdul Aziz

She received her Master's degree in Green & Energy Efficient Building in 2014, a joint award between UniKL, Malaysia and the University of Applied Sciences Rosenheim, Germany. She is currently a Vocational Training Officer at Mara Japan Industrial Institute (MJII), Beranang.

Mr. 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.

The Capabilities and Readiness of Unmanned Aerial System (UAS) implementation in Construction Work Progression

Nurfadzillah Ishak
Faculty of Civil Engineering & Technology, Universiti Malaysia Perlis, Perlis, Malaysia
Centre of Excellence for Unmanned Aerial System, Universiti Malaysia Perlis, Perlis, Malaysia

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

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: nurfadzillah@unimap.edu.my

Abstract

Nowadays, the implementation of Unmanned Aerial Systems technology is becoming increasingly widespread in the construction industry. Unmanned Aerial Systems (UAS) have evolved over the past decade as both advanced military technology and off-the-shelf consumer devices. Nowadays, the implementation of Unmanned Aerial Systems technology is becoming increasingly widespread in construction industry. The uses of drones are increasingly moving towards successful remote procedures which can take place in a range of in building industry. Unmanned Aerial Systems remote sensing equipment is a way in which current building construction work progress can be analysed and inspected. Therefore, this study would be beneficial as eye opener for building construction practitioner that this remote sensing equipment can replaced the manual paper-based supervision and unsystematic database of current conventional construction work progression process.

Keywords: Capabilities, Readiness; Unmanned Aerial System; Construction Work Progression

1. Introduction

Traditionally, monitoring construction progress is a labour-intensive task. Daily or weekly reports are mostly used to document the quantity of work done on site. Although these reports are flexible and offer many kinds of information, including resource availability, possible hazards, inventory checklists, and incidents, they fail to collect geometric three-dimensional (3D) data from the finished job [1]. Progress information collected using existing processes for manual progress measurement is prone to mistakes and takes time [2]. In addition, approximately 30 to 50 percent of the time a project manager spends on data processing and analysis. Accurate and fast data collection, analysis, and visualisation of a project's as-built state are important components of effective project progress monitoring. This increased burden leads to further mistakes, as project teams hurry to synthesise data and make timely effective project control choices. Accurate and frequent progress reporting on building projects allows the project team to know their project status and make educated choices. [3] stated that the progress reporting of the construction project is according to an important component of the management of the construction projects. From the report information, the project team can understand how the project progresses, budgets are respected, the necessary quality is reached and safety precautions are followed via a frequent and accurate status report. This information leads them via educated choices and remedial action. There is a positive connection between performance and a person is exposed to the quantity of information. The performance of an individual "declines quickly" if too much information is given to the person. Moreover, conventional progress reports offer a snapshot of progress over a period (bi-weekly, weekly, or monthly), but management must understand the progress trend of the project up to assessment date in order to make informed choices. Current methods to monitor progress are not as precise, consistent, dependable or fast to allow effective choices for controls to maintain a project on schedule. Manual non-spatial data gathering procedures, such as checklists and daily reports with a huge number of visual data from the construction site may be collected in a short period, are the methods most often employed. Substantial delays may occur before foremen are analysed on a daily or weekly basis and any relevant information is communicated to the parties involved. The entire procedure may avoid timely corrective measures and lead to delayed schedules.

2. Unmanned Aerial System (UAS) technology capabilities

Additionally, the Architecture, Engineering, and Construction (AEC) business has gradually embraced Unmanned Aerial System (UAS) technology. UASs have been used in a wide variety of AEC applications, ranging from traffic surveillance [4]; and landslide monitoring [5] to cultural heritage conservation [6]; [7] and city planning [8]. Construction, in particular, has seen an exponential increase in the use of UAS, owing to the advantages connected with their capacity to enter inaccessible or dangerous regions and their ability to accomplish jobs safely and effectively [9]. Unmanned Aerial System (UAS), Unmanned Aerial Vehicles (UAV), Unmanned Aircraft System (UAS), Aircraft Operated Remotely (ROA), and Remote Pilot (RPA) are among other terms used today for various purposes, the whole world recognizes it as a drone. Previously, Unmanned Aerial System or drone technology originally mainly developed for military purposes. Unmanned Aerial System as powered aerial vehicles not carrying a human operator, so the aerial vehicles can fly autonomously or be remotely piloted. This system presents opportunities for an Unmanned Aerial System effective with remote procedures. With that Unmanned Aerial Systems as both advanced aircraft hardware and off-shelf consumer devices developed in the last decade [10].

In recent years unmanned aircraft have become an important research area. Today, because of their mechanical simplicity, they are more and more recruited for civil applications with regard to surveillance and infrastructure inspection. Aerial vehicles are often differentiated by their capacity to fly at different speeds, to maintain their position, to hover over a destination and to carry out manoeuvres near to obstacles and, whether fixed or loitering, to fly inside and outdoors. These characteristics make them ideal for human replacement when human involvement is hazardous, difficult, costly or exhausting [11]. Previous research has examined the use of Unmanned Aerial Systems for building systems inspection, focusing on the deployment and workflow of drones, 3D model reconstruction, and anomaly detection image data analytics.

The use of the Unmanned Aerial System has increasingly grown and has now proven to be regarded as a standard analysis method for the collection of photographs and other information on demand in an area of interest. [12] discovered that drones are among the current technologies that are considered suitable to replace the monitoring of progress done manually which describes the drone technology as capable of collecting data on the construction site in a short time and monitor on site build large-scale in real time more often. The combination of the technology will continue to permeate a variety of fields because applications, innovations and the capabilities are discovered and improved according to the current situation.

3. Technology Readiness Index (TRI)

Technology readiness classifies individuals' willingness to embrace and use emerging technology to accomplish goals in home life and at work. Development can be interpreted as an overall state of mind arising from a combination of mental enablers and inhibitors that collectively decide the predisposition of a person to use new technologies. The scale of items to be easier to incorporate as part of analysis questionnaires. The initial scale also consists of four measurements that are optimism, innovation, discomfort, and insecurity [13]. With the advancements in technology, operating conditions have improved tremendously. Additionally, the intellectual challenge, recognition, increased knowledge base, growing world population, and rising economies gain are just a few drivers of technological advancement. Additionally, technology will change over time and the rapid growth of new technology has led to an updated and simplified Technology Readiness Index (TRI) being created. Hence, the method developed by Parasuraman consists of 4 which is optimism, innovation, discomfort and security. TRI is a measure that has been updated to suit the changes in the technology environment, while at the same time streamlining the measure. With that, the scale of items to be easier to incorporate as part of analysis questionnaires. The initial scale also consists of four measurements that are optimism, innovation, discomfort, and insecurity [14]. Optimism element is a positive view of technology and a belief that it offers improved power, flexibility and efficiency in people's lives. It generally captures the optimistic feelings of technology. While the innovativeness element is classified as a propensity to be a pioneer in technology and a leader in thought. In general, this dimension tests to the degree individuals consider themselves to be at the front of adoption of technology. The discomfort element is like a perceived loss of technology power and a sense of being confused by it. This aspect usually tests fear, and when faced with technology, concerns people's experience. Hence, the insecurity element is a lack of technology and uncertainty about its ability to operate properly. This aspect focuses on concerns that people might have in the face of transactions based on technology.

Technically, the use of construction automation and robotics follows the same route, with particular focus on assembling and installing components using these technologies. Any country's readiness to implement construction automation and robotics technologies can be evaluated in terms of costs or financial obligations, technological expertise and equipment such as the availability of components and machinery, compatibility with existing practices and current construction activity, labour situation, construction characteristics and market culture.

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 safety management and three practitioners or project managers who have experience in managing construction 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 74 construction organizations from total G7 contractor-grade in Pulau Pinang. Pulau Pinang has a high number of constructions and is considered as a positively developing state in Malaysia. A total of 66 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 capabilities and readiness of Unmanned Aerial System implementation in construction work progression

Table 1 shows the means and standard deviation based on each variable which is for capabilities factor of Unmanned Aerial System technology. The variables consist of Control System (mean=4.60, sd.=0.34), Data Collection (mean=4.47, sd.=0.44), Safety Element (mean=4.43, sd.=0.44), Building Factor (mean=4.41, sd.=0.47) and Visual Element (mean=4.38, sd.=0.52). Furthermore, for the readiness elements of implementation Unmanned Aerial System technology are Optimism (mean=3.56, sd.=0.44), Innovative (mean=3.24, sd.=1.38), Discomfort (mean=3.06, sd.=0.95) and Insecurity (mean=3.12, sd.=1.37). The mean value score interpretation of the variables used in this study were be interpreted based on the suggestion by [15].

Table 1. Descriptive Analysis

Variables	Mean	Std. Deviation	Mean Value Score Interpretation
Capabilities Factors			
Control System	4.60	0.34	High
Data Collection	4.47	0.44	High
Safety Element	4.43	0.44	High
Building Factor	4.41	0.47	High
Visual Element	4.38	0.52	High
Readiness Elements			
Optimism	3.56	0.44	High
Innovative	3.24	1.38	Moderate
Discomfort	3.06	0.95	Moderate
Insecurity	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 1, it can be concluded that technology advances in recent years have resulted in Unmanned Aerial System technology more capable and can offer a wide range of uses on construction sites. In addition, the usage of Unmanned Aerial Systems on workplaces also means savings in time and expenses while doing activities and improving safety compared to other more conventional methods. However, there are complex factors that should be stressed that may inhibit the readiness level of the Unmanned Aerial System implementation that cause the moderate relationship findings. However, it may be influenced by Unmanned Aerial Vehicles having general budget difficulties. They are extremely costly including regulation, maintenance, insurance, training, image processing software and software navigation related to Unmanned Aerial Vehicles operations. Therefore, the discovery of potential complex factors should need further input in order to get a solid substantial value in order to develop an effective strategy to improve the readiness level of Unmanned Aerial System implementation in construction work progression.

6. Conclusions

As summarised, the building sector is critical to the nation's economic progress. Construction projects are time-sensitive; any delays may cost a significant amount of money and must be avoided. To do this, building progress should be regularly monitored to guarantee that deadlines and objectives are reached on time. A great deal of work goes into developing a system for tracking progress, without which the project may become unorganised. The Unmanned Aerial System is regarded as an efficient assessment and reward system for quality improvement. The applicability of using Unmanned Aircraft System (UAS) can be an alternative to the conventional construction work progress at site which is in line with the era of digitization construction of the fourth industrial revolution (IR 4.0).

References

1. Puri, N., & Turkan, Y. Bridge construction progress monitoring using lidar and 4D design models. Automation in Construction, 2020, 109, 102961.
2. Vick, S. M., & Brilakis, I. A Review of Linear Transportation Construction Progress Monitoring Techniques. 16th International Conference on Computing in Civil and Building Engineering, ICCCBE 2016, July, 1106–1113.
3. Lamptey, W. N. L., & Fayek, A. R. Developing a Project Status Dashboard for Construction Project Progress Reporting. International Journal of Architecture, Engineering and Construction, 2012, 1(2), 112–120.
4. Albeaino, G., Gheisari, M., & Franz, B. W. (2019). A systematic review of unmanned aerial vehicle application areas and technologies in the AEC domain. Journal of Information Technology in Construction, 24(June), 381–405.
5. Barmpounakis, E., & Geroliminis, N. On the new era of urban traffic monitoring with massive drone data: The pNEUMA large-scale field experiment. Transportation Research Part C: Emerging Technologies, 2020, 111, 50–71.
6. Yeh, T. W., & Chuang, R. Y. Morphological Analysis of Landslides in Extreme Topography by UAS-SfM: Data Acquisition, 3d Models and Change Detection. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 2020, 43(B5), 173–178.
7. Enríquez, C., Jurado, J. M., Bailey, A., Callén, D., Collado, M. J., Espina, G., Marroquín, P., Oliva, E., Osla, E., Ramos, M. I., Sarceño, S., & Feito, F. R. The UAS- based 3D image characterization of Mozarabic Church Ruins in Bobastro (Malaga), Spain. Remote Sensing, 2020, 12(15), 1–16.
8. Uysal, M., Toprak, A. S., & Polat, N. Photo Realistic 3D Modeling With Uav:Gedik Ahmet Pasha Mosque inAfyonkarahisar. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL- 5/W2, 2013, 659–662.
9. Banaszek, A., Zarnowski, A., Cellmer, A., & Banaszek, S. Application of new technology data acquisition using aerial (UAV) digital images for the needs of urban revitalization. 10th International Conference on Environmental Engineering, ICEE 2017, September.
10. Gheisari, M., & Esmaeili, B. Applications and requirements of unmanned aerial systems (UASs) for construction safety. Safety Science, 2019, 118, 230–240.
11. Zhou, S., & Gheisari, M. Unmanned aerial system applications in construction: a systematic review. Construction Innovation, 2018, 18(4), 453–468.
12. Rakha, Tarek & Gorodetsky, Alice. Review of Unmanned Aerial System (UAS) applications in the built environment: Towards automated building inspection procedures using drones. Automation in Construction, 2018.
13. Bang, S., Kim, H., & Kim, H. UAV-based automatic generation of high- resolution panorama at a construction site with a focus on pre-processing for image stitching. Automation in Construction, 2017, 84, 70–80.
14. Darwin, N. Unmanned Aerial Vehicle Large Scale Mapping for Coastal Erosion Assessment. May, 2017.
15. Ishak, N. & Azizan, M.A.A. Sustainable Building Retrofits Potential Indicators. Proceedings of Green Design and Manufacture 2020. AIP Conf. Proc. 2339, 020154-1–020154-4.

Authors Introduction

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 Ecellence for Unmanned Aerial System (COEUAS), Universiti Malaysia Perlis.

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 Ecellence for Unmanned Aerial System (COEUAS), Universiti Malaysia Perlis.

Prof. Dr. Hazry Bin Desa

He graduated from Oita University in PhD in Materials Science and Production Engineering (Robotics) and currently the Professor in Centre of Excellence for Unmanned Aerial Systems (COEUAS), Universiti Malaysia Perlis (UniMAP).

Chaotic African Vultures Optimization Algorithm for Feature Selection

Wy-Liang Cheng[1], Li Pan[1], Mohd Rizon Bin Mohamed Juhari[1], Chin Hong Wong[2,3], Abhishek Sharma[4], Tiong Hoo Lim[5], 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*
[5]*Faculty of Engineering, Universiti Teknologi Brunei, Bandar Seri Begawan 1410, Brunei Darussalam*
E-mail: 1001437466@ucsiuniversity.edu.my, 1002060534@ucsiuniversity.edu.my, mohdrizon@ucsiuniversity.edu.my, chinhong.wong@mu.ie, abhishek15491@gmail.com, lim.tiong.hoo@utb.edu.bn, tiangss@ucsiuniversity.edu.my limwh@ucsiuniversity.edu.my

Abstract

Feature selection is a widely used technique to remove the undesirable, noisy and inaccurate information from raw input dataset while maintaining the accuracy and efficiency of classifier. Tremendous researches have explored the feasibility of metaheuristic search algorithms (MSAs) such as African Vultures Optimization Algorithm (AVOA) to solve feature selection problem. Similar with many original MSAs, the conventional initialization scheme of AVOA has undesirable drawbacks that can lead to entrapment of local optima, especially when dealing with complex dataset. In this paper, a new variant known as Chaotic African Vultures Optimization Algorithm (CAVOA) is proposed to solve feature selection problem with enhanced classification accuracy by incorporating the chaotic map concept into the initialization scheme. Twelve datasets obtained from UCI Machine Learning Repository are used to investigate the capability of CAVOA in feature selection and compared with four other peer algorithms. Simulation results show that CAVOA can produce the best classification accuracies and lowest feature numbers in most datasets.

Keywords: African Vultures Optimization Algorithm, feature selection, metaheuristic search algorithm

1. Introduction

Enormous growth of dataset is observed in the era of Big Data due to the generation of datasets from various domains of sources. Complexity of datasets has increased significantly due to the presence of unwanted and noisy data, leading to "curse of dimensionality" issue that can compromise the effectiveness of classifier. It remains a challenging issue to eliminate unwanted information from raw input datasets while maintaining the accuracy and computation time of classifier.

Feature selection[1] is a popular technique used to eliminate redundant information from raw datasets when solving real-world problems such as fault detection[2], [3] and automatic modulation recognition[4],[5] Existing feature selection can be classified as the filter and wrapper approaches. Filter approach is implemented to identify the potential useful feature subset based on predefined data or metric content. Meanwhile, wrapper approach is designed by incorporating a selected classifier to evaluate the quality of feature subset. Despite incurring longer computation time, the wrapper methods tend to produce better accuracy than those of filter approaches when solving feature selection problems.

Metaheuristic search algorithm (MSA) has emerged as a promising wrapper approach used to solve feature selection problems due to its good global search ability. Depending on the sources of inspirations, existing MSAs can be categorized into: (a) Darwin's theory of evolution, (b) swarm intelligence, (c) physics-based and (d) human-based[6]. African Vultures Optimization Algorithm (AVOA) is a new MSA proposed in 2021[7] motivated by the behavior of vulture in hunting for food sources. Similar with other MSAs, the initial population of AVOA is randomly produced by discarding the information of search environment[8]. This drawback can restrict the ability of AVOA to solve tackle real-world problem such as feature selection due to premature convergence issue.

In this paper, a new variant of Chaotic African Vultures Optimization Algorithm (CAVOA) is designed as a wrapper-based method to solve the feature selection problems competitively. A chaotic initialization method is introduced to produce the initial population with better solution quality that can lead to the better optimization results in terms of feature subset selected. The capability of CAVOA to solve feature selection problems is evaluated with twelve datasets selected from the UCI Machine Learning Repository[9].

2. Related Works

2.1. *Inspiration of AVOA*

Vulture is a hunting bird that searches for the animal carcasses as a food source. Popular physical feature of a vulture is its bald head to avoid any infection when consuming the carcasses. According to some researches, the bare skin of vulture is used to maintain its body heat. Different species of vultures have unique characteristics in terms of flight pattern and strength to fend off another animal. When hunting for food, vultures fly around to search for other species of vultures with food supply. This might cause other vultures to move to the same place and fight with each other for food sources. The weaker vultures tend to move around the stronger vultures and steal their food by tiring the stronger vultures, resulting in some vultures becoming more aggressive.

2.2. *Feature Selection*

Feature selection can be formulated as a bi-objective optimization problem by considering the classification accuracy and number of selected features. The fitness function used to measure the quality of each AVOA solution X when solving the feature selection problem is:

$$f(X) = q\gamma + n\frac{|F_s|}{|F_o|} \tag{1}$$

where $q \in [0,1]$ and $n \in [1-q]$ are parameters used to represent the weightage of classification error γ and the length of selected feature subset $|F_s|$. Meanwhile, $|F_O|$ is the lengths of original input datasets. Smaller $f(X)$ value is more desirable for feature selection problem.

3. Proposed CAVOA

The search mechanisms of proposed CAVOA to solve feature selection problems can be described in four stages. Stages 1 is the modified initialization scheme with a chaotic map to produce initial population of CAVOA. Stage 2 determines the satiation rate of each CAVOA solution. Stages 3 and 4 focus on balancing exploration and exploitative states of proposed CAVOA.

In stage 1, the initial value of a chaotic variable is randomly produced as $\alpha_0 \in [0,1]$. A circle map is then used to update the value of chaotic variable at every k-th iteration, denoted as α_k, where $k = 1, \dots, k_{max}$ and k_{max} is the maximum iteration. The circle map used to update α_{k+1} in the next $(k+1)$-th iteration is formulated as:

$$\alpha_{k+1} = mod\left(\alpha_k + c - \left(\frac{b}{2\pi}\right)\sin(2\pi\alpha_k), 1\right) \tag{2}$$

where $b = 0.5$ and $c = 0.2$ are two constant values. After obtaining the chaotic result $\alpha_{k(end)}$ at the final iteration of k_{max}, the initial position of each i-th CAVOA solution in the d-th dimension is obtained as:

$$X_{i,d} = X_d^{LB} + \alpha_{k(end)}(X_d^{UB} - X_d^{LB}) \tag{3}$$

where $X_{i,d}^{LB}$ and $X_{i,d}^{UB}$ are the lower and upper boundaries of d-th decision variable, respectively, and $d = 1, .., D$. The initial population of CAVOA can be generated using Eq. (3) and it is sorted in ascending order based on fitness values to obtain the first- and second-best results.

During Stage 2, the satiation rate used to mimic the status of vultures (i.e., hungry or fulfilled) are calculated to determine the exploration and exploitation state of CAVOA as shown below:

$$m = (2 \times r_1 + 1) \times \aleph \times \left(1 - \frac{\mu}{\mu_{max}}\right) + Q \tag{4}$$

where r_1 is a random number between 0 and 1; $\aleph$ is a random number between -1 and 1; μ and μ_{max} represents the current fitness evaluation number and maximum fitness evaluation number, respectively. Q is a variable utilized to enhance the performance of CAVOA when dealing with complex optimization problems, i.e.,

$$Q = r_2 \times \left(sin^f\left(\frac{\pi\mu}{2\mu_{max}}\right) + cos\left(\frac{\pi\mu}{2\mu_{max}}\right) - 1\right) \tag{5}$$

where r_2 indicates a random number with value between -2 and 2; f is a predefined parameter used to represents the optimization operation disrupts the exploration and exploitation state. At the end of Stage 2, a roulette wheel approach is used to randomly select one CAVOA solution from the best and second-best performing vultures. When the values of parameter $|m|$ in Stage 2 is bigger or equal to one (i.e., $|m| \geq 1$), the searching process of CAVOA will move forward to Stage 3 that aims to promote the exploration search of algorithm.

During the Stage 3 of CAVOA, all vultures are guided to perform searching in different random locations by using two possible search methods. In particular, the search method performed by every i-th CAVOA solution to update its position is randomly selected based on a probability denoted as P_1. If P_1 is less than generated random number, Eq. (6) is used to calculate the updated position of i-th CAVOA solution as X_i^{new}. Otherwise, Eq. (7) is used instead:

$$X_i^{new} = X_{row} - |(2 \times r_3) \times X_{row} - X_i| \times m \tag{6}$$

$$X_i^{new} = X_{row} - m + r_3 \times \left((X^{UB} - X^{LB}) \times r_4 + X^{LB}\right) \tag{7}$$

where m is the vulture satiation rate obtained from Eq. (4); X^{LB} and X^{UB} represent the lower and upper boundary of decision variables; X_{row} refers to the CAVOA solution randomly selected from the best or second-best vultures using the roulette wheel approach; r_3 and r_4 are two random numbers between 0 and 1.

Stage 4 of CAVOA is triggered when the $|m|$ value obtained in Stage 2 is smaller than one (i.e.,$|m| < 1$). Note that this stage is more exploitative and it has two subphases that can be triggered based on the value of $|m|$. The searching procedure of first subphase starts when $|m|$ is found larger than or equal to 0.5 (i.e., $|m| \geq 0.5$). Referring to Eq. (8), a search method is chosen based on the probability $P_2 \in [0,1]$ to calculate the new position of every i-th CAVOA solution X_i^{new} as follow:

$$X_i^{new} = \begin{cases} (C + D)/2 & , P_2 > rand \\ X_{row} - |X_{row} - X_i| \times m \times L & , \text{otherwise} \end{cases} \tag{8}$$

where C and D represent the movements of the best vulture X_1^{best} and second-best vulture X_2^{best} when they are competing for the food sources, i.e.,

$$C = X_1^{best} - \frac{X_1^{best} \times X_i}{X_1^{best} \times (X_i)^2} \times m \tag{9}$$

$$D = X_2^{best} - \frac{X_2^{best} \times X_i}{X_2^{best} \times (X_i)^2} \times m \tag{10}$$

It is also notable that L presented in Eq. (8) refers to the levy flight characteristic used to further enhance the search potential of CAVOA as follow:

$$L = \frac{0.01}{|Y|^{\frac{1}{\omega}}} \left(T \left(\frac{\delta(1+\omega) \times sin\left(\frac{\pi\omega}{2}\right)}{\delta(1+2\omega) \times \omega \times 2\left(\frac{\omega - 1}{2}\right)} \right)^{\frac{1}{\omega}} \right) \tag{11}$$

where T and Y are two random numbers between 0 and 1; ω is predefined constant and its value is set as 1.5; $\delta(\cdot)$ refers to the gamma function.

Meanwhile, the searching mechanism for the second subphase of Stage 4 is triggered when $|m|$ is lesser than 0.5 (i.e., $|m| < 0.5$). The search strategy method in the second subphase is also selected based on a predefined probability $P_3 \in [0,1]$. If P_3 is larger than the randomly generated value (*rand*), Eq. (12) is used to calculate the updated position X_i^{new} of every i-th CAVOA solution. Otherwise, Eq. (13) is applied as shown below:

$$X_i^{new} = |(2 \times r_5) \times X_{row} - X_i| \times (m + r_6) - X_{row} - X_i \tag{12}$$

$$X_i^{new} = X_i - (S_1 - S_2) \tag{13}$$

where r_5 and r_6 are two random numbers with value between 0 and 1; X_{row} represent a randomly chosen vulture using roulette wheel selected; S_1 and S_2 are calculated using Eqs. (14) and (15), respectively, where r_7 and r_8 are two random numbers between 0 and 1.

$$S_1 = X_i \times \left(\frac{r_7 \times X_i}{2\pi}\right) \times cos(X_i) \tag{14}$$

$$S_2 = X_i \times \left(\frac{r_8 \times X_i}{2\pi}\right) \times sin(X_i) \tag{15}$$

The overall flow for CAVOA is summarized in Fig 1. The search process continues until stopping condition is achieved, i.e., when the current fitness evaluation number exceeds the maximum fitness evaluation number ($\mu > \mu_{max}$).

```
Algorithm 1: CAVOA based Feature Selection
Inputs: $i$, $k$, $\mu$, $\mu_{max}$, $X^{UB}$, $X^{LB}$
Load parameters and selected dataset;
for each $i$-th vulture do
    for each $d$-th dimension do
        Generate chaotic variable using Eq. (2);
        Initialize the position using Eq. (3);
    end for
end for
Evaluate vulture's fitness using Eq. (1)
Select the first and second-best vultures as $X_1^{best}$ and $X_2^{best}$, respectively;
while $\mu \leq \mu_{max}$ do
    for each $i$-th vulture do
        Calculate $m$ using Eq. (4) and (5);
        Perform the roulette wheel selection on $X_1^{best}$ and $X_2^{best}$ to select $X_{row}$;
        if $|m| \geq 1$ then
            Calculate $X_i^{new}$ using Eq. (6) and (7);
        else if $|m| < 1$ then
            if $|m| \geq 0.5$ then
                Calculate $X_i^{new}$ using Eq. (8);
            else if $|m| < 0.5$ then
                Calculate $X_i^{new}$ using Eq. (12);
            end if
        end if
        Boundary check of $X_i^{new}$;
        Fitness evaluation of $X_i^{new}$ using Eq. (1);
        Update the $X_1^{best}$ and $X_2^{best}$;
        $\mu \leftarrow \mu + 1$;
    end for
end while
Output: $X_1^{best}$ (i.e., optimal feature subset)
```

Fig 1. Overall flow of CAVOA for feature selection.

4. Performance Evaluation on CAVOA

4.1. *Simulation Settings*

The capability of proposed CAVOA to solve various feature selection problems is investigated using twelve datasets obtained from UCI Machine Learning Repository[9], i.e., (a) Breast Cancer Wisconsin (Original), (b) Dermatology, (c) Lymphography, (d) Iris, (e) Arrhythmia, (f) Echocardiogram, (g) Waveform Database Generator (Version 1), (h) Zoo, (i) Indian Liver Patient Dataset (ILPD), (j) Semeion Handwritten Digit, (k) Letter Recognition and (l) Balance Scale. The performance of the proposed CAVOA to solve these selected datasets are also compared with another four peer feature selection algorithms developed using the state-of-art MSAs, known as the African Vultures Optimization Algorithm (AVOA)[7], Bezier Search Differential Evolution Algorithm (BeSD)[10], Generalized Normal Distribution Optimization (GNDO)[11] and Sperm Swarm Optimization (SSO)[12]. The performances of all compared feature selection algorithms in solving all selected datasets are evaluated using two metrics, i.e., the mean classification accuracy Acc^{mean} and average number of selected features $N^{feature}$. For fair comparison, the population sizes and maximum fitness evaluation numbers of all algorithms are set as $N = 20$ and $\tau^{max} = 2000$, respectively. Each algorithm is used simulated for 30 independent runs to solve each dataset during performance comparative studies.

4.2. *Performance Comparisons*

An algorithm that can produce higher Acc^{mean} and lower $N^{feature}$ values simultaneously is considered to be better when solving feature selection problem. The value with boldface represents the best result among the compared result while italic and underlined valued indicate the second-best result. Table 1 reports that the CAVOA has the best feature selection performance by solving 11 out 12 datasets with best Acc^{mean} results. This is followed by DSSA and BeSD with two best Acc^{mean}, AVOA with one best Acc^{mean}, GNDO with no best Acc^{mean}. Table 2 result reports that the proposed CAVOA has the best performance in selecting the least number of features when solving four datasets, followed by AVOA, BeSD and SSO that solve lesser datasets with the least numbers of features.

Table 1. Comparison of Acc_{mean}.

Dataset	AVOA	BeSD	GNDO	SSO	CAVOA
(a)	0.987	0.988	*0.991*	0.963	**1.000**
(b)	0.990	0.981	0.985	*0.993*	**1.000**
(c)	0.590	0.544	*0.617*	0.535	**0.655**
(d)	**1.000**	**1.000**	0.967	*0.994*	0.967
(e)	0.710	0.687	0.673	*0.711*	**0.759**
(f)	**1.000**	**1.000**	**1.000**	*0.995*	**1.000**
(g)	*0.831*	0.815	0.830	0.813	**0.838**
(h)	*0.927*	0.888	0.890	0.773	**0.963**
(i)	*0.760*	0.726	0.724	0.739	**0.793**
(j)	*0.932*	0.886	0.928	0.882	**0.949**
(k)	0.953	0.946	*0.954*	0.923	**0.955**

(l)	*0.840*	0.808	0.712	0.822	**0.888**

Table 2. Comparison of $N_{feature}$.

Dataset	AVOA	BeSD	GNDO	SSO	CAVOA
(a)	4.07	5.20	*3.83*	4.97	**3.30**
(b)	**11.43**	16.57	13.97	16.87	*11.87*
(c)	**3.37**	9.40	7.23	6.23	*3.90*
(d)	2.00	2.40	**1.00**	*1.97*	**1.00**
(e)	*38.13*	134.73	126.97	**17.90**	47.87
(f)	**1.00**	4.37	**1.00**	*3.90*	**1.00**
(g)	16.97	**9.67**	14.23	*10.47*	16.90
(h)	*5.23*	7.63	6.80	6.57	**3.77**
(i)	2.90	4.63	*2.87*	**1.73**	3.37
(j)	144.70	*129.10*	129.37	**127.83**	185.03
(k)	12.17	**7.87**	12.30	*10.53*	12.20
(l)	4.00	**2.17**	4.00	*3.83*	4.00

5. Conclusions

A modified algorithm known as CAVOA is presented in this paper to solve feature selection problems more efficiently. In particular, the ergodicity and non-repetitive characteristics of circle chaotic map are leveraged during the initialization stage to produce the initial population of CAVOA with enhanced solution quality. This is followed by the search mechanisms used to achieve the better balancing of exploration and exploitation states of proposed. A total of 12 datasets are selected to evaluate the performance of CAVOA and its peer algorithm to solve feature selection problems. The simulation results show that CAVOA can outperform another four feature selected algorithms developed with state-of-art MSAs by solving more selected benchmark problems with the highest classification accuracy and the least number of selected features.

Acknowledgements

This research is supported by Ministry of Higher Education Malaysia (MOHE) under the Fundamental Research Grant Scheme with the project codes of FRGS/1/2019/TK04/UCSI/02/1 and FRGS/1/2020/TK0/UCSI/02/4. This research is also supported by UCSI Research Excellence & Innovation Grant (REIG) with project code of REIG-FETBE-2022/038.

References

1. Ang, K. M. et al. New Hybridization Algorithm of Differential Evolution and Particle Swarm Optimization for Efficient Feature Selection. (2022)
2. Alrifaey, M., Lim, W. H. & Ang, C. K. A Novel Deep Learning Framework Based RNN-SAE for Fault Detection of Electrical Gas Generator. IEEE Access 9, 21433-21442 (2021).
3. Alrifaey, M. et al. Hybrid Deep Learning Model for Fault Detection and Classification of Grid-Connected Photovoltaic System. IEEE Access 10, 13852-13869 (2022).
4. Jdid, B., Hassan, K., Dayoub, I., Lim, W. H. & Mokayef, M. Machine Learning Based Automatic Modulation Recognition for Wireless Communications: A Comprehensive Survey. IEEE Access 9, 57851-57873 (2021).
5. Jdid, B., Lim, W. H., Dayoub, I., Hassan, K. & Juhari, M. R. B. M. Robust Automatic Modulation Recognition Through Joint Contribution of Hand-Crafted and Contextual Features. IEEE Access 9, 104530-104546 (2021).
6. Ahmad, M. F., Isa, N. A. M., Lim, W. H. & Ang, K. M. Differential evolution: A recent review based on state-of-the-art works. Alexandria Engineering Journal 61, 3831-3872 (2022).
7. Abdollahzadeh, B., Gharehchopogh, F. S. & Mirjalili, S. African vultures optimization algorithm: A new nature-inspired metaheuristic algorithm for global optimization problems. Computers & Industrial Engineering 158, 107408 (2021).
8. Ahmad, M. F., Isa, N. A. M., Lim, W. H. & Ang, K. M. Differential evolution with modified initialization scheme using chaotic oppositional based learning strategy. Alexandria Engineering Journal 61, 11835-11858 (2022).
9. Dua, D. & Graff, C. UCI machine learning repository. (2017).
10. Civicioglu, P. & Besdok, E. Bezier Search Differential Evolution Algorithm for numerical function optimization: A comparative study with CRMLSP, MVO, WA, SHADE and LSHADE. Expert Systems with Applications 165, 113875 (2021).
11. Zhang, Y., Jin, Z. & Mirjalili, S. Generalized normal distribution optimization and its applications in parameter extraction of photovoltaic models. Energy Conversion and Management 224, 113301 (2020).
12. Shehadeh, H., Ahmedy, I. & Idris, M. Sperm Swarm Optimization Algorithm for Optimizing Wireless Sensor Network Challenges. (2018).

Authors Introduction

Mr. Wy – Liang Cheng

He received the B.Eng. degree in Mechatronic Engineering with Hnours from UCSI University, Malaysia in 2020. He is currently pursuing Master of Philosophy in Engineering in UCSI University, Malaysia. His research interests are swarm intelligence and feature selection.

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.

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.

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. 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.

Multi Chaotic Flow Direction Algorithm for Feature Selection

Wy-Liang Cheng[1], Li Pan[1], Mohd Rizon Bin Mohamed Juhari[1], Abhishek Sharma[2], Hameedur Rahman[3], Chun Kit Ang[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 Computing and Artificial Intelligence, Air University, Islamabad Capital Territory 44000, Pakistan*
E-mail: 1001436889@ucsiuniversity.edu.my, 1002060534@ucsiuniversity.edu.my, mohdrizon@ucisuniversity.edu.my, abhishek15491@gmail.com, rhameedur@mail.au.edu.pk, angck@ucsiuniversity.edu.my, tiangss@ucsiuniversity.edu.my, limwh@ucsiuniversity.edu.my,

Abstract

Feature selection is a crucial pre-processing step used to remove redundant information from original datasets while preserving the accuracy and processing time of classifier. The feasibility of using metaheuristic search algorithms (MSAs) such as Flow Directional Algorithm (FDA) to solve feature selection problems is one of the active research topics. Similar with other MSAs, FDA also employs conventional initialization scheme that generates initial solutions in random basis. The absence of intelligent mechanisms in conventional initialize scheme tends to generate initial populations in local optima, hence compromising the performance of algorithm to handle datasets with complex features. In this paper, a modified algorithm known as Multi Chaotic Flow Directional Algorithm (MCFDA) is proposed to solve feature selection problems with enhanced performances by leveraging the strengths of multiple chaotic maps for population initialization. A total of 12 datasets from UCI Machine Learning Repository are selected for performance evaluation of MCFDA and another four peer algorithms to solve feature selection problems. The proposed MCFFA is revealed to deliver best performances by solving 7 out of 12 datasets with the best mean classification accuracy and 6 out of 12 datasets with the least numbers of selected features.

Keywords: chaotic map, feature selection, Flow Directional Algorithm (FDA), metaheuristic search algorithm

1. Introduction

Due to the emerging of big data era, data-driven methods such as machine learning and deep learning have become more popular to solve various real-world problems. The performances of these data-driven methods depend on the quality of datasets used for model training. In the real-world scenario, redundant and noisy information tend to be observed from the input datasets and their presence tends to degrade the performance of data-driven methods. Feature selection[1] is a popular pre-processing method used to remove the undesirable information from original datasets without compromising the accuracy and speed of data-driven methods. Some popular applications of feature selection are fault detection[2],[3] and automatic modulation recognition[4],[5]. Feature selection is considered as a challenging NP-type optimization problem because it involves the optimal selection of feature subset from the original input datasets.

Existing feature selection can be divided into two main approaches, i.e., filter- and wrapper-based methods. For filter-based methods, the potential useful feature subsets are determined based on statistical characteristics of datasets. For wrapper-based methods, a data-driven classifier is incorporated along to carefully evaluate the quality of potential feature subsets. Both approaches have their pros and cons. Filter-based methods are more computationally efficient with lower accuracy. Opposite behaviors are found from wrapper-based methods.

Motivated by their promising characteristics such as strong global search and easy implementation, many metaheuristic search algorithm (MSAs) inspired by various natural phenomena[6] (e.g., theory of evolution, animal forging behaviors, physics-based phenomena and human-based activities) are used to handle numerous complex optimization problems[7],[8],[9],[10],[11] including the feature selection. Flow Directional Algorithm (FDA) is a physic-based algorithm proposed in 2021 by emulating the flow direction within a drainage basin. Similar with other MSAs, FDA generates its initial population in random basis without considering any intelligent mechanisms that can prevent the generation of initial solution in local optima[12]. This undesirable behavior can compromise the performance of FDA due to premature convergence.

A modified algorithm known as Multi-Chaotic Flow Directional Algorithm (MCFDA) is designed to solve feature selection problems competitively. Multiple chaotic maps with the ergodicity and non-repetitive characteristics are incorporated into initialization phase of MCFDA, aiming to generate initial solutions in the more promising solution regions. The performance of MCFDA to solve feature selection problems is evaluated using 12 datasets obtained from UCI Machine Learning Repository and compared with 4 peer algorithms.

2. Related Works

2.1. *Inspiration of FDA*

Excessive rainfall happens when the rainfall pours over the surface of ground and does not seep into the soil. Direct runoff refers to the additional water that is not absorbed after rainfall and losses. A method known as μ–index approach[13] is proposed based on this inspiration.

An index μ refers to the average number of water loss during precipitation with a unit of centimeter per hour, where this left-over water will change into runoff. The direct runoff value can be obtained by the difference of index μ at each interval of rainfall. The direct runoff w_d can be calculated as:

$$w_d = \sum_{g=1}^{G} (W_g - \mu\Delta t) \tag{1}$$

where Δt and G refer as the length of each time interval and total number of time steps, respectively.

The movement of runoff towards the outlet of basin is influenced by the angle of slope and can be emulated by splitting the drainage basin into several set of cells. The slope height and angle in a surrounding cell can also affect the amount of runoff in each cell moving to other cells. D8 method[14] can be used to verify the variation of runoff direction. Accordingly, each cell consists of eight neighbors with unique height and distance. Comparison between the distance and height of each cell with its adjacent cells is used to distinguish the flow direction. Referring to all cell slops obtained, the cell flows toward the adjacent cell with highest slope. The flow path is determined by using D8 for the basin. After the flow path is indicated, a variable corresponds to the number of cells that flow into that cell is defined. The greatest number is assigned to the outlet point of basin.

2.2. *Feature Selection Optimization Problem*

Feature selection is a bi-objective problem used to maximize the classification accuracy and minimize the numbers of selected features, simultaneously. The fitness function used to evaluate the quality of each solution X during feature selection process is represented as:

$$f(X) = kz + g\frac{|F_s|}{|F_o|} \tag{2}$$

where $k \in [0, 1]$ and $g = 1 - k$ are the parameters used to represent the weightage of classification error z and the length of selected feature subset $|F_s|$; $|F_O|$ refers to the length of original input dataset.

3. Proposed MCFDA

During the initialization phase, an initial chaotic variable $\sigma_0 \in [0,1]$ is randomly generated for each i-th MCFDA solution. At any k-th iteration, the chaotic variable σ_k can be updated with the chosen chaotic map, where $k = 1, \dots, k_{max}$ and k_{max} refers to the maximum iteration numbers. In this paper, four chaotic maps are used to update σ_k. If $\sigma_k < 0.25$, a circle map is selected. For $0.25 \le \sigma_k < 0.5$, a Gauss map is used. Singer map is chosen to update chaotic variable if $0.5 \le \sigma_k < 0.75$. Finally, Sinusoidal map is selected when $\sigma_k \ge 0.75$. The Circle, Gauss, Singer and Sinusoidal maps used to this study are presented in Eqs. (3), (4), (5) and (6), respectively, as follows:

$$\sigma_{k+1} = mod\left(\sigma_k + 0.2 - \left(\frac{0.5}{2\pi}\right)\sin(2\pi\sigma_k), 1\right) \tag{3}$$

$$\sigma_{k+1} = \begin{cases} 1 & , \sigma_k = 0 \\ \dfrac{1}{mod(\sigma_k, 1)} & , otherwise \end{cases} \tag{4}$$

$$\sigma_{\mathrm{k+1}} = 1.07(7086\sigma_k - 23.31\sigma_k^2 + 28.75\sigma_k^3 - 13.302875{\sigma_k}^4) \tag{5}$$

$$\sigma_{\mathrm{k+1}} = c\sigma_k^2 \sin(\pi\sigma_k), \quad \mathrm{c} = 2.3 \tag{6}$$

At the final iteration of k_{max}, the final chaotic variable $\sigma_{k(final)}$ is obtained and used to initialize the d-th dimension of i-th MCFDA solution as follow:

$$X_{i,d}^{flow} = X_d^{LB} + \sigma_{d(final)}(X_d^{UB} - X_d^{LB}) \tag{7}$$

where X_d^{LB} and X_d^{UB} represent as the minimum and maximum values of d-th decision variable. As compared to the conventional initialization scheme described in Eq. (8), the ergodicity and non-repetition characteristics of chaotic map can produce more thorough search in the solution space and prevent premature convergence. The initial population of MCFDA with the population size I and dimensional size D is generated using Eq. (7) for $i = 1,..,I$ and $d = 1,..,D$, before moving to next stage.

$$X_{i,d}^{flow} = X_d^{LB} + r_1 * (X_d^{UB} - X_d^{LB}) \tag{8}$$

During the iterative searching process, it is assumed that there are β neighborhoods around every flow, where the position is represented as shown:

$$X_j^{flow} = X_i^{flow} + r_2 * \Delta \tag{9}$$

where X_j^{flow} is refer the position of each j-th neighbour for every i-th MCFDA solution; X_i^{flow} is current position of i-th flow; r_2 is a random number between 0 and 1 generated by uniform distribution; Δ is a parameter used to control the behavior of MCFDA, i.e.,

$$\Delta = \left(r_3 X_i^r - r_4 X_i^{flow}\right) * \left\|X^{best} - X_i^{flow}\right\| * W \tag{10}$$

where r_3 and r_4 are the random numbers between 0 and 1 generated by uniform distribution; X_i^r is a randomly flow obtained in related with Eq. (1); X^{best} is the flow position with the best (i.e., lowest) fitness; W is a nonlinear weight with random number, i.e.,

$$W = \left(\left(1 - \frac{\gamma}{\gamma_{max}}\right)^{(2*r_5)}\right) * \left(\overline{r_6} * \frac{\gamma}{\gamma_{max}}\right) * \overline{r_7} \tag{11}$$

where r_5 is a random number between 0 and 1 generated by uniform distribution; $\overline{r_6}$ and $\overline{r_7}$ refer as a random number generated by normal distribution; γ and γ_{max} represent the current fitness evaluation number and maximum fitness evaluation number, respectively.

The velocity of flow V towards the neighbour with the lowest fitness and is related to the slope as shown:

$$V = r_8 * S_0 \tag{12}$$

where r_8 is a random number between 0 and 1 generated by uniform distribution; S_0 represents the slope vector between the i-th flow and its j-th neighthour, where

$$S_0 = \frac{f\left(X_i^{flow}\right) - f\left(X_j^{flow}\right)}{\left\|X_{i,d}^{flow} - X_{j,d}^{flow}\right\|} \tag{13}$$

where $f\left(X_i^{flow}\right)$ and $f\left(X_j^{flow}\right)$ represent the fitness value of i-th flow and its j-th neighbour with position vectors of X_i^{flow} and X_j^{flow}, respectively. The new position X_i^{newf} of each i-th flow is calculated as:

$$X_i^{newf} = X_i^{flow} + V * \frac{X_i^{flow} - X_j^{flow}}{\left\|X_i^{flow} - X_j^{flow}\right\|} \tag{14}$$

When a randomly selected r-th neighbour has better fitness the current i-th flow, it will follow the r-th neighbour flow. Otherwise, the i-th flow will move in the dominant slope direction as shown below:

$$X_i^{newf} = \begin{cases} X_i^{flow} + \overline{r_9}\left(X_r^{flow} - X_i^{flow}\right), \text{if } f\left(X_r^{flow}\right) < f\left(X_i^{flow}\right) \\ X_i^{flow} + r_{10}\left(X^{best} - X_i^{flow}\right), \quad \text{Otherwise} \end{cases} \tag{15}$$

The psuedocode of MCFDA to solve feature selection is explained in Fig. 1. The algorithm is repeated until the stopping condition of $\gamma > \gamma_{max}$ is attained, where γ and γ_{max} represents current fitness evaluation number and maximum fitness evaluation number, respectively.

4. Performance Evaluation on MCFDA

4.1. *Simulation Settings*

The performance of MCAVOA to solve feature selection problem is evaluated with 12 datasets taken from UCI Machine Learning Repository[15], i.e., (a). Breast Cancer Wisconsin (Diagnostic), (b) Dermatology, (c) Statlog (Heart), (d) Ionosphere, (e) Iris, (f) Arrhythmia, (g) Echocardiogram, (h) Haberman's Survival, (i) Diabetes, (j) MONK 1 Problem, (k) Zoo and (l) Letter Recognition. The performance of MCFDA is compared with four feature selection algorithms developed using Bezier Search Differential Evolution Algorithm (BeSD)[16], Coronavirus Herd Immunity Optimizer Algorithm (CHIO)[17], Flow Direction Algorithm (FDA)[18] and Improved Flame Generation Mechanism

(ODSFMFO)[19]. The performances of all algorithms are measured with mean accuracy Acc^{mean} and average number of selected features $N^{feature}$. Same population size of $I = 20$ and maximum fitness evaluation of $\tau^{max} = 2000$ are set for all algorithms. Each algorithm is simulated for 30 independent runs to solve datasets.

```
Algorithm: MCFDA-based Feature Selection Algorithm
Inputs: $i, j, \gamma, \gamma_{max}, X^{LB}, X^{UB}, k$
Define fitness functions and input variables;
for each $i$-th flow do
    for each $d$-th dimension do
    while $k < k_{max}$ do
        if $\sigma_k < 0.25$ then
            Select Circle map using Eq. (3);
        else if $0.25 \le \sigma_k < 0.5$ then
            Select Gauss map using Eq. (4);
        else if $0.50 \le \sigma_k < 0.75$ then
            Select Singer map using Eq. (5);
        else if $\sigma_k \ge 0.75$ then
            Select Sinusoidal map using Eq. (6);
        end if
        Update $\sigma_{k+1}$ and $k \leftarrow k + 1$;
    end while
    Initialize $X_{i,d}^{flow}$ based on Eq. (7);
    end for
end for
Evaluate fitness of flows and select the best flow
with lowest fitness value;
while $\gamma < \gamma_{max}$ do
    for each $i$-th flow do
        Update $Q$ based on Eq. (11);
        for each $j$-th neighbour do
            Update $\Delta$ and $X_j^{flow}$ with Eqs. (9) and (10);
            Evaluate $f\left(X_j^{flow}\right)$ using Eq. (2);
            $\gamma \leftarrow \gamma + 1$;
        end for
        Sort all $X_j^{flow}$ from best to worst fitness;
        if $f\left(X_j^{flow}\right) < f\left(X_i^{flow}\right)$ then
            Calculate $W_i$ using Eq. (13);
            Update each velocity flow $V$ with Eq. (12);
            Limit velocity for $V$ within $[v^{low}, v^{up}]$;
            Calculate $X_i^{newf}$ using Eq. (14);
        else
            Generate random integer number of $r$;
            Calculate $X_i^{newf}$ using Eq. (15);
        end if
        Evaluate $f\left(X_i^{newf}\right)$ using Eq. (2);
        $\gamma \leftarrow \gamma + 1$;
        Update current flow and best flow.
    end for
end while
Output: $X_{best}$ (i.e., optimal feature subset)
```

Fig. 1. Feature selection with MCFDA.

4.2. *Comparisons between competing algorithms*

It is desirable for a compared algorithm to produce higher Acc^{mean} and lower $N^{feature}$ when solving feature selection problem. Boldface style value in Table 1 and Table 2 is categorized as the best result, while the second-best result is presented in underlined and italic fronts. Table 1 shows that MCFDA has the highest accuracy by producing seven best Acc^{mean}, followed by BeSD and CHIO with four best Acc^{mean}, ODSFMFO with three best Acc^{mean} and FDA with one best Acc^{mean}. Table 2f shows that the proposed MCFDA solve 6 out of 12 datasets with best $N^{feature}$, followed by FDA with three best $N^{feature}$, BeSD and CHIO with two best $N^{feature}$ and ODSFMFO without any best $N^{feature}$.

Table 1. Comparison of Acc_{mean}.

Dataset	BeSD	CHIO	FDA	ODSF MFO	MCFDA
(a)	*0.972*	0.968	0.930	0.955	**0.973**
(b)	0.981	**1.000**	0.996	*0.997*	**1.000**
(c)	0.838	0.854	*0.910*	0.810	**0.912**
(d)	0.901	**0.951**	*0.944*	0.900	0.925
(e)	**1.000**	**1.000**	0.857	**1.000**	*0.967*
(f)	0.687	*0.736*	0.673	0.605	**0.772**
(g)	**1.000**	**1.000**	*0.979*	**1.000**	**1.000**
(h)	**0.836**	0.828	*0.830*	0.787	0.705
(i)	0.771	*0.781*	0.742	0.765	**0.789**
(j)	**0.933**	0.906	*0.913*	*0.913*	0.875
(k)	0.888	0.955	**0.995**	*0.987*	0.850
(l)	0.946	*0.951*	0.950	**0.958**	**0.958**

Table 2. Comparison of $N_{feature}$

Dataset	BeSD	CHIO	FDA	ODSF MFO	MCFDA
(a)	15.93	9.17	*5.93*	18.33	**5.57**
(b)	16.57	14.53	**12.20**	22.90	*13.07*
(c)	5.67	**5.10**	*5.63*	8.10	5.77
(d)	15.90	13.80	*10.27*	16.87	**9.30**
(e)	2.40	2.03	*2.00*	3.73	**1.00**
(f)	134.73	**112.17**	131.47	155.87	*126.07*
(g)	4.37	2.60	**1.00**	3.63	*2.00*
(h)	*1.87*	6.13	5.33	2.30	**1.00**
(i)	**3.47**	*4.07*	4.97	6.20	4.53
(j)	3.67	*3.10*	**3.00**	5.33	**3.00**
(k)	7.63	6.27	*5.00*	9.97	**4.20**
(l)	**7.87**	*10.57*	11.60	15.33	11.83

5. Conclusion

In this paper, a modified algorithm known as MCFDA is introduced to solve feature selection problems more competitively. The initial population of MCFDA with better solution quality is produced by using the ergodicity and non-repetition characteristics of multiple chaotic maps are used in the initialization to improve the initial position. This is followed by the search processes with balanced exploration and exploitation searches as inspired by the rainfall behavior. A total of 12 datasets are selected to evaluate the performance of MCFDA and its peer algorithms to solve feature selection problems. Based on the simulation results, it is concluded that MCFDA can surpass other competing feature selection algorithms with better classification accuracy and lesser numbers of selected features when solving majority of selected datasets.

Acknowledgements

This research is supported by Ministry of Higher Education Malaysia (MOHE) under the Fundamental Research Grant Scheme with the project codes of FRGS/1/2019/TK04/UCSI/02/1 and FRGS/1/2020/TK0/UCSI/02/4. This research is also supported by UCSI Research Excellence & Innovation Grant (REIG) with project code of REIG-FETBE-2022/038.

References

1. Visalakshi, S. & Radha, V. A literature review of feature selection techniques and applications: Review of feature selection in data mining. *2014 IEEE International Conference on Computational Intelligence and Computing Research*, 1-6 (2014).
2. Alrifaey, M., Lim, W. H. & Ang, C. K. A Novel Deep Learning Framework Based RNN-SAE for Fault Detection of Electrical Gas Generator. *IEEE Access* **9**, 21433-21442 (2021).
3. Alrifaey, M. *et al.* Hybrid Deep Learning Model for Fault Detection and Classification of Grid-Connected Photovoltaic System. *IEEE Access* **10**, 13852-13869 (2022).
4. Jdid, B., Hassan, K., Dayoub, I., Lim, W. H. & Mokayef, M. Machine Learning Based Automatic Modulation Recognition for Wireless Communications: A Comprehensive Survey. *IEEE Access* **9**, 57851-57873 (2021).
5. Jdid, B., Lim, W. H., Dayoub, I., Hassan, K. & Juhari, M. R. B. M. Robust Automatic Modulation Recognition Through Joint Contribution of Hand-Crafted and Contextual Features. *IEEE Access* **9**, 104530-104546 (2021).
6. Ahmad, M. F., Isa, N. A. M., Lim, W. H. & Ang, K. M. Differential evolution: A recent review based on state-of-the-art works. *Alexandria Engineering Journal* **61**, 3831-3872 (2022).
7. Hassan, C., Durai, V., Sapuan, S., A.A, N. & Mohamed Yusoff, M. Z. Mechanical and Crash Performance of Unidirectional Oil Palm Empty Fruit Bunch Fibre-reinforced Polypropylene Composite. *Bioresources* **13**, 8310-8328 (2018).
8. Tee, Z. Y., Yeap, S. P., Hassan, C. S. & Kiew, P. L. Nano and non-nano fillers in enhancing mechanical properties of epoxy resins: a brief review. *Polymer-Plastics Technology and Materials* **61**, 709-725 (2022).
9. Jamaludin, F. A. *et al.* Considering the effects of a RTV coating to improve electrical insulation against lightning. *2016 33rd International Conference on Lightning Protection (ICLP)*, 1-5 (2016).
10. Shaari, M. *et al.* Supervised evolutionary programming based technique for multi-DG installation in distribution system. *IAES International Journal of Artificial Intelligence (IJ-AI)* **9**, 11 (2020).
11. Berghout, T., Benbouzid, M., Muyeen, S. M., Bentrcia, T. & Mouss, L. H. Auto-NAHL: A Neural Network Approach for Condition-Based Maintenance of Complex Industrial Systems. *IEEE Access* **9**, 152829-152840 (2021).
12. Ahmad, M. F., Isa, N. A. M., Lim, W. H. & Ang, K. M. Differential evolution with modified initialization scheme using chaotic oppositional based learning strategy. *Alexandria Engineering Journal* **61**, 11835-11858 (2022).
13. Chow, V. T. Handbook of applied hydrology. *New York: Mc Graw Hilld* (1959).
14. Huang, P.-C. Analysis of Hydrograph Shape Affected by Flow-Direction Assumptions in Rainfall-Runoff Models. *Water* **12** (2020).
15. Dua, D. & Graff, C. *{UCI} Machine Learning Repository*, 2019).
16. Civicioglu, P. & Besdok, E. Bezier Search Differential Evolution Algorithm for numerical function optimization: A comparative study with CRMLSP, MVO, WA, SHADE and LSHADE. *Expert Systems with Applications* **165**, 113875 (2021).
17. Al-Betar, M. A., Alyasseri, Z. A. A., Awadallah, M. A. & Abu Doush, I. Coronavirus herd immunity optimizer (CHIO). *Neural Computing and Applications* **33**, 5011-5042 (2021).
18. Karami, H., Anaraki, M. V., Farzin, S. & Mirjalili, S. Flow Direction Algorithm (FDA): A Novel

Optimization Approach for Solving Optimization Problems. *Computers & Industrial Engineering* **156**, 107224 (2021).

19. Li, Z., Zeng, J., Chen, Y., Ma, G. & Liu, G. Death mechanism-based moth–flame optimization with improved flame generation mechanism for global optimization tasks. *Expert Systems with Applications* **183**, 115436 (2021).

Authors Introduction

Mr. Wy – Liang Cheng

He received the B.Eng. degree in Mechatronic Engineering with Hnours from UCSI University, Malaysia in 2020. He is currently pursuing Master of Philosophy in Engineering in UCSI University, Malaysia. His research interests are swarm intelligence and feature selection.

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.

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.

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. 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. 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. 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.

Performance Comparison of Convolutional Neural Network for COVID-19 Diagnosis

Suhaim Parvez Wadekar[1], Koon Meng Ang[1], Chin Hong Wong[2,3], Abhishek Sharma[4], Tiong Hoo Lim[5], Chun Kit Ang[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*
[5]*Faculty of Engineering, Universiti Teknologi Brunei, Bandar Seri Begawan 1410, Brunei Darussalam*
E-mail: 1001748138@ucsiuniversity.edu.my, 1001436889@ucsiuniversity.edu.my, chinhong.wong@mu.ie, abhishek15491@gmail.com, lim.tiong.hoo@utb.edu.bn, angck@ucsiuniversity.edu.my, tiangss@ucsiuniversity.edu.my limwh@ucsiuniversity.edu.my

Abstract

COVID-19 has devastated the global healthcare system as well as the world economy with more than 600 million confirmed cases and 6 million deaths globally. A timely and accurate diagnosis of the disease plays a vital role in the treatment and preventative spread of disease. Recently, deep learning such as Convolutional Neural Networks (CNNs) have achieved extraordinary results in many applications such as medical classifications. This work focuses on investigating the performance of nine state-of-the-art architectures: Alexnet, Googlenet, Inception-v3, Mobilenet-v2, Resnet-18, Resnet-50, Shufflenet, Squeezenet and Resnet-50 RCNN for COVID-19 classification by comparing with performance metrics such as accuracy, precision, sensitivity, specificity and F-score. The datasets considered in current study are divided into three different classes namely Normal Chest X-Rays (CXRs), Pneumonia patient CXR and COVID-19 patient CXR. The results achieved shows that Resnet-50 RCNN achieved an accuracy, precision, sensitivity, specificity and F-score of 95.67%, 95.71%, 95.67%, 97.84% and 95.67% respectively.

Keywords: Chest X-ray, convolutional neural network, COVID-19 diagnosis, deep learning

1. Introduction

COVID-19 is a highly infectious and contagious viral disease caused by the SARS-COV-2 virus. Over the past three years, this disease has caused ravage throughout the world by creating catastrophic damage on global economic as well as irreversible losses of human life. This disease first emerged in Wuhan, China in December 2019. As of January 2023, a total of 667,510.073 infection cases and 6,707,374 death cases are reported by World Health Organization (WHO)[1].

Despite being considered as the golden test used for COVID-19 diagnosis, reverse transcription polymerase chain reaction (RT-PCR) is a time-consuming test and it could take up to 48 hours to confirm infected cases. Thus, it is necessary to explore for other efficient alternatives that can address the drawbacks of RT-PCR. Existing studies show that the feasibility of using clinical imaging techniques such as chest X-ray (CXR) images for early diagnosis of COVID-19 with promising accuracy level[2]. From Fig 1, notable opacities can be observed from the CXR images of COVID-19 patients due to the SARS-COV-2 virus induced pneumonia on their lungs. These opacities are considered as the important features used by deep learning or machine learning based computer-aided diagnosis (CAD) systems to perform rapid diagnosis of COVID-19 based on the CXR images of patients.

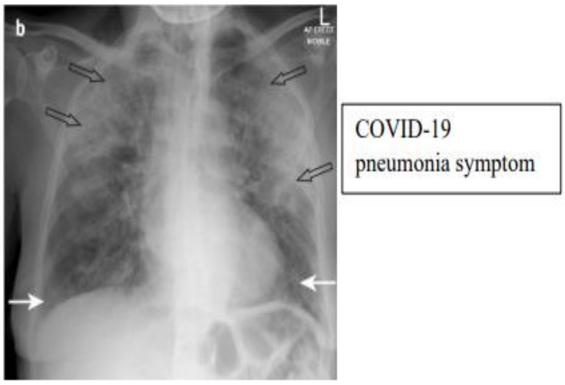


Figure 1: Opacities observed from the CXR images of COVID-19 patients

Convolutional neural network (CNN) is widely used to solve various real-world applications [2], [3], [4], [5], [6],[7] given its excellent capability to extract meaningful information from input sources and learn the nonlinear relationships between the input and expected outputs. The fundamental architecture of CNN consists of feature extractor (i.e., convolutional layers and pooling layers) and classifier (fully-connected layers). The breakthrough of deep learning researches motivated the development of various CNN architectures such as AlexNet[8], GoogleNet[9], Inception-v3[10], MobileNet-v2[11], ResNet-18[12], ResNet-50[12], ShuffleNet[13], SqueezeNet[14] and ResNet-50 RCNN[15].

While numerous works of COVID-19 diagnosis with deep learning technology were reported[16], the network performances of many existing architectures remain unexplored. This scenario opens an opportunity to further enhance the accuracy of COVID-19 diagnosis through the optimal selection of CNN architectures. Therefore, a comprehensive study is conducted in this paper to investigate the performance of previously mentioned network architectures in COVID-19 classification. The best CNN architecture is determined objectively via a set of statistical metrices and then implemented into an automated CAD system designed to identify COVID-19 diseases, which could be of great help to radiologist and other medical officers. Similar approach can also be used to design the automated CAD systems for other medical and non-medical applications.

2. Materials and Methods

2.1. *Datasets*

The CXR images used to train the abovementioned CNN architectures are obtained from a public database[17] contains 2700 images that can be equally divided into 3 classes, namely Normal, Pneumonia and COVID-19 cases. These images datasets are randomly divided into a ratio of 70:30 for training and testing sets, respectively. Before training the selected architecture with transfer learning for COVID-19 diagnosis, these input images are resized based on the requirements of respective network architecture as mentioned in Table 1. These input datasets are also pre-processed using data augmentation and converted into color images before the model training.

Table 1. Image resolution used by each architecture

CNN	Resolution
AlexNet	227×227×3
GoogleNet	224×224×3
Inception-v3	299×299×3
MobileNet-v2	224×224×3
ResNet-18	224×224×3
ResNet-50	224×224×3
ShuffleNet	224×224×3
SqueezeNet	227×227×3
ResNet-50 RCNN	224×224×3

2.2. *Transfer Learning of Pretrained Networks*

It is nontrivial to train the CNN networks from scratch for specific tasks because it involves tremendous amount of resources (e.g., training time, input datasets and etc.). Transfer learning is a promising solution used to tackle these drawbacks due to its ability to transfer the knowledge from one or more domains and apply the knowledge to another domain with a different target task. During transfer learning process, learnable parameters of these pretrained network architectures (i.e., AlexNet, GoogleNet, Inception-v3, MobileNet-v2, ResNet-18, ResNet-50, ShuffleNet, SqueezeNet and ResNet-50 RCNN) are extracted and applied on the same type of network for different purposes. The original output layers of these pretrained networks are also replaced with the new output layers containing three classes (i.e., COVID-19, pneumonia and normal cases) and trained with new datasets in Section 2.1 for COVID-19 diagnosis.

2.3. *Hyperparameters Settings*

Classification performances of pretrained networks used for COVID-19 can be governed by the hyperparameter settings of transfer learning process. Stochastic gradient descent (SGD) is used to train the selected architectures by minimizing the cross-entropy loss function. Five hyperparameters known as Momentum (MOM), Initial Learning Rate (ILR), Max Epoch (ME), Mini Batch Size (MBS), L2Regularization (L2) and Validation Frequency (VF) are considered. The best hyperparameter settings of each CNN architecture are determined based on trial and error and their values are presented in Table 2.

2.4. *Evaluation*

The performances of CNN architectures in COVID-19 diagnosis are evaluated using five metrics known as Accuracy, Sensitivity or Recall, Specificity, Precision and F1 score. These five metrics can be computed based on true positive (TP), true negative (TN), false positive (FP) and false negative (FN) results obtained from the confusion matrices of respective CNN architectures during the testing stage. Mathematical formulations of all performance metric are defined as follows.

Table 2: Hyperparameters used for all CNN architectures

CNN	MOM	ILR	ME	L2	MBS	VF
AlexNet	0.2	0.002	5	5e-4	32	35
GoogleNet	0.9	0.04	5	5e-4	32	35
Inception-v3	0.9	0.04	5	5e-4	32	35
MobileNet-v2	0.9	0.04	5	25e-4	32	35
ResNet-18	0.9	0.04	5	5e-4	32	35
ResNet-50	0.9	0.04	5	5e-4	32	35
ShuffleNet	0.9	0.04	5	25e-4	32	35
SqueezeNet	0.2	0.003	5	5e-4	32	35
ResNet-50 RCNN	0.9	0.04	5	5e-4	32	35

Accuracy represents the total numbers of correct predictions made by the CNN network, i.e.,

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

The probability of a CNN network to obtain true positive results is quantified as sensitivity or recall, where

$$Sensitivity = \frac{TP}{TP + FN} \quad (2)$$

In contrary to sensitivity or recall, specificity refers to the probability of a CNN network in obtaining true negative results, i.e.,

$$Specificity = \frac{TN}{TP + FN} \quad (3)$$

Precision metric is used to evaluate the accuracy of positive results obtained by a CNN network out of all positive predictions made as defined below:

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

Finally, the F1 score represents the harmonic score evaluation of precision and sensitivity obtained by a CNN network, where

$$F1\ Score = 2 \times \left(\frac{Sensitivity \times Precision}{Sensitivity + Precision}\right) \quad (5)$$

Apart from the quantitative performance metrics, the performance of CNN networks in COVID-19 diagnosis is also evaluated qualitatively using the confusion matrix and receiver operating characteristic (ROC) curve.

3. Performance Evaluations of CNN Networks

The performances of all CNN networks (i.e., AlexNet, GoogleNet, Inception-v3, MobileNet-v2, ResNet-18, ResNet-50, ShuffleNet, SqueezeNet and ResNet-50 RCNN) to diagnose COVID-19, pneumonia and normal cases based on CXR images of patients are evaluated. All pretrained networks are trained with 70% of training datasets and tested with 30% of testing datasets. During the transfer learning process, SDG is used to adjust the learnable parameters of CNN models based on the hyperparameter settings reported in Table 2.

Fig 2 illustrates the confusion matrices used to qualitatively compare the classification performances all CNN architectures after completing their training and testing processes. The effectiveness of all selected CNN network architectures in COVID-19 diagnosis are also qualitatively analyzed based on ROC curves presented in Fig 3. Note that the ROC curve of an CNN network is constructed based on the corresponding TP and FP values obtained when solving the COVID-19 image datasets. Accordingly, ResNet-50 RCNN is observed to dominate most pretrained networks (i.e., GoogleNet, MobileNet-v2, ResNet-50 and ShuffleNet) by producing more correct classifications on the COVID-19, pneumonia and normal cases. Although AlexNet, ResNet-18 and SqueezeNet can perform slightly better than ResNet-50 RCNN by producing relatively more correct classification on pneumonia case, the latter network architecture can produce significantly better results when classifying the remaining two cases (i.e., COVID-19 and normal). Similar observations are found when comparing Inception-v3 and ResNet-50 RCNN, where the former network has slightly better result on COVID-19 case but the latter architecture has more competitive classification performance on the pneumonia and normal cases. Based on the qualitative analysis, ResNet-50 RCNN is found to be most reliably used for COVID-19 diagnosis.

Five performance metrics explained in Eqs. (1) to (5) are also used to quantitatively compare the performances of all pretrained networks in COVID-19 diagnosis as shown in Table 3, where the best results are highlighted with bold font. Accordingly, ResNet-50 RCNN can produce the best results in terms of accuracy, sensitivity, specificity, precision and F1 score when classifying the COVID-19, pneumonia and normal cases from CXR images. The quantitative results reported in Table 3 are consistent with the qualitative ones as shown in Fig 2 and Fig 3. ResNet-50 is another competitive classifier used to solve COVID-19 datasets by producing second best results for all performance metrics. Meanwhile, both of AlexNet and ShuffleNet are the worst-performing CNN models in COVID-19 diagnosis.

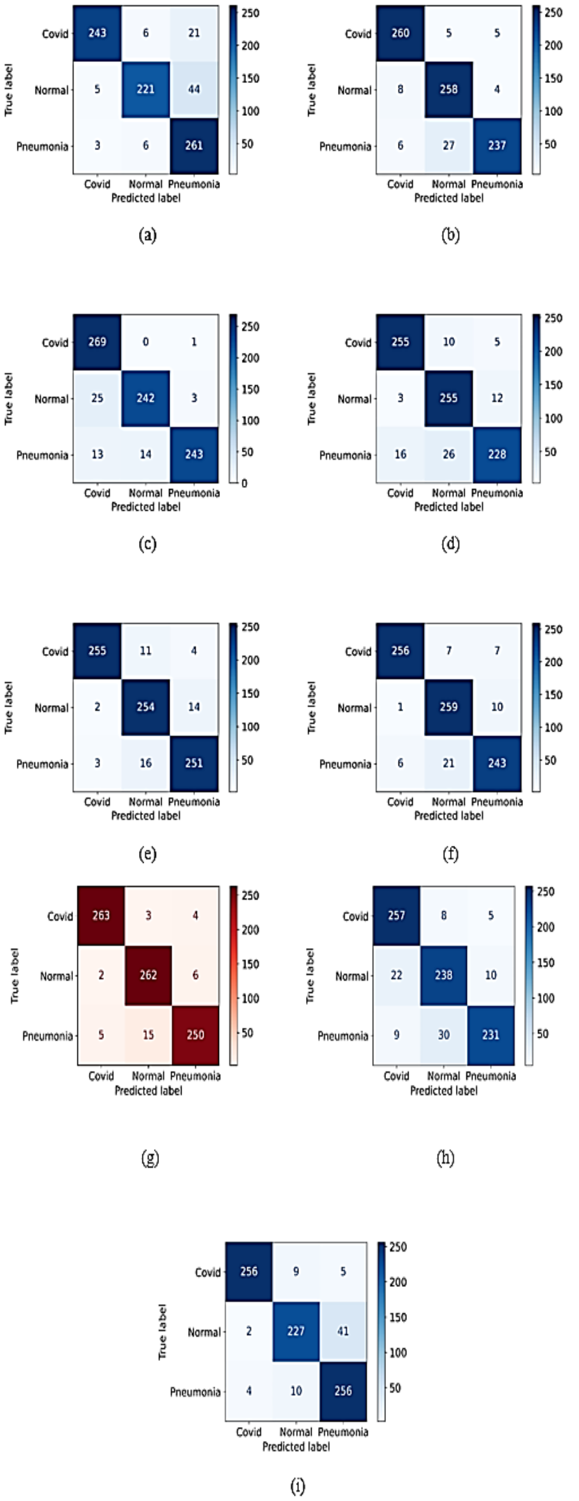


Figure 2: Confusion matrices of: (a) Alexnet, (b) Googlenet, (c) Inception-v3, (d) Mobilenet-v2, (e) Resnet-18, (f) Resnet-50, (g) Restnet-50 R-CNN, (h) Shufflenet and (i) Squeezenet in COVID-19 diagnosis using CXR images.

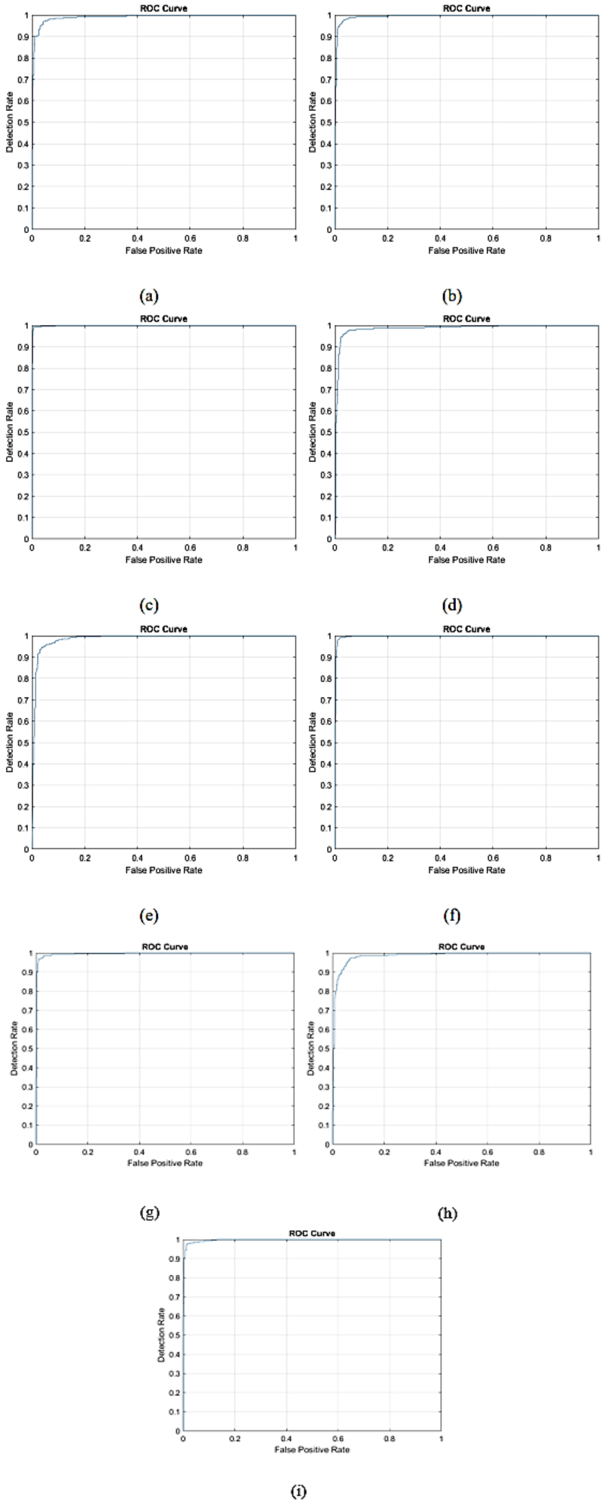


Figure 3: ROC curves of: (a) Alexnet, (b) Googlenet, (c) Inception-v3, (d) Mobilenet-v2, (e) Resnet-18, (f) Resnet-50, (g) Restnet-50 R-CNN, (h) Shufflenet and (i) Squeezenet in COVID-19 diagnosis using CXR images.

Table 3: Quantitative Performance Comparison Results

CNN	Acc.	Prec.	Sens.	Spec.	F1
Alexnet	89.50	90.50	89.50	94.75	89.58
Googlenet	93.21	93.39	93.21	96.60	93.19
Inception-v3	93.09	93.51	93.08	96.54	93.08
Mobilenet-v2	91.11	91.25	91.11	95.56	91.00
Resnet-18	91.73	92.10	91.73	95.86	91.75
Resnet-50	93.58	93.68	93.58	96.70	93.59
Shufflenet	89.63	89.79	89.63	94.81	89.61
Squeezenet	91.24	91.58	91.23	95.61	91.25
Resnet-50 RCNN	**95.67**	**95.71**	**95.67**	**97.84**	**95.67**

4. Conclusion

This paper proposes a deep learning-based CAD system by combining pretrained CNN models and transfer learning for the classifications of COVID-19, pneumonia and normal cases based on CXR images of patients. This current study aims to perform comprehensive analyses on the performances of AlexNet, GoogleNet, Inception-v3, MobileNet-v2, ResNet-18, ResNet-50, ShuffleNet, SqueezeNet and ResNet-50 RCNN to tackle COVID-19 diagnosis task based on different performance metrices. Transfer learning employed during the training process enables fast convergence of network without required long training time and large numbers of datasets. Simulation studies show that ResNet-50 RCNN achieves the best results with in terms of accuracy of 95.67%, precision of 95.71%, sensitivity of 95.67%, specificity of 97.84% and F1-score of 95.67%.

Acknowledgment

This research is supported by Ministry of Higher Education Malaysia (MOHE) under the Fundamental Research Grant Scheme with the project codes of FRGS/1/2019/TK04/UCSI/02/1 and FRGS/1/2020/TK0/UCSI/02/4. This research is also supported by UCSI Research Excellence & Innovation Grant (REIG) with project code of REIG-FETBE-2022/038.

References

1. COVID Live – Coronavirus Statistics Worldometer. https://www.worldometers.info/coronavirus/?utm_campaignomeAdvegas1?. Accessed 6 Jan 2023.
2. Berghout, T., Benbouzid, M., Muyeen, S. M., Bentrcia, T., Mouss L. Auto-NAHL: A Neural Network Approach for Condition-Based Maintenance of Complex Industrial Systems. *IEEE Access* **9**, 152829 – 152840 (2021).
3. Berghout, T., Benbouzid, M., Muyeen, S. M. Machine learning for cybersecurity in smart grids: A comprehensive review-based study on methods, solutions, and prospects. International Journal of Critical Infrastructure Protection, 38, 100547 (2022).
4. Alrifaey, M., Lim, W. H., Ang, C. K. A Novel Deep Learning Framework Based RNN-SAE for Fault Detection of Electrical Gas Generator. 21433-21442 (2021).
5. Alrifaey, M. *et al.* Hybrid Deep Learning Model for Fault Detection and Classification of Grid-Connected Photovoltaic System. *IEEE Access* **10**, 13852-13869 (2022).
6. Jdid, B., Hassan, K., Dayoub, I., Lim, W. H. & Mokayef, M. Machine Learning Based Automatic Modulation Recognition for Wireless Communications: A Comprehensive Survey. *IEEE Access* **9**, 57851-57873 (2021).
7. Jdid, B., Lim, W. H., Dayoub, I., Hassan, K. & Juhari, M. R. B. M. Robust Automatic Modulation Recognition Through Joint Contribution of Hand-Crafted and Contextual Features. *IEEE Access* **9**, 104530-104546 (2021).
8. Krizhevsky, A., Sutskever, I., Hinton., G. E. ImageNet Classification with Deep Convolutional Neural Networks. *Advances in neural information processing systems* (2012).
9. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1-9 (2015).
10. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2818-2826 (2016).
11. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A. and Chen, L.C. MobileNetV2: Inverted Residuals and Linear Bottlenecks. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 4510-4520, (2018).
12. He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770-778 (2016).
13. Zhang, X., Zhou, X., Lin, M., Sun, J., ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices. In *2018 IEEE/CVF*

Conference on Computer Vision and Pattern Recognition, 6848-6856 (2018).

14. Iandola, F. N., Song, H., Matthew W. M., Khalid, A., William, J. D., Kurt, K. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size. Preprint, submitted November 4 (2016).
15. Renjun, X., Junliang. Y., Yi, W., MengCheng, S. Fault Detection Method Based on Improved Faster R-CNN: Take ResNet-50 as an Example. *Geofluids*. (2022).
16. Siddiqui, S. et. al. Deep Learning Models for the Diagnosis and Screening of COVID-19: A Systematic Review. *SN Computer Science 3*, 1303 (2022).
17. Goel, T., Murugan, R., Mirjalili, S., Chakrabartty, D. K., OptCoNet: an optimized convolutional neural network for an automatic diagnosis of COVID-19. *Applied Intelligence* 51, 1351-1366 (2021).

Authors Introduction

Mr. Suhaim Parvez Wadekar

He received the B.Eng. degree in Mechatronic Engineering with Honours from UCSI University, Malaysia, in 2022. His research interests are swarm intelligence, machine learning and deep learning, and medical diagnosis system.

Mr. Koon Meng Ang

He received the B.Eng. degree in Mechatronic Engineering with Honours from UCSI University, Malaysia, in 2019. He is currently pursuing Ph.D. degree in UCSI University, Malaysia. 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. 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. 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. 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.

Classification of Wafer Defects with Optimized Deep Learning Model

Koon Hian Ang[1], Koon Meng Ang[1], Mohd Rizon Bin Mohamed Juhari[1], Chin Hong Wong[2,3], Abhishek Sharma[4], Chun Kit Ang[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*
E-mail: 1001850063@ucsiuniversity.edu.my, 1001436889@ucsiuniversity.edu.my, mohdrizon@ucsiuniversity.edu.my, chinhong.wong@mu.ie, abhishek15491@gmail.com, angck@ucsiuniversity.edu.my, tiangss@ucsiuniversity.edu.my, limwh@ucsiuniversity.edu.my

Abstract

Wafer defect inspection is one of the crucial semiconductor processing technologies because it can help to identify the surface defects in the process and eventually improve the yield. Manual inspection using human eye is subjective and long-term fatigue can lead to erroneous classification. Deep learning technology such as convolutional neural network (CNN) is a promising way to achieve automated wafer defect classification. The training of CNN is time consuming and it is nontrivial to fine tune its hyperparameters to achieve good classification performance. In this study, Arithmetic Optimization Algorithm (AOA) is proposed to optimize the CNN hyperparameters, such as momentum, initial learn rate, maximum epochs, L2 regularization, to reduce the burden brought by trial-and-error methods. The hyperparameters of a well-known pretrained model, i.e., GoogleNet, are optimized using AOA to perform wafer defects classification task. Simulation studies report that the AOA-optimized GoogleNet achieves promising accuracy of 91.32% in classifying wafer defects.

Keywords: Arithmetic optimization algorithms, wafer defects classification, convolutional neural networks, hyperparameters optimization.

1. Introduction

Emergence of Industrial Revolution 4.0 (IR4.0) era has led to the growing demands in semiconductor industries to produce sophisticated integrated circuit chips. More integrated circuit components are patterned and etched onto semiconductor wafers to meet various requirements of chips in terms of its accessing speed, lifespan, memory storage, size and etc. Likelihood of having manufacturing process-based defects on the wafer surface tends increase with the pressures of satisfying the growing demands of customer sides and this undesirable scenario can reduce the yields. One of the important steps used to address the manufacturing yield issue is to identify and classify the wafer defect patterns that can be associated with different steps of manufacturing process. Some typical issues of chip fabrication process include flow leakages, robot handoffs, contamination and etc. Engineers can improve the yield by locating the manufacturing problems of chips based on wafer defect patterns observed[1]. For most semiconductor industries, the wafer defect identification is still performed through human visual inspection. This manual process has undesirable drawbacks such as lack of objectivity and the high tendency of making false classifications due to long term fatigue issue.

To address the drawbacks of human inspection, an automated machine vision system incorporated with optimized deep learning model is designed for wafer defects classification in reliable manners. Convolutional neural network (CNN) is a popular deep learning method used to solve real-world problems[2],[3],[4],[5],[6],[7], including wafer defect classification[8],[9], motivated by its promising ability to learn the nonlinear relationships of input and output based on useful information extracted from raw data. Tremendous success of deep learning has driven the designs of various popular CNN architectures (e.g., AlexNet[10], ResNet-50[11], GoogleNet[12], VGG-16[13], etc.). Transfer learning is used to train these CNN networks to solve new tasks with lesser datasets and time. A crucial factor that governs the performance of pretrained network to solve new tasks is hyperparameter settings used during the training process.

Conventionally, CNN hyperparameters are manually tuned in trial-and-error basis but it is time-consuming. Given their strong global search ability, metaheuristic search algorithms (MSAs) inspired by various natural phenomena[15] (e.g., evolution theory, animal behaviors, physics principles and human activities) are used to solve many complex optimization problems[16],[17],[18],[19],[20],[21], including the hyperparameter tuning of CNN. Arithmetic optimization algorithm (AOA)[14] is an emerging MSA inspired by the distribution behaviors of four major arithmetic operators (i.e., addition, subtraction, multiplication and division). In this study, GoogleNet is selected as a pretrained network and trained with new datasets via transfer learning for solving wafer defect classification task. AOA is used to optimize four hyperparameters if CNN, i.e., momentum, initial learn rate, maximum epochs and L2 regularization. Performance of optimized CNN model in classifying wafer defects is investigated.

2. Related Works

2.1. *Conventional CNN and GoogleNet*

A typical CNN model consists of three different layers, i.e., convolutional layer, pooling layer and fully connected layer[22]. Convolution layer is mainly used to extract desired features from input images. Pooling layer is applied to reduce the size of the feature maps produced by the convolution layer. Subsequently, a fully connected layer acts as the classification module of a CNN model. Table 1 presents an example of a typical CNN model.

Table 1. Architecture of a typical CNN.

Layer	Type of Layer	#Feature maps	Feature map size	Filter size
1	Input	1	14 × 14	-
2	Convolutional 1	6	7 × 7	5 × 5
3	Pooling 1	6	4 × 4	2 × 2
4	Convolutional 2	16	4 × 4	5 × 5
5	Pooling 2	16	2 × 2	2 × 2
6	Fully connected 1	1	120	-
7	Fully connected 2	1	84	-

GoogleNet was proposed by Szegedy et al.[12] and it has 7 million parameters. The network architecture of GoogleNet consists of nine inception modules, four convolutional layers, four max-pooling layers, three average pooling layers, five full-connected layers, and three SoftMax layers for the main auxiliary classifiers in network. GoogleNet also has the dropout regularization and ReLU activation functions in its fully connected layers and convolutional layers.

2.2. *Basic AOA*

AOA was proposed by Abualigah, *et al.*[14] in year 2020 and its search mechanism was inspired by the distribution behaviors of four popular arithmetic operators, known as Addition (*A*), Subtraction (*S*), Multiplication (*M*) and Division (*D*). The search process AOA is implemented with three main phases, i.e., initialization, exploration and exploitation phases to solve an optimization problem.

During initialization phase, a set of possible solutions are generated randomly within the predefined boundaries of each dimensional component. A Math Optimizer Accelerated (*MOA*) that controls the search behavior of algorithm in either exploration phase or exploitation phase, is then calculated as follow:

$$MOA(C_{Iter}) = Min + C_{Iter} \times \left(\frac{Max-Min}{M_{Iter}}\right) \tag{1}$$

where $MOA(C_{Iter})$ is the function value at *t*-th iteration; C_{Iter} is current iteration number; M_{Iter} is the predefined maximum number of iterations; *Min* and *Max* indicate the minimum and maximum values of accelerated function, respectively. If *MOA* is smaller than a random number with value between 0 to 1 that are produced by a uniform, the AOA population is assigned with exploration phase. Otherwise, exploitation phase is performed.

During exploration phase, both *M* and *D* operators are applied to increase the coverage of AOA population in search space with good solution diversity given their high-distributed values with high dispersion. The *d*-th dimension for the new position of each *i*-th AOA solution is updated in exploration phase as:

$$x_{i,j}(C_{Iter}+1) = \begin{cases} best_j \div (MOP + \in) \times \left(\left(UB_j - LB_j\right) \times \mu + LB_j \right), & r1 < 0.5 \\ best_j \times MOP \times \left(\left(UB_j - LB_j\right) \times \mu + LB_j \right), & otherwise \end{cases} \tag{2}$$

where $x_{i,j}(C_{Iter}+1)$ refers to the j-th dimension of i-th solution in next iteration of $C_{Iter}+1$, where $j=1,\dots,D$, $i=1,\dots,I$ with I refers to population size, $C_{Iter}=1,\dots,M_{Iter}$ with M_{Iter} refers to maximum iteration number of algorithm; $best_j$ is the j-th dimension of the best solution; UB_j and LB_j are the upper and lower boundary limits in the j-th dimension, respectively; $r1$ is a random number within 0 to 1 generated by uniform distribution; μ is a parameter used to adjust the range of search process. Given the values of C_{Iter} and M_{Iter}, a Math Optimizer Probability (*MOP*) can be formulated as:

$$MOP(C_{Iter}) = 1 - \frac{C_{Iter}^{\frac{1}{\alpha}}}{M_{Iter}^{\frac{1}{\alpha}}} \tag{3}$$

where $MOP(C_{Iter})$ is a function value at *t*-th iteration; C_{Iter} is the current number iteration; α is a sensitive parameter that used to define exploitation accuracy.

For exploitation phase, *S* and *A* operators are applied to fine tune the promising solution regions around global optimum by leveraging its high-dense results with low dispersion. The *d*-th dimension for the new position of each *i*-th solution is updated in exploitation phase as:

$$x_{i,j}(C_{Iter}+1) = \begin{cases} best_j - MOP \times \left(\left(UB_j - LB_j\right) \times \mu + LB_j\right), & r2 < 0.5 \\ best_j + MOP \times \left(\left(UB_j - LB_j\right) \times \mu + LB_j\right), & otherwise \end{cases} \tag{4}$$

where *r*2 is a random number within 0 to 1 generated by uniform distribution. The exploration and exploitation processes of AOA are repeated iteratively until the stopping criteria are satisfied. The best solution produced by AOA at the end of the optimization process is returned to solve the given problem.

3. Optimization of CNN Models Using AOA

3.1. *Dataset Preprocessing*

In this study, the WM-811K dataset is used to perform the training and evaluation processes of the optimized deep learning model. It is notable that the data distributed across all classes are heavily unbalanced. A total of 30,000 None defect images is extracted from the WM-811K dataset and compiled with all other defect images[7]. Table 2 summarizes the number of all defect and non-defect images observed in the newly formed dataset.

Table 2. Number of defect and non-defect images

Types of Defects	# Labeled Images
Center	4,294
Donut	555
Edge-Loc	5,189
Edge-Ring	9,680
Loc	3,593
Near-full	149
Random	866
Scratch	1,193
None	30,000

3.2. *AOA-Optimized GoogleNet*

Transfer learning used to train GoogleNet with updated WM-811K datasets presented in Table 2, enabling this pretrained network to solve wafer defects classification problems. In order to perform the new classification task effectively, several modifications are made on the hyperparamters of GoogleNet. Particularly, the size of its input layer is modified from the original value of 224 × 224 × 3 to 64 × 64 × 1. Besides that, the padding size of first convolution layer is changed from its original value of 3 × 3 × 3 × 3 to be same as the input size. Finally, the output size of fully connected layer is changed from its original value of 1000 to be 9 (i.e., defect types).

To further enhance the performance of GoogleNet in wafer defect classification task, AOA is implemented to optimize the hyperparameters of GoogleNet during the transfer learning process. Each *i*-th AOA solution vector of AOA is encoded with four decision variables known as momentum, initial learn rate, maximum epochs and L2 regularization, where their boundary limits are presented in Table 3. Fitness function used to evaluate the quality of each AOA solution can be defined based on the classification accuracy obtained. The overall framework of the proposed AOA-optimized GoogleNet used for wafer defect classification is described in Fig 1.

Table 3. Search range of four hyperparameters.

Hyperparameters	Lower Boundary	Upper Boundary
Momentum	0.5	0.9
Initial Learn Rate	0.01	0.1
Maximum Epochs	5	10
L2 Regularization	1×10^{-4}	5×10^{-4}

4. Performance Evaluations

4.1. *Simulation settings*

The pre-processed WM-811K dataset is randomly split for training, validating and testing purposes with the ratio of 70%, 10% and 20%, respectively. All the input images

```
Algorithm 1: AOA-Optimized GoogleNet
Input: $N$, $D$, $UB_j$, $LB_j$
Initialize $C_{Iter} = 0$;
for $i$ =1 to $I$ do
    Randomly generate solution $x_i$;
    $C_{Iter} = C_{Iter} + 1$;
end for
while $C_{Iter} \leq M_{Iter}$ do
    Decode the hyperparameters from $x_i$;
    Evaluate the accuracy $f(x_i)$ of GoogleNet;
    Calculate $MOP$ using Eq. (3);
    Calculate $MOA$ using Eq. (1);
    for $i$ =1 to $I$ do
        for $j$ =1 to $D$ do
            if $rand > MOA$ then
                Update $x_{i,j}(C_{Iter} + 1)$ with Eq. (2);
            else
                Update $x_{i,j}(C_{Iter} + 1)$ with Eq. (4)
            end if
        end for
        Decode the hyperparameters from $x_i$;
        Evaluate the accuracy $f(x_i)$ of GoogleNet;
        Update $x_i$, $f(x_i)$, $best$, $f(best)$
        $C_{Iter} \leftarrow C_{Iter} + 1$;
    end for
end while
Output: $best$
```

Fig.1 Pseudocode of optimizing GoogleNet using AOA.

are resized into 64 × 64 × 1. Besides that, the minimum batch size is set as 32.

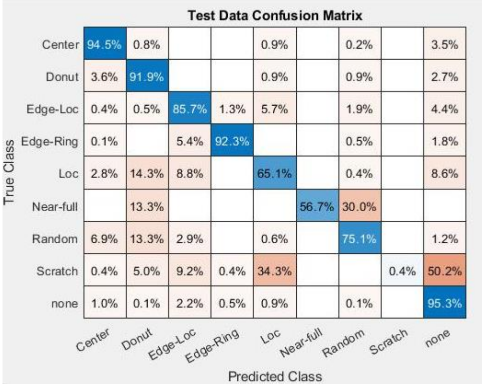
4.2. *Performance Comparisons*

The performance of unoptimized GoogleNet and AOA-Optimized GoogleNet used to classify the wafer defects are evaluated and compared based on the values of recall, accuracy, precision, F1 score and area under curve (AUC). Recall measures the model's ability to detect positive samples. Accuracy reports the numbers of data are predicted correctly to the class it supposed to be. Precision refers to the accuracy of CNN in classifying a sample as positive. F1 score is the harmonic mean of CNN's precision and recall. AUC represents the area under receiver operating characteristic curve. Table 4 compares the classification performance of GoogleNet and AOA-Optimized GoogleNet in quantitative manner based on five performance metrics described. Qualitative performance analyses are also conducted based on the confusion matrices produced by unoptimized GoogleNet and AOA-Optimized GoogleNet as shown in Fig 2 and Fig 3, respectively.

Table 4 reports that the proposed AOA-optimized GoogleNet shows significantly better performance over the unoptimized GoogleNet, in terms of recall, accuracy, precision, F1 score and AUC, when solving the wafer defects classification problems. According to Fig 2 and Fig 3, GoogleNet without hyperparameters optimization is reported to suffer severe issue by misclassifying the scratch defect to be Loc defects (34.3%) and non-defected (50.2%). This misclassification behavior has been rectified by optimizing the hyperparameters of GoogleNet using AOA as illustrated in Fig 3.

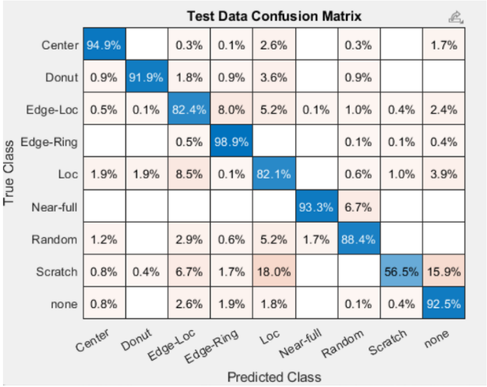
Table 4. Quantitative performance comparison.

Performance	Unoptimized-GoogleNet	AOA-Optimized GoogleNet
Recall	0.7300	**0.8676**
Accuracy	0.8936	**0.9132**
Precision	0.8185	**0.8523**
F1 Score	0.7000	**0.8566**
AUC	0.7735	**0.8151**



	Center	Donut	Edge-Loc	Edge-Ring	Loc	Near-full	Random	Scratch	none
Center	94.5%	0.8%			0.9%		0.2%		3.5%
Donut	3.6%	91.9%			0.9%		0.9%		2.7%
Edge-Loc	0.4%	0.5%	85.7%	1.3%	5.7%		1.9%		4.4%
Edge-Ring	0.1%		5.4%	92.3%			0.5%		1.8%
Loc	2.8%	14.3%	8.8%		65.1%		0.4%		8.6%
Near-full		13.3%				56.7%	30.0%		
Random	6.9%	13.3%	2.9%		0.6%		75.1%		1.2%
Scratch	0.4%	5.0%	9.2%	0.4%	34.3%			0.4%	50.2%
none	1.0%	0.1%	2.2%	0.5%	0.9%		0.1%		95.3%

Fig.2 Confusion matrix of unoptimized GoogleNet.



	Center	Donut	Edge-Loc	Edge-Ring	Loc	Near-full	Random	Scratch	none
Center	94.9%		0.3%	0.1%	2.6%		0.3%		1.7%
Donut	0.9%	91.9%	1.8%	0.9%	3.6%		0.9%		
Edge-Loc	0.5%	0.1%	82.4%	8.0%	5.2%	0.1%	1.0%	0.4%	2.4%
Edge-Ring			0.5%	98.9%			0.1%	0.1%	0.4%
Loc	1.9%	1.9%	8.5%	0.1%	82.1%		0.6%	1.0%	3.9%
Near-full						93.3%	6.7%		
Random	1.2%		2.9%	0.6%	5.2%	1.7%	88.4%		
Scratch	0.8%	0.4%	6.7%	1.7%	18.0%			56.5%	15.9%
none	0.8%		2.6%	1.9%	1.8%		0.1%	0.4%	92.5%

Fig.3 Confusion matrix of AOA-Optimized GoogleNet.

5. Conclusions

An optimized deep learning model is proposed to solve wafer defect classification problems more effectively. In particularly, transfer learning process is performed on the GoogleNet based on the wafer defects datasets and its hypermeter settings are optimized by AOA. Simulation results show that AOA-optimized GoogleNet has better performance than unoptimized GoogleNet to solve wafer defects classification by producing more competitive values of recall, accuracy, precision, F1 score and AUC.

Acknowledgements

This research is supported by Ministry of Higher Education Malaysia (MOHE) under the Fundamental Research Grant Scheme with the project codes of FRGS/1/2019/TK04/UCSI/02/1 and FRGS/1/2020/ TK0/UCSI/02/4. This research is also supported by UCSI Research Excellence & Innovation Grant (REIG) with project code of REIG-FETBE-2022/038.

References

1. Chien, J.C., Wu, M. T., Lee, J. D. Inspection and Classification of Semiconductor Wafer Surface Defects Using CNN Deep Learning Networks. *Applied Sciences*, **10**, 5340 (2020).
2. Berghout, T., Benbouzid, M., Muyeen, S. M., Bentrcia, T., Mouss L. Auto-NAHL: A Neural Network Approach for Condition-Based Maintenance of Complex Industrial Systems. *IEEE Access* **9**, 152829 – 152840 (2021).
3. Berghout, T., Benbouzid, M., Muyeen, S. M. Machine learning for cybersecurity in smart grids: A comprehensive review-based study on methods, solutions, and prospects. International Journal of Critical Infrastructure Protection, **38**, 100547 (2022).
4. Alrifaey, M., Lim, W. H., Ang, C. K. A Novel Deep Learning Framework Based RNN-SAE for Fault Detection of Electrical Gas Generator. *IEEE Access* **9**, 21433-21442 (2021).
5. Alrifaey, M. *et al.* Hybrid Deep Learning Model for Fault Detection and Classification of Grid-Connected Photovoltaic System. *IEEE Access* **10**, 13852-13869 (2022).
6. Jdid, B., Hassan, K., Dayoub, I., Lim, W. H. & Mokayef, M. Machine Learning Based Automatic Modulation Recognition for Wireless Communications: A Comprehensive Survey. *IEEE Access* **9**, 57851-57873 (2021).
7. Jdid, B., Lim, W. H., Dayoub, I., Hassan, K. & Juhari, M. R. B. M. Robust Automatic Modulation Recognition Through Joint Contribution of Hand-Crafted and Contextual Features. *IEEE Access* **9**, 104530-104546 (2021).
8. Batool, U., Shapiai, M. I., Tahir, M., Ismail, Z. H., Zakaria, N. J., Elfakharany, A. A Systematic Review of Deep Learning for Silicon Wafer Defect Recognition, *IEEE Access*, **9**, 116572-116593 (2021).
9. Bhatt, D., et al. CNN Variants for Computer Vision: History, Architecture, Application, Challenges and Future Scope. *Electronics*, **10**, 2470, (2021) .
10. Krizhevsky, A., Sutskever, I., Hinton., G. E. ImageNet Classification with Deep Convolutional Neural Networks. *Advances in neural information processing systems* (2012).
11. He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770-778 (2016).
12. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1-9 (2015).
13. Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014).
14. Abualigah, L., Diabat, A., Mirjalili, S., Abd Elaziz, M., & Gandomi, A. H., "The Arithmetic Optimization Algorithm", *Computer methods in applied mechanics and engineering* **376**, 113609, (2021).
15. Ahmad, M. F., Isa, N. A. M., Lim, W. H. & Ang, K. M. Differential evolution: A recent review based on state-of-the-art works. *Alexandria Engineering Journal* **61**, 3831-3872 (2022).
16. A. Sharma, A. Sharma, V. Jately, M. Averbukh, S. Rajput, B. Azzopardi. A Novel TSA-PSO Based Hybrid Algorithm for GMPP Tracking under Partial Shading Conditions," *Energies* **15**, 3164 (2022).
17. 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**, 564 (2022) .
18. Shaari, M. et al. Supervised evolutionary programming-based technique for multi-DG installation in distribution system. *IAES International Journal of Artificial Intelligence (IJ-AI)* **9**, 11 (2020).
19. Tee, Z. Y., Yeap, S. P., Hassan, C. S. & Kiew, P. L. Nano and non-nano fillers in enhancing mechanical properties of epoxy resins: a brief review. *Polymer-*

Plastics Technology and Materials **61**, 709-725 (2022).

20. Hassan, C., Durai, V., Sapuan, S., A.A, N. & Mohamed Yusoff, M. Z. Mechanical and Crash Performance of Unidirectional Oil Palm Empty Fruit Bunch Fibre-reinforced Polypropylene Composite. *Bioresources* **13**, 8310-8328 (2018).
21. Jamaludin, F. A. et al. Considering the effects of a RTV coating to improve electrical insulation against lightning. *2016 33rd International Conference on Lightning Protection (ICLP)*, 1-5 (2016).
22. Ang K. M. et al. Optimal Design of Convolutional Neural Network Architectures Using Teaching–Learning-Based Optimization for Image Classification. *Symmetry* **14**, 2323 (2022).

Authors Introduction

Mr. Koon Hian Ang

He is currently pursuing Bachelor of Mechatronics Engineering with Honours as final year student in Faculty of Engineering, Technology and Built Environment, UCSI University, Malaysia. His research interests are machine learning, deep learning, and optimization algorithm.

Mr. Koon Meng Ang

He received the B.Eng. degree in Mechatronic Engineering with Honours from UCSI University, Malaysia, in 2019. He is currently pursuing Ph.D. degree in UCSI University, Malaysia. His research interests are swarm intelligence, machine learning and deep learning.

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.

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. 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.

Compact Wearable Antenna for Millimeter-Wave (mm-Wave) Fifth Generation

Wai Kiat Wong[1], **Sew Sun Tiang**[1, *], **Wei Hong Lim**[1], **Mastaneh Mokayef**[1], **Eryana Eiyda**[1], **Chin Hong Wong**[2,3]
([1]UCSI University, Malaysia) ([2]Fuzou University, China) ([3]Maynooth University, Ireland)
E-mail: 1001954282@ucsiuniversity.edu.my, tiangss@ucsiuniversity.edu.my, limwh@ucsiuniversity.edu.my,
mastaneh@ucsiuniversity.edu.my, eryanaeiyda@ucsiuniversity.edu.my, chinhong.wong@mu.ie
http://www.ucsiuniversity.edu.my/

Abstract

The need for networking, communication, and data sharing capabilities among users of wearable terminal devices has increased, and this has made the new wearable antenna one of the most active research areas. This work presents a wearable antenna for 5G applications based on a microstrip patch antenna operating at 28GHz millimeter-wave (mm-wave). The operating frequency of 28GHz is expected to be appropriate for 5G mm-wave wearable antenna design. The design is made of the semi-flexible Rogers Duroid RO3003 substrate, which has a thickness of 0.75mm, a loss tangent of 0.001 and a relative permittivity of 3. CST Microwave Studio software is used to analyze and evaluate the proposed antenna's performance to other existing designs in terms of return loss, bandwidth, gain, directivity, and point SAR value.

Keywords: 5G, Millimeter-Wave (mm-wave), Wearable Antenna, Specific Absorption Rate (SAR).

1. Introduction

Due to high demand for secure, fast, and large data transmission rates in many recent and advanced applications, such as broadcasting, Internet of Things (IoT), automobiles, smart cities, energy, and wearable devices, global mobile data traffic has grown significantly in recent years. The capacity and performance of each generation of mobile and wireless communication systems have been enhanced to fulfil those expectations [1].

The bandwidth is one important technique to increase capacity and data rates in current and future mobile and wireless generations. The bandwidth is directly proportional to the data rates. Higher data rates are possible because of the increased bandwidth. For current frequency ranges, such as the 1.7 GHz GSM band, the 1.8 GHz 4G/LTE band, the 2.0 GHz 4G/LTE band, the 2.1 GHz LTE band, and the 2.6 GHz band, all have restricted capacity. For 5G applications, high-frequency bands such as 24, 28, 37, and 39 GHz, as well as certain future suggested bands such as 47 GHz and 60 GHz, have recently been explored [1]. These high-frequency bands, also known as millimeter-Wave (mm-Wave) bands, can provide a huge amount of bandwidth, exceeding 500 MHz. The key benefit of mm-Wave technology is that it reduces the volume of devices by having very high resonance frequencies, resulting in a smaller antenna and higher speed and capacity [2]. Despite this, current 5G communication continues to use the sub-6 GHz spectrum which include 3.3 GHz to 4.2 GHz and 4.4 GHz to 5 GHz. In the recent decade, the demand for wearable devices has skyrocketed. Wearable devices that work on many bands, such as 3G, 4G, Wi-Fi, and GPS, are already crowded, but they are getting smarter and smaller. As a result, multiband antennas and smaller antennas are preferred. The propagation losses of mm-wave communications, which operate at high frequencies, are quite large [3]. To reduce propagation losses, the employed antenna should have high directional gain radiation patterns which is directed in the direction of wave propagation. Wearable antennas are one of the most important components of wearable electronics, which are used in a variety of applications from medical to military to entertainment and other everyday wearable devices [4]. Medical equipment for patient monitoring, smartwatches with embedded miniature antennas, military tracking and navigation systems, body-worn camera with Wi-Fi and Bluetooth, glasses for augmented reality, and wearable sporting gear are only a few examples of wearable antennas.

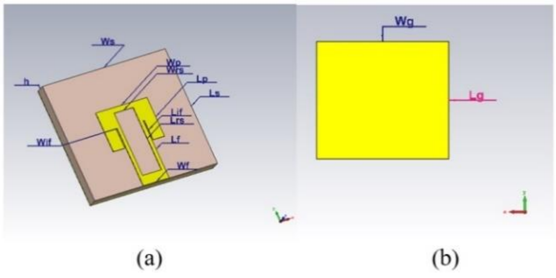
Nonetheless, the wearable antenna's design is crucial, especially for 5G mm-wave and IoT applications, where the manufacturing process and tolerances at higher frequencies have a significant impact on performance. When developing a wearable antenna for 5G applications for use as an integrable element of worn devices, numerous factors must be considered. They must be conformal or flexible, durable, and capable of operating in close proximity to the human body with little performance deterioration. The human body is well recognized for degrading antenna efficiency and gain owing to natural body tissue losses; thus, the implementation environment must be addressed during the design phase to develop a highly stable and durable 5G wearable antenna [5]. As one of the key criteria for such devices, the wearable antenna should be able to work successfully under various bending situations. On the other hand, the materials used as substrates and conductive portions for wearable antennas are critical [6]. They must be carefully designed to give the essential mechanical or physical properties, such as bending, wrapping, and occasionally washing, with little performance impact.

Other than that, wearable antennas are antennas that function within the human body, allowing the human body to absorb part of the radiated energy [5]. Human tissues will be damaged and burned because of the absorption of these waves. Therefore, when wearable antennas are employed, it is essential to decrease the interaction of electromagnetic radiation with human body tissues [3]. A specific absorption rate (SAR) is used to evaluate the absorption of electromagnetic waves in human tissue. SAR calculations may be made using either the point SAR technique or the average of SAR in terms of mass or volume. Point SAR is determined for each grid cell by dividing the absorbed power in each grid by the grid mass. Point SAR is also the value without mass averaging and defines the highest SAR of all the grid cells. For averaged SAR measurements, each point is represented by a cube with a given mass, either 1g or 10 g, and the power loss density is integrated on this cube. The cube's mass is divided by the power loss in integral form at the end [7]. Therefore, the SAR limit is determined by the Standardization Committee and varies by area throughout the world. The Federal Communications Commission (FCC) regulates SAR in the United States, with a maximum allowable value of 1.6 W/kg averaged over 1 gram of tissue at the frequency less than 6 GHz [8]. However, in Europe, the International Commission on Non-Ionizing Radiation Protection (ICNIPR) established a maximum tolerable SAR of 2.0 W/kg averaged over 10 grams of tissue at the frequency of less than 10 GHz. Due to the benefits of low profile, small size, light weight, and ease of manufacturing, the microstrip patch antenna has been selected for the 5G wearable antenna design in this work. The proposed antenna is operated in mm-wave spectrum since its operational frequencies is 28 GHz. Transmission line model calculation was used to determine the size of the patch antenna and inset feed. CST Microwave Studio software is used to simulate the antenna and analyze its parameters including SAR.

2. Related Works

2.1. *Antenna Design*

Fig. 1 shows the geometry of the proposed antenna. The patch antenna with a copper thickness of 0.035 mm is modelled on a semi-flexible Rogers Duroid RO3003 substrate with a relative permittivity of 3.0, a loss tangent of 0.001, and a thickness of 0.75 mm. The patch and ground plane of a microstrip antenna is crucial for designing an antenna since it determines the antenna's bandwidth, size, gain, and efficiency. Table 1 shows all the optimized dimension of the proposed antenna.



(a) (b)

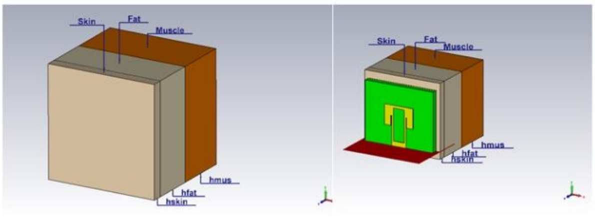
Fig. 1. Geometry of 5G Wearable Antenna Operating at 28GHz: (a) Front View, (b) Rear View.

Table 1. Dimensions of 5G Wearable Antenna Operating at 28 GHz

Parameters	Annotation	Value (mm)
Substrate Width	Ws	9
Substrate Length	Ls	8
Ground Width	Wg	9
Ground Length	Lg	8
Patch Width	Wp	3.75
Patch Length	Lp	2.7
Feed Line Width	Wf	1.9
Feed Line Length	Lf	4
Inset Feed Width	Wif	0.09
Inset Feed Length	Lif	1.35
Slot Width	Wrs	1.4
Slot Length	Lrs	4.35
Substrate Height	h	0.75
Copper Height	ht	0.035

2.2. *Antenna Simulation of Specific Absorption Rate (SAR) Level*

To calculate the SAR value as the antenna is to be worn on the body, a numerical human phantom model has been created in CST Microwave Studio software, as illustrated in Fig. 2. The human phantom model consists of three layers of human tissues: skin, fat, and muscle. Table 2 lists the dielectric properties of human tissues at 28 GHz, and Table 3 lists the tissues' thermal characteristics [7].



(a) (b)

Fig. 2. Numerical Human Phantom Model: (a) Human Tissue Layers, (b) Perspective View.

Table 2. Dielectric Properties of Human Tissues at 28 GHz.

Tissue	Mass Density (kg/m3)	Relative Permittivity (Ɛr)	Conductivity (S/m)
Skin	1109	16.55	25.82
Fat	911	6.09	5.04
Muscle	1090	24.43	33.6

Table 3. Thermal Parameters of Human Tissues.

Tissue Properties	Skin	Fat	Muscle
Heat Capacity (kJ/K/kg)	3.391	2.348	3.421
Thermal Conductivity (W/K/m)	0.37	0.21	0.49
Metabolic Rate (W/m3)	1620	300	480
Thickness (mm)	1	4	6

3. Results and Discussion

In this work, the performance of the 5G wearable antenna was evaluated through the simulation using CST Microwave Studio software. Through the simulation, the S-Parameters, Bandwidth, Gain, Directivity, VSWR, Smith Chart, Surface Current, and SAR analysis of proposed antenna are obtained as shown in below. Fig. 3. shows the simulated return loss of the proposed 5G wearable antenna. From the simulation result, the 5G wearable antenna resonates at 27.995 GHz with a return loss of -21.451 dB and a wide bandwidth covering from 27.221 GHz to 28.738 GHz with respect to -10 dB.

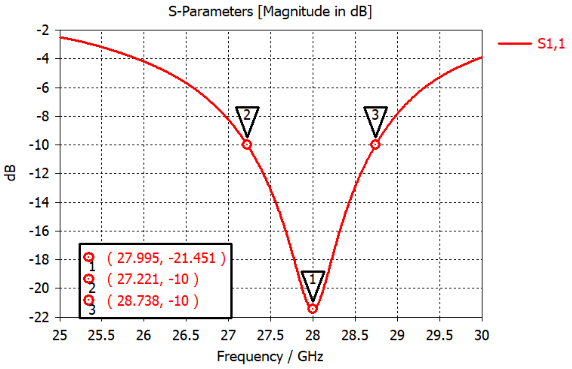
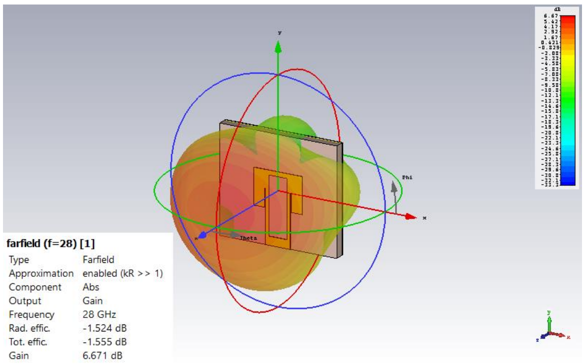


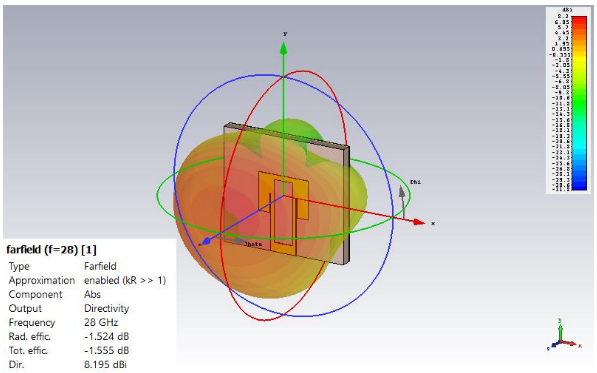
Fig. 3. Simulated Return Loss (S11) of the Proposed Antenna

Fig. 4 shows the simulated gain of the proposed 5G wearable antenna. From the simulation result, the proposed 5G wearable antenna has a gain of 6.671 dB at 27.995 GHz. Fig. 5 illustrates the simulated directivity of the proposed 5G wearable antenna. From the simulation result, the proposed 5G wearable antenna has a directivity of 8.195 dBi at 27.995 GHz.



farfield (f=28) [1]
Type Farfield
Approximation enabled (kR >> 1)
Component Abs
Output Gain
Frequency 28 GHz
Rad. effic. -1.524 dB
Tot. effic. -1.555 dB
Gain 6.671 dB

Fig. 4. Simulated Gain of the Proposed Antenna



farfield (f=28) [1]
Type Farfield
Approximation enabled (kR >> 1)
Component Abs
Output Directivity
Frequency 28 GHz
Rad. effic. -1.524 dB
Tot. effic. -1.555 dB
Dir. 8.195 dBi

Fig. 5. Simulated Directivity of the Proposed Antenna

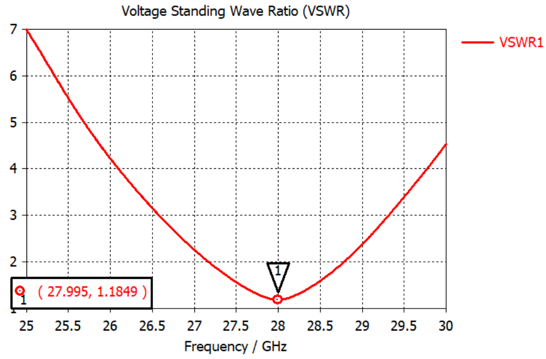
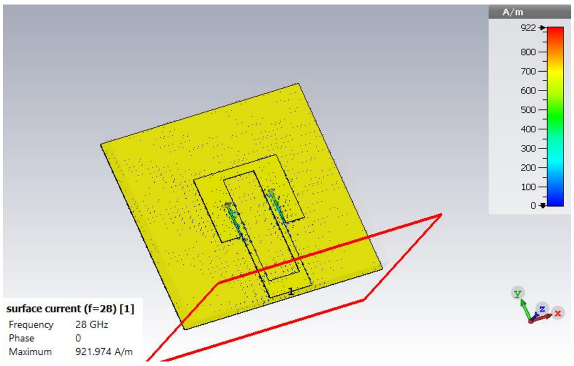


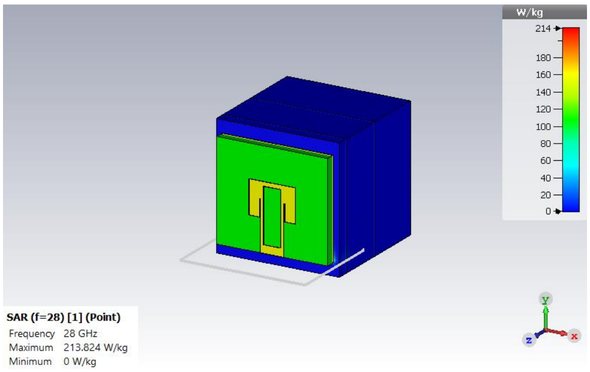
Fig .6. Simulated VSWR of the Proposed Antenna

Fig. 6 illustrates the simulated VSWR of the proposed 5G wearable antenna. It has a good VSWR value of 1.1849 at 27.995 GHz. Fig. 7 displays the surface current of the proposed 5G wearable antenna. From the simulation result, the proposed 5G wearable antenna has a surface current of 921.974 A/m surrounded at the feeding line.



surface current (f=28) [1]
Frequency 28 GHz
Phase 0
Maximum 921.974 A/m

Fig. 7. Surface Current of the Proposed Antenna



SAR (f=28) [1] (Point)
Frequency 28 GHz
Maximum 213.824 W/kg
Minimum 0 W/kg

Fig. 8. Simulated Point SAR of the Proposed Antenna

Fig. 8 shows the simulated point SAR of the proposed 5G wearable antenna. The proposed antenna has a point SAR of 213.824 W/kg at 27.995 GHz.

Table 4. Summary of Simulated Results of the Proposed Antenna

Antenna Parameter	Values
Operating Frequency (GHz)	27.995
Bandwidth (GHz)	1.517
Return Loss (dB)	-21.451
Gain (dB)	6.671
Directivity (dBi)	8.195
VSWR	1.1849
Surface Current (A/m)	921.974
Point SAR (W/kg)	213.824

Table 4 shows the summary of the simulated results of the proposed antenna. At the operating frequency of 28 GHz, a return loss of -21.451 dB is achieved with VSWR of 1.1849. In addition, the operating bandwidth for 28 GHz is 1.517 GHz between 27.221 GHz and 28.738 GHz with respect to -10 dB. The proposed antenna achieved with a gain of 6.671 dB, directivity of 8.195 dBi, and surface current of 921.974 A/m. SAR which is employed with wearable antenna, is used to measure how well electromagnetic waves are absorbed by human tissues. To assess he point SAR value of the 5G wearable antenna, a numerical human phantom model was created. The point SAR produced from the proposed antenna is 213.824 W/kg based on the results of the simulation.

Table 5. Comparison of Antenna Parameters between Proposed 5G Wearable Antenna with Other Existing Designs

Paper	Ahmed et.al.[2]	Vivek et.al.[9]	Proposed antenna
Substrate	Ultralam 3850HT	Polycarbonate	Rogers 3003
Relative Permittivity (Ɛr)	3.14	2.57	3
Operating Frequency (GHz)	38, 60	28	27.995
Bandwidth (GHz)	0.442, 0.657	1.49	1.517
Return Loss (dB)	-13.996, -19.268	-65.76	-21.451
Gain (dB)	2.19, 4.43	8.886	6.671
Directivity (dBi)	-	-	8.195
VSWR	-	1.001	1.185
Surface Current (A/m)	-	-	921.974
Point SAR (W/kg)	-	-	213.824
Average SAR (1g/10g) (W/kg)	0.33/0.15, 0.63/0.35	-	-

Table 5 shows the comparison of antenna parameters between the proposed antenna with other existing designs. In contrast to the design by Ahmed and Ahmed [2], they developed a dual-band flexible wearable antenna enabling modern 5G applications to be incorporated on a smartwatch. The substrate is made of Ultralam 3850HT, which has a low and steady relative permittivity of 3.14 and thin flexible cores. For wearable and high frequency designs, the steady dielectric constant is an important requirement. The frequency range of the mm-waves includes both 38 GHz and 60 GHz, and the antenna can operate at any one of these frequencies. As compared to this paper, the proposed antenna in this work has a greater bandwidth, as a broader bandwidth may effectively function across a wider range of frequencies. The antenna designed by Ahmed and Ahmed has a low average SAR value for SAR analysis of 0.33 W/kg and 0.63 W/kg averaged across 1g and 10g of tissues, respectively, at 38 GHz. As opposed to this, the proposed antenna in this work only has a point SAR value of 213.824 W/kg without averaging across the mass of tissues. For a different design by Vivek, Kumar, and Shambavi [9], they developed a polycarbonate substrate with a relative permittivity of 2.57, thickness of 0.5mm, and a loss tangent of 0.0069 for a 5G wearable antenna. Although the wearable antenna has a peak gain of 8.886 dB and a return loss of -65.76 dB, however the bandwidth is slightly less than the proposed antenna in this work. Moreover, the wearable antenna is lack of SAR analysis, which is important for determining how wearable antennas may affect human.

4. Conclusions

In this paper, a wearable mm-wave antenna with a semi-flexible Rogers Duroid RO3003 substrate is proposed for use in 5G applications. The antenna can operate at 28 GHz with a bandwidth of 1.517 GHz. Within the frequency band of operation, the antenna radiates in a directional pattern with a high gain of 6.671 dB and a high directivity of 8.195 dBi. In addition, the point SAR value on human tissue has been modelled. The proposed antenna in this work is suited to be a wearable antenna due to low return loss, a low point SAR value, compact size, light weight, and ease of manufacturing.

Acknowledgements

This work was supported by the Ministry of Higher Education Malaysia under the Fundamental Research Schemes with project codes of Proj-FRGS/1/2019/TK04/UCSI/02/1 and the UCSI University Research Excellence & Innovation Grant (REIG) with project code of REIG-FETBE-2022/038.

References

1. M. Ikram, K. Sultan, M. F. Lateef, and A. S. M. Alqadami, “A Road towards 6G Communication—A Review of 5G Antennas, Arrays, and Wearable Devices,” Electronics, vol. 11, no. 1, Art. no. 1, Jan. 2022.

2. M. I. Ahmed and M. F. Ahmed, “A Wearable Flexible antenna integrated on a Smart Watch for 5G Applications,” *J. Phys. Conf. Ser.*, vol. 1447, no. 1, p. 012005, Jan. 2020.
3. L. Bariah et al., “A Prospective Look: Key Enabling Technologies, Applications and Open Research Topics in 6G Networks.”, Sep. 09, 2020.
4. N. F. M. Aun, P. J. Soh, A. A. Al-Hadi, M. F. Jamlos, G. A. E. Vandenbosch, and D. Schreurs, “Revolutionizing Wearables for 5G: 5G Technologies: Recent Developments and Future Perspectives for Wearable Devices and Antennas,” *IEEE Microw. Mag.*, vol. 18, no. 3, pp. 108–124, May 2017.
5. S. N. Mahmood et al., “Recent Advances in Wearable Antenna Technologies: a Review,” *Prog. Electromagn. Res. B*, vol. 89, pp. 1–27, 2020.
6. R. Karim, A. Iftikhar, B. Ijaz, and I. Ben Mabrouk, “The Potentials, Challenges, and Future Directions of On-Chip-Antennas for Emerging Wireless Applications—A Comprehensive Survey,” *IEEE Access*, vol. 7, pp. 173897–173934, 2019.
7. T. Hamed and M. Maqsood, “SAR Calculation & Temperature Response of Human Body Exposure to Electromagnetic Radiations at 28, 40 and 60 GHz Mm-Wave Frequencies,” *Prog. Electromagn. Res. M*, vol. 73, pp. 47–59, 2018.
8. K. Zhao, S. Zhang, C.-Y. Chiu, Z. Ying, and S. He, “SAR study for smart watch applications,” in 2014 IEEE Antennas and Propagation Society International Symposium (APSURSI), Jul. 2014, pp. 1198–1199.
9. N. Vivek, S. K. B, and K. Shambavi, “Design of Wearable Antennas for 5G Applications,” *Int. J. Electr. Eng. Technol.*, vol. 12, no. 5, Jun. 2021.

Authors Introduction

Mr. Wai Kiat Wong

He is currently pursuing Bachelor of Engineering Communication and Electronics with Honours as final year student in Faculty of Engineering, Technology and Built Environment, UCSI University, Malaysia. His 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. 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. 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. Eryana Eiyda Hussin

She is a lecturer from the Department of Electrical and Electronics of UCSI University in Malaysia. She received her master’s degree from the Faculty of Electronics and Computer System, Universiti Teknikal Malaysia Melaka. She is currently persuing her study under Doctor of Philisophy in Electrical and Electronic Engineering in Universiti Teknology Petronas, Malaysia.

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.

Driver’s Fatigue Recognition Using Convolutional Neural Network Approach

Samer Abdullah Deeb Abbas[1], **Sew Sun Tiang**[1, *], **Wei Hong Lim**[1], **Li Sze Chow**[1], **Chin Hong Wong**[2,3]
([1]UCSI University, Malaysia) ([2]Fuzou University, China) ([3]Maynooth University, Ireland)
E-mail: 1001851168@ucsiuniversity.edu.my, tiangss@ucsiuniversity.edu.my, limwh@ucsiuniversity.edu.my, chowls@ucsiuniversity.edu.my, chinhong.wong@mu.ie
http://www.ucsiuniversity.edu.my/

Abstract

Drowsy driving is a serious issue that has been leaking in our communities since long time, the definition of drowsy driving is when the driver is not aware enough to proceed with driving the vehicle causing catastrophic accidents. Multiple methods were found to approach this complication across the years. Convolution Neural Network has approved to be a reliable approach to treat this issue by using face feature detection. In this paper, the effect of key parameters of the trained framework based on the driver’s fatigue recognition model are analyzed, and the accuracy of the driver’s fatigue recognition model is investigated, as well as a driver’s fatigue recognition is studied under different conditions using CNN. Transfer learning is used to develop a reliable method for detection, Mediapip Face Mesh model is used to extract the features from the face. MAR (Mouth Aspect Ratio) as well as EAR (Eyes Aspect Ratio) are obtained through the detection, these terms are responsible for detecting the eye and mouth closure ratio, the model has proved to work with accuracy of 98.3% and in different light conditions with accuracy of 94.7% outperforming several past models.

Keywords: Drowsy detection, Eye detection, Mouth Detection, Eye Aspect Ratio, Mouth Aspect Ratio.

1. Introduction

World Health Organization stated that, road accidents cost the lives of millions of individuals every year [1]. Statistics have shown that most deadly accidents are caused by driver drowsiness and carelessness. According to the American Automobile Association [2], drowsy drivers are responsible for 7% of all accidents and 21% of fatal traffic accidents. In another research done by the Foundation for Traffic Safety in 2017, 42.4% of drivers travel without getting at least one day of sleep or less than six hours of sleep in a normal week [3]. For the majority of individuals (87.9%), these difficulties are serious, and they perceive what they observe as improper behavior (95.2%). However, nearly one-third of ten drivers (30.8%) admit to driving when too fatigued to keep their eyes open in the previous months. As stated by the American Automobile Association's Road Safety Foundation, driver weariness is to blame for 16–21% of traffic accidents. The risk of a traffic collision induced by driver drowsiness is 46 times higher than while driving normally [3].

This paper aims to develop a CNN model that can detect fatigue drivers to decrease the accidents that may occur as a consequence of drowsy driving. In addition will aim to provide a CNN model which can be used in the industry, with few conditions to expand the field of analyzing face detection under different circumstances, looking into the fact that the day goes by, and the night will come, and according to most drowsy driving occur between midnight and 6 am [4]. Lastly key parameters of the trained model will be supervised to enhance the predication output of the model while keeping an eye on the accuracy to ensure high detection efficiency [5]. Studies have showed different approaches to resolve this issue and they can be categorized as shown in Table 1. When a driver is fatigued, the blinking frequency is dramatically increased compared to usual. Additionally, because the driver's head posture changes depending on whether he or she is awake or weary, the frequency with which the driver nods over some time may be utilized as a measure for assessing fatigue [6].

Table 1. Summary of the Past Approaches

Methods	Used by	Limitation
Physiological Measures	Detecting bioelectric signals using sensors mounted on the head [5].	Uncomfortable for the driver. Expensive equipment
Vehicle-Based Measures	Analyzing lane offset, SWA, and vehicle speed, among other vehicle based variables [6].	Affected by External Factors. Different driving styles affects accuracy
Behavioral Measures	Determining the driver's fatigue level based on the fluctuation of the driver's head, eyes, lips, and other characteristics [7].	Can be affected by surrounding environment.

Recently, deep learning systems, particularly those based on Convolutional Neural Networks (CNNs), have gained importance in addressing difficult categorization issues. Most of them are ground-breaking advances in a variety of Computer Vision tasks, including scene segmentation, emotion identification, object detection, and picture classification. The advancement of machine learning technology (particularly CNN) enables increased accuracy and performance. This technique applies to a broad number of applications, including driver awareness testing [8]. PERCLOS, MAR and EAR are presented as an index parameter for measuring exhaustion and developed a matching fatigue detection system, in which the closure of the eyelids and frequent yawing might partially represent the fatigue condition. The Carnegie Mellon Institute has frequently proved through experimentation and demonstration that the physical quantity "PERCLOS" indicates drowsiness. Additionally, pupil features, eye gaze direction, and blink frequency all contribute to the identification of exhaustion[8]. The ability of computer vision models to detect the existence of human faces in digital pictures means that the system can recognize the presence of a human face in an image or video and distinguish it from other things [9].

The system begins by analyzing the input image before identifying the face of the user, locating their eyes and mouth from there, and detecting the movement of their eyelids. In order to distinguish between voluntarily and involuntarily blinking, the eyelids are first detected. Then, the Eye Aspect Ratio is determined. Multiple (x, y) coordinates are used to represent each eye, beginning at the left corner and moving clockwise around the rest of the area [10]. Similar to how the eye aspect ratio is dependent on these variables to determine when the eye is open and closed. the mouth aspect ratio can be determined using the coordinated information from the area around the upper and lower lip. To create the detection system, OpenCV and Mediapipe face mesh libraries were used. The libraries are used to analyses photos and live stream video to recognize specific objects, including faces, hands, and feet.

2. Related Works

2.1. *Methodology*

The system architecture mainly outlines how data is used, how it is changed, and how the results are affected by these changes. The suggested method starts by taking a video frame with the system's camera as shown in Fig. 1.

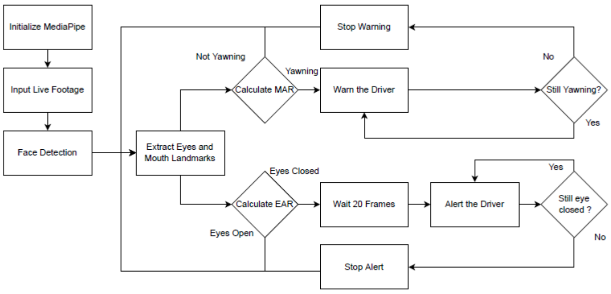


Fig. 1. Flowchart of the Fatigue Detection System.

2.2. *Frame Capturing*

OpenCV (Open Source Computer Vision Library) is a programming function library geared mostly at real-time computer vision. It's helpful in obtaining live footage through the camera, as well as its compatible with important libraries like Dlib and Mediapipe, also it is used for real-time activities like plotting the obtained data into a graph form. OpenCV is very crucial for the performance of the model, it's used to get the input live footage which will be processed to acquire the important terms related to the study like MAR and EAR. OpenCV also provides a way to monitor the MAR and EAR and make it user-friendly, by using a few lines of code EAR and MAR will be plotted into *x* and *y* graphs which makes it easier to determine when the driver hits the threshold and also to monitor the driver's behavior as well.

2.3. *Mediapipe Landmark*

Shape prediction includes identifying landmarks on the face. When an input image is given, the Shape Predictor seeks to locate the localization points around the structure, which is critical for a variety of facial analysis tasks. The Facial Landmark Algorithm is used to regionalize the face. In terms of facial landmarks, our goal is to identify the facial structures on the face using shape prediction methods. Using plot 468-point system[11] as shown in Fig. 2, the Mediapipe Face attributes are identified using the face mesh library [11]. In face landmarks, transfer learning is used to train a network with multiple objectives: the network simultaneously predicts 2D semantic contours on annotated real-world data and 3D landmark coordinates on artificially rendered data. The developed network allowed the models to predict 3D landmarks with high accuracy using both synthetic and real-world data.

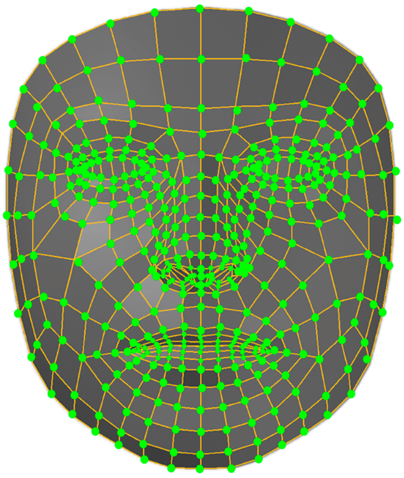


Fig. 2. The 468-point system.

2.4. *Face Detection, Eye and Mouth Key Points*

A facial landmark technique is used to identify a face inside a frame in order to detect faces. Only facial structures are shown, and all other foreign objects are disregarded by this algorithm. To complete this stage, Mediapipe Face Mesh is used to find and reflect notable facial characteristics. The model generates 468-xy points using the Mediapipe Face Mesh facial landmark detector. To label these 468 spots, the shape predictor approach is employed. Using these locations, the facial region can be identified. According to the 468 points surrounding the face, a placement of the points is computed in Fig. 3 as shown.

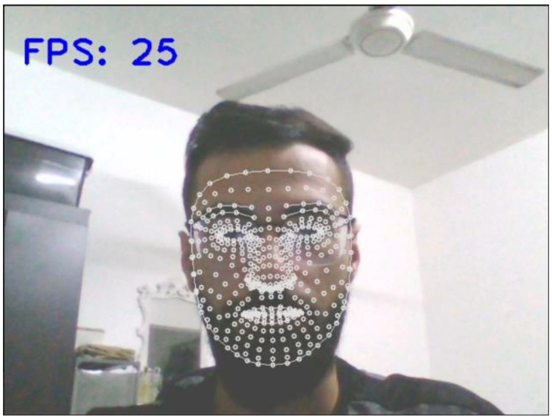


Fig. 3. Face detected using 468-point system.

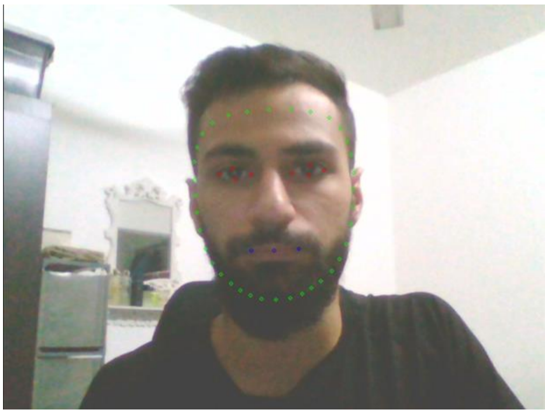


Fig. 4. Eye and Mouth detected.

An important step in this method is localization of the area around the lips and eyes. Eye and mouth areas were observed and used for eye and mouth monitoring and detection from the face which was captured. Based on the 468-point system the indexes for the left eye [362, 382, 381, 380, 374, 373, 390, 249, 263, 466, 388, 387, 386, 385,384, 398] and for the right eye [33, 7, 163, 144, 145, 153, 154, 155, 133, 173, 157, 158, 159, 160, 161, 246], and for lips [61, 146, 91, 181, 84, 17, 314, 405, 321, 375,291, 308, 324, 318, 402, 317, 14, 87, 178, 88, 95, 185, 40, 39, 37,0 ,267 ,269 ,270 ,409, 415, 310, 311, 312, 13, 82, 81, 42, 183, 78]. The convex hull computed for the right eye, left eye, and mouth is displayed in Fig. 4.

2.5. *Eye Aspect Ratio and Mouth Aspect Ratio*

Fig. 5 shows the eye vertical and horizontal coordinates. Once the left eye and right eye coordinates have been extracted, they are used to calculate the Eye aspect ratio based on the formula:

$$\delta = \frac{||E_2 - E_6|| + ||E_3 - E_5||}{2||E_1 - E_4||} \qquad (1)$$

where δ represent EAR and E_n represent the landmark on the eyelids and n represents the location of the landmark.

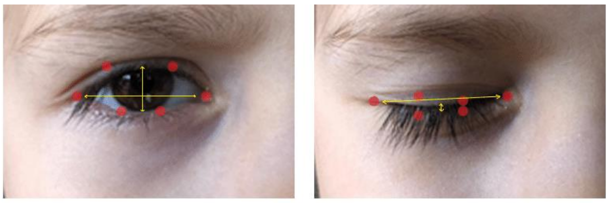


Fig. 5. Eye vertical and horizontal coordinates

This equation's denominator, which is weighted correctly because there is only one set of horizontal points but two sets of vertical points, computes the distance between horizontal eye landmarks while the numerator computes the distance between vertical eye landmarks. When a blink occurs, the eye aspect ratio will quickly drop below the threshold even if it is roughly constant when the eye is open. The ratio of eye landmark distances can be used to assess whether someone is blinking by using this straightforward equation. The average EAR is determined after the individual EARs for each eye are obtained; in this approach, the EAR threshold is set at 4.

Fig. 6 shows the mouth vertical and horizontal coordinates. The mouth aspect ratio can be calculated using the formula:

$$\delta = \frac{||P_2 - P_6|| + ||P_3 - P_5||}{2||P_1 - P_4||} \qquad (2)$$

where δ represent MAR and E_n represent the landmark on the lips and n represents the location of the landmark.

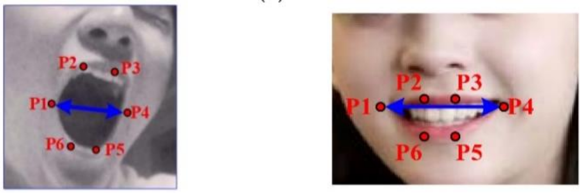


Fig. 6. Mouth vertical and horizontal coordinates

Similar to EAR, the ratio of mouth landmark distances can be used to assess whether someone is yawing by using this straightforward equation. The average MAR is determined after the individual MARs for mouth are obtained; in this approach, the MAR threshold is set at 1.8 if the ratio was below the threshold the individual will be considered yawing.

2.6. *Face Detection Training*

As transfer learning is used the model for detecting facial landmarks is a pre-trained that is able or detect 468 landmarks on the face. the model was trained using 17 evenly distributed samples (based on the United Nations geo-scheme) to perform a fairness evaluation [12]. Table 2 shows the dataset used for training the model and the distribution among 17 geographical regions.

Table 2. Dataset Used for Training

Dataset	Samples	Distribution
Dataset I	720 Samples	40 images per area plus 40 images without faces
Dataset II	800 Samples	400 Male + 400 Female
Dataset II	425 Samples	350 photos with one face on each, and 75 photographs for each category of skin tone

3. Results and Discussion

The distance between the camera and the user and the brightness of the room's lighting are the two factors that determine the system output. A laptop camera with a resolution of 640 x 480 at 30 frames per second is employed in this setup. The number of frames set for each warning is 20 frames, and these values are intended to prevent needless notifications when a user's eyes are naturally closing due to blinking, or any other facial emotion [13]. Fig. 7, Fig. 8, and Fig. 9 illustrate the angles which are best for detection. As a result, face roll and pitch (tilt) angles cannot deviate from the straight alignment by more than 45 degrees. Yaw (pan) angles must not be greater than 90 degrees. These angles are best fitted to detect the driver's behavior, while maintaining the range of these angles to ensure the accuracy increases.

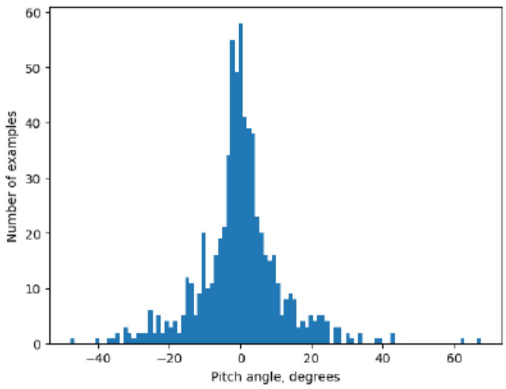


Fig. 7. Detection Pitch Angle

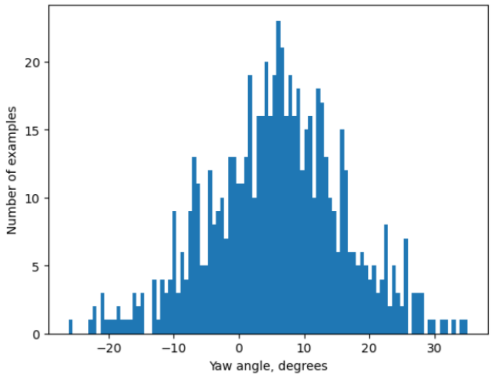


Fig. 8. Detection Yaw Angle

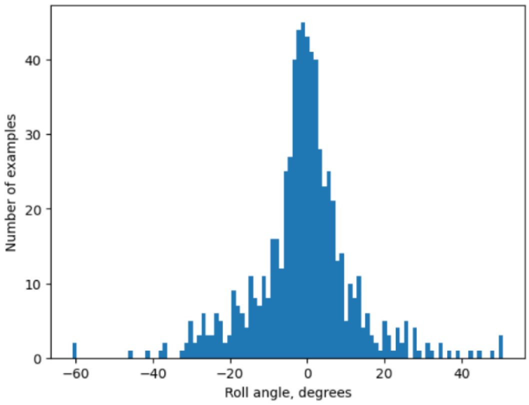


Fig. 9. Detection Roll Angle

The EAR was monitored to check the Eye closure, if the person's eyes closure exceeds the threshold the model will hold for 20 frames, and then alert the driver. After supervising the EAR, Fig. 10 shows the EAR value when the eyes are open, the values are ranged from 2.8 to 3.2, which considered as standard to open eyes. MAR was observed to check if the Mouth was considered yawning or not yawning, if the person's MAR ratio dropped below the threshold the model warns the driver and prompt the driver to take a rest. After supervising the MAR, when the mouth is closed the MAR value is undefined using the formula, however it can vary from a person to another due to different face features. Fig. 11 shows the values while the person is talking giving a range above 1.8.

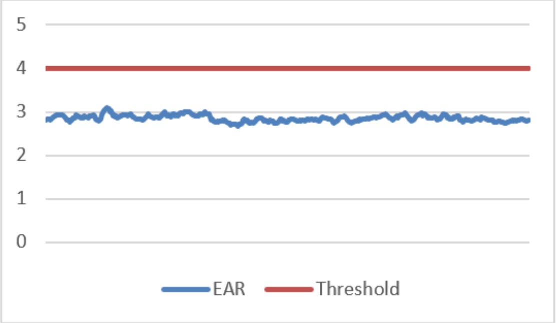


Fig. 10. Eye Aspect Ratio Monitoring

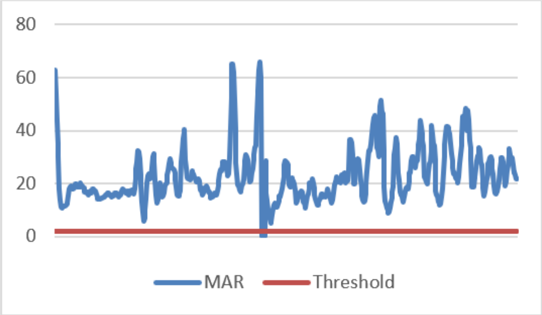


Fig. 11. Mouth Aspect Ratio Monitoring

The last stage is the communication between the system and the user. As the EAR and MAR computed, based on its results the system will voice out alerts to the user. Since the driver is drowsy sending out voice signals is very crucial to the state to warn the driver and the surrounding of the situation. The resulting eye or mouth reaction will be divided into 2 alerts once the driver yawn the system will send out warning by using pyttxs3 library which transforms writing into a speech sound, telling the driver to rest (Drowsy Warning … Take rest), the second alert is when the driver has closed his eyes for 20 frames counter then the system will voice out alerts to wake up the driver (Drowsy Alert … Wake up) the action will be repeated as long the driver is considered drowsy. As the day goes the model will adapt to the surrounding luminance around the driver, the accuracy of the detection is dropped to 91% at night due to the luminance while the accuracy will increase to 98.3% in good light condition. Table 3 displays the system's detection accuracy results when the ambient light level is 300 lux or more, which indicates that the interior of the automobile is sufficiently lit. A lux meter is used to evaluate the brightness of the surrounding environment. According to the results, accuracy is highest between 80 and 100 cm, then decreases as distance rises. 94.33% accuracy as the average detection is achieved through different lux settings. Fig. 12, Fig. 13, and Fig. 14 show the detection of the model tested on a different set of light.

Table 3. Detection Accuracy with Different Light and Distance

Luminance in Lux	Accuracy
> 300 lx	98.3%
100 > lx >300	95%
< 100 lx	91%

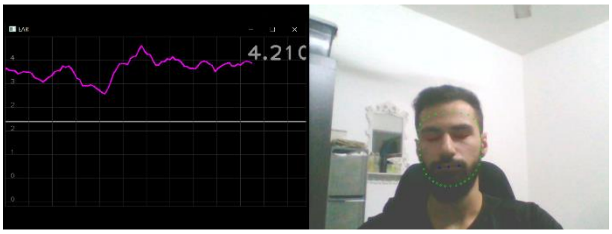


Fig. 12. Detection in > 300 lx.

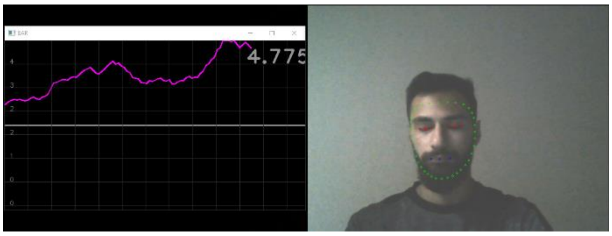


Fig. 13. Detection in 100 > lx > 300.

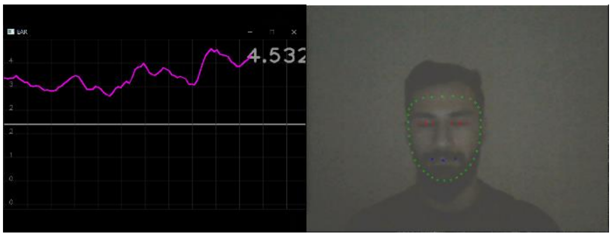


Fig. 14. Detection in < 100 lx.

A confusion matrix is a special table structure that enables visualization of the performance of an algorithm. It is sometimes referred to as an error matrix. In Fig. 15 and Fig. 16 illustrate the eye confusion matrix results. Table 4 shows the evaluation metrics of the model. Table 5 shows the results compared between our model, and a model trained using YawDD dataset [13] to detect mouth and eyes. Our model has the highest set of accuracy and precision; however, our model is considered the least among the other models if comparing the recall score.

Table 4. Evaluation Model Results

Class	Accuracy	Precision	Recall	F1 Score
Eyes	99%	99.5%	77.4%	87%
Mouth	97.7%	98.6%	87.6%	92.7%

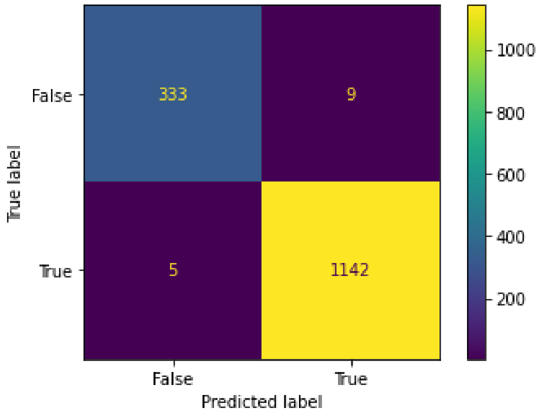


Fig. 15. Eyes Confusion Matrix

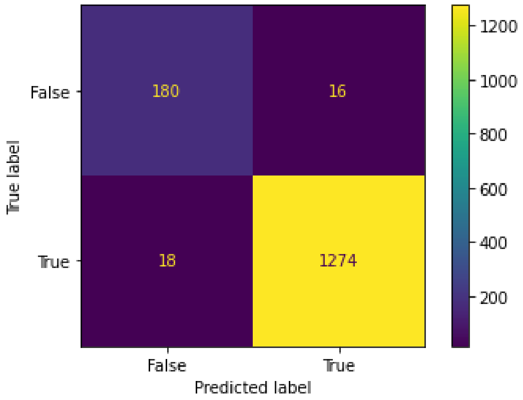


Fig. 16. Mouth Confusion Matrix

Table 5. Comparing Detection in Frame Per Second.

Models	Accuracy	Recall	Precision	F1-Score
[13]	95%	95%	95%	94.5%
Trained YOLOv5 Model	89.7%	93.7%	92.2%	92.9%
Our Model	98.3%	82.5%	99%	89.85%

Table 6 shows the comparison between other models, while Mohammad's [12] and Jonathan's [13] model achieved higher accuracy, our model has taken into accountability different environments which affects the accuracy. Table 7 compares between the detection in real-time, as shown our model did not have the highest fps between other models, because the camera used to test the model were limited to 30 – 40 fps. Better camera can achieve up 275 fps which makes the model significantly faster in detection than other models.

Table 6. Comparison with other methods

Models	Method	Accuracy
Walizad, et. al. [13]	CNN	95%
Monroy, et. al. [14]	SS-CNN	98.95%
Our Model	Mediapipe, OpenCV	98.3%

Table 7. Comparing Detection in Frame Per Second.

Authors	Detection Per FPS
Liu, et. al. [15]	60 FPS
Li, et. al. [16]	58 FPS
You, et. al. [17]	20 FPS
Our Model	30 FPS

Based on the mentioned data, the system's overall detection accuracy is 98.%. taken into consideration different light condition the accuracy decreases to 94.73%. the model did not have a high accuracy detecting glasses due to the reflection of light. In addition, better camera resolution can aid when the room's brightness is less than 300 lx. and a camera with higher frame rates will be able to detect the eye closure and mouth yawing much quicker up to 275 FPS as mentioned by the authors of MediaPipe model, producing better results with greater precision.

4. Conclusions

In this paper, OpenCV and Mediapipe are used to create a system for detecting drowsy driving. The system's goal is to provide the drivers with a constant warning for their behavior in on spot situation. The method is built around the detection of faces using a 468-point system and the detection of eyes and mouth using the eye aspect ratio and mouth aspect ratio. Results from the system revealed an average 90% percent accuracy in detection. A camera with faster frame rates and sharper quality can increase the accuracy percentage.

Acknowledgements

This work was supported by the Ministry of Higher Education Malaysia under the Fundamental Research Schemes with project codes of Proj-FRGS/1/2019/TK04/UCSI/02/1 and the UCSI University Research Excellence & Innovation Grant (REIG) with project code of REIG-FETBE-2022/038.

References

1. WHO, “Road traffic injuries,” 2022. https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries.
2. AAA, “Drowsy Driving: Don’t Be Asleep at the Wheel,” 2018, [Online].
3. Foundation of Traffic Safety, “Prevalence of Drowsy Driving Crashes: Estimates from a Large-Scale Naturalistic Driving Study,” 2018.
4. NHTSA, “Drowsy Driving,” 2022. https://www.nhtsa.gov/risky-driving/drowsy-driving.
5. X. Hu and G. Lodewijks, “Detecting fatigue in car drivers and aircraft pilots by using non-invasive measures: The value of differentiation of sleepiness and mental fatigue,” *J. Safety Res.*, vol. 72, pp. 173–187, Feb. 2020.
6. C. C. Ukwuoma and C. Bo, “Deep Learning Review on Drivers Drowsiness Detection,” 2019.
7. F. Liu, X. Li, T. Lv, and F. Xu, “A Review of Driver Fatigue Detection: Progress and Prospect,” 2019.
8. M. Ngxande, J. R. Tapamo, and M. Burke, “Driver drowsiness detection using behavioral measures and machine learning techniques: A review of state-of-art techniques,” in *2017 Pattern Recognition Association of South Africa and Robotics and Mechatronics International Conference*, PRASA-RobMech 2017, 2017, vol. 2018-Janua, pp. 156–161.
9. A. Kost, W. A. Altabey, M. Noori, and T. Awad, “Applying neural networks for tire pressure monitoring systems,” *SDHM Struct. Durab. Heal. Monit.*, vol. 13, no. 3, pp. 247–266, 2019.
10. Z. Li, S. E. Li, R. Li, B. Cheng, and J. Shi, “Online detection of driver fatigue using steering wheel angles for real driving conditions,” *Sensors (Switzerland)*, vol. 17, no. 3, Mar. 2017.
11. D. R. Carvalho, W. D. A. Fonseca, J. Hollebon, P. H. Mareze, and F. M. Fazi, “Head tracker using webcam for auralization,” *Proc. INTER-NOISE 2021 - 2021 Int. Congr. Expo. Noise Control Eng.*, no. August, 2021.
12. V. Bazarevsky, Y. Kartynnik, A. Vakunov, K. Raveendran, and M. Grundmann, “BlazeFace: Sub-millisecond Neural Face Detection on Mobile GPUs,” pp. 3–6, 2019, [Online].
13. M. Elham Walizad, M. Hurroo, and D. Sethia, “Driver Drowsiness Detection System using Convolutional Neural Network,” *2022 6th Int. Conf. Trends Electron. Informatics, ICOEI 2022 - Proc.*, no. Icoei, pp. 1073–1080, 2022.
14. J. Flores-Monroy, M. Nakano-Miyatake, G. Sanchez-Perez, and H. Perez-Meana, “Visual-based Real Time Driver Drowsiness Detection System Using CNN,” *CCE 2021 - 2021 18th Int. Conf. Electr. Eng. Comput. Sci. Autom. Control*, pp. 18–22, 2021.
15. M. Z. Liu, X. Xu, J. Hu, and Q. N. Jiang, “Real time detection of driver fatigue based on CNN-LSTM,” *IET Image Process.*, vol. 16, no. 2, pp. 576–595, Feb. 2022.
16. X. Li, J. Xia, L. Cao, G. Zhang, and X. Feng, “Driver fatigue detection based on convolutional neural network and face alignment for edge computing device,” *Proc. Inst. Mech. Eng. Part D J. Automob. Eng.*, vol. 235, no. 10–11, pp. 2699–2711, Sep. 2021.
17. F. You, X. Li, Y. Gong, H. Wang, and H. Li, “A Real-time Driving Drowsiness Detection Algorithm with Individual Differences Consideration,” *IEEE Access*, vol. 7, pp. 179396–179408, 2019.

Authors Introduction

Mr. Samer Abdullah Deeb Abbas

He is currently pursuing Bachelor of Engineering Electrical and Electronics with Honours as final year student in Faculty of Engineering at UCSI University, Malaysia. His research interests are neural network 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. 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. Li Sze Chow

She is an Assistant Professor in Faculty of Engineering at UCSI University in Malaysia. lecturer She has received Ph.D degree in Electrical and Electronic Engineering from The University of Sheffield, UL in 2006. Her research interests are detection in brain waves 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.

Deep Residual Neural Network for Efficient Traffic Sign Detection

Hanlin Cai[1], Zheng Li[1], Jiaqi Hu[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: hanlin.cai@ieee.org, zheng.li.2021@mumail.ie, jiaqi.hu.2021@mumail.ie, limwh@ucsiuniversity.edu.my, tiangss@ucsiuniversity.edu.my, mastaneh@ucsiuniversity.edu.my, chinhong.wong@mu.ie

Abstract

This paper established three deep residual neural network models with different architectures for traffic sign detection. Also, a new systematic analytic hierarchy process method for model performance evaluation has been proposed, which was utilized to determine the configuration of the deep learning model. In this paper, four evaluation metrics were used for analytic hierarchy process measurement, they are accuracy, stability, response time, and system capability. Based on the Tsinghua-Tencent 100K dataset, experimental results verified the feasibility of the proposed models for traffic sign detection and recognition which has training and testing accuracy of 99.03% and 98.01% respectively.

Keywords: Traffic Sign Detection System; Residual Neural Network (RNN); Analytic Hierarchy Process (AHP).

1. Introduction

Deep Learning (DL) plays a non-negligible role in current frontier science, which has been widely used in agriculture, and transportation industries. As for the application in urban transportation, it is of great significance to utilise related DL approaches for the Traffic Signs Detection System (TSDS), which normally consists of two related domains: Traffic Sign Detection (TSD) and Traffic Sign Recognition (TSR). However, TSDS requires high accuracy and precision while exploiting the shortest possible detection and recognition time. As an alternative to conventional machine learning schemes, deep learning-based schemes appear to be a promising option for efficient traffic sign detection[1],[2],[3]. According to recent literature works, to address the challenge of traffic sign detection and recognition, Suriya Prakash, *et al.*[4] proposed a LeNet-5 Convolutional Neural Network (CNN) model that possessed a high detection accuracy of nearly 98.8%. Changzhen, *et al.*[5] implemented an advanced detection method based on a deep CNN model which also achieved a satisfactory result of above 99.0% recognition precision.

Despite the good performance of deep learning models, the effectiveness of this CNN model will decrease when facing trickier image recognition challenges which require deeper layers and more computing resources [3]. The Residual Neural Network (RNN) or in short, ResNet, approach was proposed by He, *et al.*[6] to resolve the problem that the performance decreases with the deepening of network training. RNN models adopt a residual learning methodology that significantly reduces the difficulty of the deep networks training process. Besides, RNN models have been widely applied in some research works, Zakaria, *et al.*[7] utilized RNN models in the medical field to recognize and classify medical images. Li and Raim made optimization and improvement combined with the actual applications, which obtained very good results of over 98.2% accuracy for fruit leaves detection and recognition.

This paper applies RNN approaches to explore state-of-the-art solutions for efficient traffic sign detection. In this work, a deep RNN is built to address the TSDS challenges. Experiments have been conducted to verify the feasibility of implementing the RNN model for TSD and TSR problems. Also, another contribution of this paper is to propose a new performance evaluation method for a RNN with different parameters and optimizers. An optimal configuration scheme for RNN models was suggested through a large number of experiments based on representative datasets.

2. Related Works

This section goes through the key concepts of this paper, including TSDS and DL. Furthermore, many related research works have been explicitly reviewed, and the gaps in existing knowledge have been identified.

2.1. *Traffic Sign Detection System (TSDS)*

Traffic signs provide paramount information for real-world driving, and a variety of methods and algorithms have been implemented to detect and recognize different traffic signs in different countries and regions. The TSDS concern two related subjects: TSD and TSR, where TSD aims to find an accurate location of the sign in the physical transportation environment and TSR mainly focus on identifying the meaning of specific traffic signs (e.g., Speed Limit, Stop and Direction). As for existing knowledge in the area of TSDS, Lu, *et al.*[9] Wali, *et al.*[10] and Arcos-García, *et al.*[11] have proposed comprehensive surveys of some state-of-the-art techniques for TSDS purposes.

As shown in Table 1, many related works have been explicitly reviewed in this paper. As for the conventional methods for TSDS, most of the research works focus on the methods of colour segmentation, image shape and texture features[12],[13],[14],[15],[16],[17]. However, these traditional approaches are highly dependent on the quality of the images, which can be easily affected by daylight conditions and the reaction of the paint to the pollutants in the air. Fleyeh and Dougherty[1] proposed an exhaustive overview of the traditional methods and pointed to many problems regarding traditional image detection and recognition methods. Considering the validity of the TSDS, most of the conventional methods have been gradually replaced by new learning-based models, which can optimize model performance and effectiveness through learning the existing datasets and previous experience.

Over the past two decades, many learning-based approaches have been proposed to address complex TSDS problems. Support Vector Machine (SVM) models have been applied in Spanish TSDS to provide alerts to the drivers[18]. Neural network models also gained extensive attention in this domain, which can be combined with Hough transformation, corner detection and projection methods. The Neural Network (NN) models proposed by Kuo and Lin[19] have achieved good accuracy of nearly 95.5% based on the traffic sign datasets in Taiwan, China. However, since the emergence of trickier traffic scenarios and the increase of different signal categories, general ML models get exhausted when facing more complicated challenges, such as contaminated, multi-object and large-scale sign detection and recognition[9].

2.2. *Deep Learning Technique*

DL technique has been the core topic in computer vision, which has been highly applied in image detection and classification[2]. CNN and RNN models are the most prominent DL approaches in the field of traffic sign detection and recognition.

Suriya Prakash, *et al.*[4] extended and developed a classical LeNet-5 CNN model, which makes use of Gabor based kernel followed by a normal convolutional kernel after the pooling layer. Their proposed CNN model was evaluated using the German Traffic Sign Benchmark and gave an accuracy of nearly 98.9%.

Also, Changzhen, *et al.*[5] suggested a new algorithm based on deep CNN using Region Proposal Network (RPN) to detect all Chinese traffic signs. Experiments show that their model has real-time detection speed and above 99.0% precision.

Considering better detection response time, K R, *et al.*[20] have proposed a combined scheme utilizing Faster Region-based Convolution Neural Network (RCNN) and RPN network. Besides, the Random Forest algorithm is used to perform classification and regression in the given dataset. Their composite methods significantly reduced the resource requirements used for training the deep learning models and the accuracy increased up to 99.9%.

However, most of the existing methods suggested are based on a limited number of traffic signs (about 50 classes out of several hundred in different regions). Tabernik and Skocaj[21] proposed several improvements using CNN and mask R-CNN approach to resolve the issue of detecting large-scale traffic sign categories. The experiments are conducted on large-scale traffic signs detection and results show that the detection has a 2–3% average error rate in actual detections.

Table 1 A summary of related literature works

Techniques	Descriptions
Colour Segmentation [12][13]	Easily affected by daylight conditions.
Texture Features [16][17]	Highly depending on the quality of the images.
SVM Classifier [18]	Good classification accuracy, but low speed.
NN Models [19]	High accuracy, but a large resource is required.
LeNet-5 CNN [4]	Utilizing Gabor Based Kernel, high accuracy
CNN+RPN [5]	Very high real-time detection speed.
R-CCN+RPN [20]	Very high accuracy, close to 99.9%.
Mask R-CNN [21]	Based on highly challenging datasets.

2.3. *Research Gap*

To date, there is limited research that has focused on training and testing the RNN model based on representative traffic sign datasets including large-scale categories. In this paper, a deep RNN is proposed to address the large-scale detection challenges. Experiments are conducted to verify the effectiveness of implementing the RNN model to advance TSDS.

Although the existing research works have achieved a good detection result, most of the works are only evaluated by detection accuracy, precision, and response time, which are all in the same key. In this paper, a new analytic hierarchy process (AHP) method for RNN with different parameters and optimizers is proposed to deploy in the practical performance measurement of the deep learning model.

3. The Proposed Method

This work utilizes the Tsinghua-Tencent 100k dataset proposed in [22], which provides 100k images containing 30k different traffic signs. The RNN with respectively 50, 101, and 152 layers of CNN-based architecture has been established for training and testing. Fig. 1 illustrates the workflow of the model training and testing process. The training dataset is composed of 80% of the total dataset, which is used to train the proposed RNN model. While the testing dataset consists of the remaining 20%, which is used to evaluate the performance of the trained RNN model.

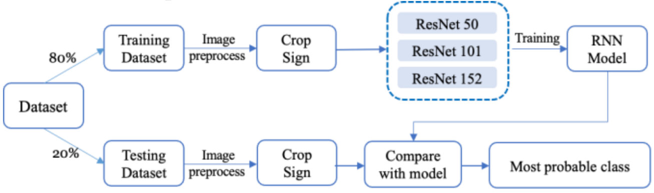

80%
Training Dataset
Image preprocess
Crop Sign
ResNet 50
ResNet 101
ResNet 152
Training
RNN Model
Dataset
20%
Testing Dataset
Image preprocess
Crop Sign
Compare with model
Most probable class

Fig. 1. Workflow of model training and testing process

3.1. *Image Pre-processing*

The first step of the TSDS system workflow is image pre-processing. Considering that a collection of pixels in an image with a sharp change in brightness is often the outline of an object. Being able to locate them accurately means that the actual signs can be located and predicted[21]. To extract the features in the image, this paper mainly utilized edge detection and corrosion expansion. Fig. 2 shows the specific workflow of the image pre-processing, including noise processing, gradient computation, non-maximum suppression, double threshold detection and corrosion expansion.

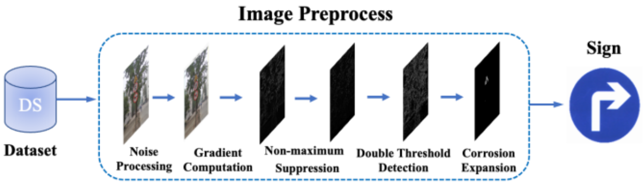


Image Preprocess
DS
Dataset
Noise Processing
Gradient Computation
Non-maximum Suppression
Double Threshold Detection
Corrosion Expansion
Sign

Fig. 2. Flowchart of the image pre-processing

Edge Detection

Noise processing and gradient computation

In the image, the pixels with large grayscale changes appear randomly and gene the noise. And the noise comes from various situations, including image acquisition, transmission, and quantization. In order to minimize the noise, this paper applied the Gaussian filter to minimise the noise. The gaussian filter will give a negative impact on the definition and sharpness of images, therefore, the gradient computation is conducted to extract the outline of the objects in the image. These two steps make the features of the image more prominent and suitable for subsequent processing.

Non-maximum suppression

After the gradient computation, the features of the main objects in the images will be highlighted. In this case, non-maximal value suppression is utilized to eliminate the influence of other non-target objects. It suppresses all the gradient values other than the local maximum and indicates the position of the strongest strength of the image. In the location of the traffic sign detection, a large number of traffic signs in the location of the same target will be detected, and these traffic signs may overlap between them. The non-maximal values can be used to prevent the finding of the best target border frame and eliminate the redundant boundary boxes.

Double threshold detection

The final step of edge detection is to distinguish the objects from the background. The grey difference between the target and the corresponding background in the image is used to extract the targeted traffic signs and divide the pixel level into several classes by setting the threshold in order to achieve the separation of the objects and the background.

Corrosion expansion

After the edge detection, the next procedure is corrosion expansion, which selects the maximum value in the neighbourhood of each position as the output grey value. After expansion, the overall brightness of the image will be improved. The size of the brighter object in the graph will be larger, while the size of the darker object will be reduced or even disappear. According to the result of image processing and the location determined by the colour threshold, the traffic sign without background noise is obtained.

3.2. *Residual Neural Network (RNN)*

The proposed RNN architecture has two layers, called *conv* block and identity block, which serve as shortcuts in residual blocks and are included in an order, as shown in Fig.3.

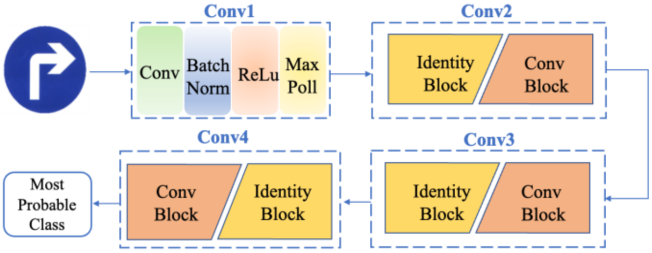

Conv1
Conv
Batch Norm
ReLu
Max Poll
Conv2
Identity Block
Conv Block
Conv3
Identity Block
Conv Block
Conv4
Conv Block
Identity Block
Most Probable Class

Fig. 3. Architecture of proposed ResNet

Fig. 4 presents the structures of the identity and *conv* blocks respectively. A stack of three layers was used for each residual block. The 1×1, 3×3, and 1×1 layers are three convolution layers. The 1×1 layers focus on first reducing and then increasing the dimensions, and the 3×3 layer has smaller input and output dimensions.

4. Analytic Hierarchy Process

To accurately measure the maturity level of different process parts, the Analytic Hierarchy Process (AHP) is proposed to establish a maturity evaluation model. AHP applies simple mathematical tools combined with operational thoughts to decompose complex issues into individual constituents, and form hierarchies according to the disposable relationship group.

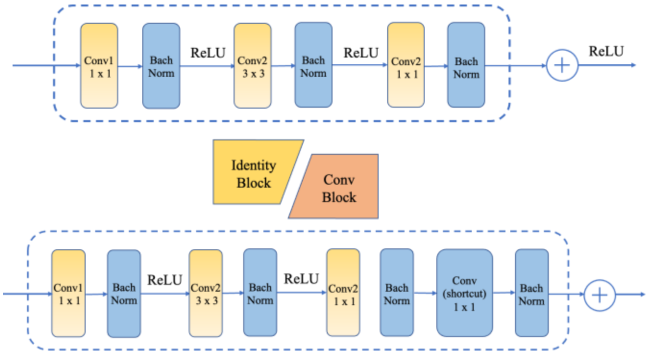


Fig. 4. Structures of the identity and *conv* blocks

The key mathematical notations used in this sector are listed in Table 2.

Table 2. Notations of the AHP method

Symbol	Description	Unit
Score	Total score	1
RT	Response time score	1
SB	Stability score	1
AC	Accuracy score	1
SC	System capability score	1
PS	Process Speed	Piture/ms
CPR	Computing Requirement	FLOPS

4.1. *Maturity level evaluation model*

Indicator selection

This paper designs four sub-indicators to calculate the maturity level score of different parts. These four sub-indicators are Accuracy (*AC*), Stability (*SB*), Response Time (*RT*) and System Capability (*SC*).

Accuracy (AC)

The most important standard to judge the model is the accuracy. Therefore, the ratio between the number of correct detection and the total set is used to normalize the performances of different scale pictures. The corresponding score in Table 3 is calculated as follows:

$$AC = \frac{Correct\ number}{Total\ number} \times 100\% \tag{1}$$

Table 3 Accuracy score table

Accuracy	*AC (Score)*
Less than 0.75	0
0.75-0.80	1
0.80-0.85	2
0.85-0.90	3
0.90-0.95	4
More than 0.95	5

Stability (SB)

The model must keep relatively high stability in all working conditions. But unexpected things would disturb the detection process in many realistic cases, like dim environments, incomplete pictures, and broken traffic signs. So, the ratio between the accuracy of realistic and theoretical conditions is considered to obtain the *SB*. The *SB* score in Table 4 is calculated using:

$$SB = \frac{Practical\ accuracy}{Theatrical\ accuracy} \times 100\% \tag{2}$$

Table 4 Stability score table

Stability	*SB (Score)*
Less than 75%	0
75%-80%	1
80%-85%	2
85%-90%	3
90%-95%	4
95%-100%	5

Process Speed (PS) & Response Time (RT)

To guarantee the efficiency of the process parts, the response time is a very important factor. The training speed is determined by the time of processing 1000 images using eq. 3. The shorter the training time, the higher the score, as shown in Table 5.

$$PS = \frac{Response\ time}{1000} \tag{3}$$

Table 5 Response time score table

Process Speed	*RT (Score)*
More than 5s	0
4-5s	1
3-4s	2
2-3s	3
1-2s	4
0-1s	5

System capability (SC)

Hardware requirements are essential factors that limit the performance of the model. The requirement is determined by computing power requirement (CPR). Floating-point operations per second (FLOPS) is often used to estimate the performance of a computer, especially in scientific computing where many floating-point arithmetic is used. Therefore, this paper also adopts FLOPS to evaluate the CPR level, as shown in Table 6.

Table 6 System capability score table

CPR	*SC (Score)*
More than 3.5G	0
2.0-3.5G	1
1.5-2.0G	2
1.0-1.5G	3
0.5-1.0G	4
0-0.5G	5

4.2. *Weight Determination*

In this work, the judgment matrix of the indicators is constructed according to the nine-point scale, as shown in Table 7, to compare the five sub-indicators in the scores pairs to obtain the judgment matrix. Among them, the elements in the matrix should satisfy:

$$x_{ij} = \frac{1}{x_{ji}}, (i, j = 1,2,3,4) \tag{4}$$

Table 7 Nine-point table

Scaling	Definition
1	Factor i is as important as factor j
3	Factor i is slightly more important than factor j
5	Factor i is significantly more important than factor j
7	Factor i is much more important than factor j
9	Factor i is extremely more important than factor j
2,4,6,8	The scale value of the importance of factor i over factor j is between the above two adjacent levels
Reciprocal of scaling value	Inverse comparison of factor i and factor j: $x_{ij} = 1/x_{ji}$

The weight vector can be obtained by the arithmetic mean below

$$\omega_{1i} = \frac{1}{n}\sum_{j=1}^{5} \frac{x_{ij}}{\sum_{k=1}^{5} x_{ki}} (i = 1,2,3,4) \tag{5}$$

And the geometric mean method to find the weight vector is

$$\omega_{2i} = \frac{\left(\prod_{j=1}^{5} x_{ij}\right)^{\frac{1}{5}}}{\sum_{k=1}^{5} \left(\prod_{j=1}^{5} x_{ij}\right)^{\frac{1}{5}}}, (i = 1,2,3,4) \tag{6}$$

The arithmetic mean and the geometric mean are used to explore the weight vector, ω_{1i}, ω_{2i} $(i = 1,2,3,4)$, then average them to obtain the weight vector. It can be seen from Table 8 that the weight of *RT*, *SB*, *AC*, and *SC* are similar, and *AC* has the largest weight on $Score$.

The model must guarantee the accuracy of traffic sign detection. The detection must get the command during the right driving time. The consistency test of the judgment matrix is less than 0.1, indicating that the weight data obtained are valid. The evaluation model is shown as follows:

$$\begin{aligned} Score &= 0.275RT + 0.211SB \\ &\quad + 0.342AC + 0.172SC \end{aligned} \tag{7}$$

Table 8 Weight bar graph

Indicator	*Weight*
RT	0.275
SB	0.211
AC	0.342
SC	0.172
RT	0.275

5. Experiment and Analysis

This section describes the results of experiments using selected models and the application of the analysis of the results with the performance evaluation indicators presented in this paper.

5.1. *Basic result analysis*

The experiment result of training and testing is usually assessed by metrics derived from the confusion matrix, as shown in Table 9.

Table 9 Confusion matrix for performance evaluation

Input Image	Positive Predictive	Negative Predictive
Positive Sample	True Negative (TN)	False Negative (FN)
Negative Sample	False Positive (FP)	True Positive (TP)

In order to avoid a biased analysis, credible metrics namely False Alarm Rate (FAR) and Un-Detection Rate (UND) metrics were used to evaluate the performance. Table 10 illustrates several evaluation metrics and their corresponding formulas.

Table 10 Evaluation metrics and explanations

Evaluation Metrics	Corresponding Formula	
Accuracy	$\frac{1}{n}\sum_{1}^{n}\frac{TP+TN}{TP+TN+FP+FN}\times 100\%$	(8)
False Alarm Rate (FAR)	$\frac{1}{n}\sum_{1}^{n}\frac{FP}{FP+TN}\times 100\%$	(9)
Un-Detection Rate (UND)	$\frac{1}{n}\sum_{1}^{n}\frac{FN}{FN+TP}\times 100\%$	(10)

5.2. *Result Analysis*

This paper provides a unique performance evaluation standard by combining practical application scenarios and various factors. For the effects between different parameters in the same case, the comparison results were tabulated in Table 11.

Furthermore, this paper selects the RNN model with the best execution effect as the training model. By comparing the results of different parameters in the training process, the optimal parameters were listed in Table 12. The RNN model has the highest training and testing accuracy which are 99.03% and 98.01% respectively.

6. Conclusion

In this paper, the RNN with different architectures has been established for resolving traffic signal detection challenges. Also, this paper designed a new AHP method to evaluate the performance of the proposed RNN models with different parameters, so that the configuration of the start-of-the-art model can be determined. Four evaluation metrics were utilized for AHP measurement: accuracy, stability, response time and system capability.

Based on the Tsinghua-Tencent 100k dataset, the training and testing of the proposed models were conducted and analysed. The RNN model has the highest training and testing accuracy which are 99.03% and 98.01% respectively. Experimental results verified the feasibility of the proposed model for traffic sign detection and recognition.

Table 11 Experimental results of training and testing

Evaluation Metrics		**Deep Learning Models**		
		VGG	**GoogleNet**	**RNN**
Training (80%)	**Accuracy**	98.24%	98.89%	**99.03%**
	FAR	0.06%	0.03%	**0.01%**
	UND	0.87%	0.86%	**0.41%**
Testing (20%)	**Accuracy**	83.60%	96.62%	**98.01%**
	FAR	2.47%	0.19%	**0.09%**
	UND	56.73%	2.94%	**1.28%**

Table 12 Parameter list of the AHP method

Training Parameters	**Parameter 1**	**Parameter 2**	**Parameter 3**
Convolution Layers	**RNN 50**	RNN 101	RNN 152
Learning Rate	**Step**	Low	High
Split Strategy	**Classification Split**	Random Split	
Image Enhancement	**Brightness Enhancing**	**Image Scale**	Contrast Enhancing
Colour Processing	**HSV**	RGB	

References

1 H. Fleyeh and M. Dougherty, "Road and traffic sign detection and recognition," in Proceedings of the 16th Mini-EURO Conference and 10th Meeting of EWGT, 2005, pp. 644-653.

2 M. Pak and S. Kim, "A review of deep learning in image recognition," presented at the 2017 4th international conference on computer applications and information processing technology (CAIPT), 2017.

3 V. Feng. "An overview of resnet and its variants." (accessed on Jan 6, 2023).

4 A. Suriya Prakash, D. Vigneshwaran, R. Seenivasaga Ayyalu, and S. Jayanthi Sree, "Traffic Sign Recognition using Deeplearning for Autonomous Driverless Vehicles," presented at the 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), 2021.

5 X. Changzhen, W. Cong, M. Weixin, and S. Yanmei, "A traffic sign detection algorithm based on deep convolutional neural network," in 2016 IEEE International Conference on Signal and Image Processing (ICSIP), 2016: IEEE, pp. 676-679.

6 K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770-778.

7 N. Zakaria, F. Mohamed, R. Abdelghani, and K. Sundaraj, "Three ResNet Deep Learning Architectures Applied in Pulmonary Pathologies Classification," presented at the 2021 International Conference on Artificial Intelligence for Cyber Security Systems and Privacy (AI-CSP), 2021.

8 X. Li and L. Rai, "Apple Leaf Disease Identification and Classification using ResNet Models," presented at the 2020 IEEE 3rd International Conference on Electronic Information and Communication Technology (ICEICT), 2020.

9 Y. Lu, J. Lu, S. Zhang, and P. Hall, "Traffic signal detection and classification in street views using an attention model," Computational Visual Media, vol. 4, no. 3, pp. 253-266, 2018, doi: 10.1007/s41095-018-0116-x.

10 S. B. Wali *et al.*, "Vision-Based Traffic Sign Detection and Recognition Systems: Current Trends and Challenges," Sensors (Basel), vol. 19, no. 9, May 6 2019, doi: 10.3390/s19092093.

11 Á. Arcos-García, J. A. Álvarez-García, and L. M. Soria-Morillo, "Evaluation of deep neural networks for traffic sign detection systems," Neurocomputing, vol. 316, pp. 332-344, 2018, doi: 10.1016/j.neucom.2018.08.009.

12 A. Ruta, Y. Li, and X. Liu, "Detection, tracking and recognition of traffic signs from video input," in 2008 11th International IEEE Conference on Intelligent Transportation Systems, 2008: IEEE, pp. 55-60

13 S. Varun, S. Singh, R. S. Kunte, R. S. Samuel, and B. Philip, "A road traffic signal recognition system based on template matching employing tree classifier," in International Conference on Computational Intelligence and Multimedia Applications (ICCIMA 2007), 2007, vol. 3: IEEE, pp. 360-365.

14 H. Fleyeh, "Color detection and segmentation for road and traffic signs," in IEEE Conference on Cybernetics and Intelligent Systems, 2004., 2004, vol. 2: IEEE, pp. 809-814.

15 M. Sridharan and P. Stone, "Towards on-board color constancy on mobile robots," in First Canadian Conference on Computer and Robot Vision, 2004. Proceedings., 2004: IEEE, pp. 130-137.

16 S. Xu, "Robust traffic sign shape recognition using geometric matching," IET Intelligent Transport Systems, vol. 3, no. 1, pp. 10-18, 2009.

17 N. Barnes, A. Zelinsky, and L. S. Fletcher, "Real-time speed sign detection using the radial symmetry detector," IEEE Transactions on Intelligent Transportation Systems, vol. 9, no. 2, pp. 322-332, 2008.

18 S. Maldonado-Bascon, S. Lafuente-Arroyo, P. Gil-Jimenez, H. Gomez-Moreno, and F. Lopez-Ferreras, "Road-Sign Detection and Recognition Based on Support Vector Machines," IEEE Transactions on Intelligent Transportation Systems, vol. 8, no. 2, pp. 264-278, 2007, doi: 10.1109/tits.2007.895311.

19 W.-J. Kuo and C.-C. Lin, "Two-stage road sign detection and recognition," in 2007 IEEE international conference on multimedia and expo, 2007: IEEE, pp. 1427-1430.

20 S. K R, D. M, and V. D. R S, "Traffic Sign Recognition with Faster RCNN and RPN for Advanced Driver Assistance Systems," presented at the 2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2021.

21 D. Tabernik and D. Skocaj, "Deep Learning for Large-Scale Traffic-Sign Detection and Recognition," IEEE Transactions on Intelligent Transportation Systems, vol. 21, no. 4, pp. 1427-1440, 2020, doi: 10.1109/tits.2019.2913588.

22 Z. Zhu, D. Liang, S. Zhang, X. Huang, B. Li, and S. Hu, "Traffic-sign detection and classification in the wild," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 2110-2118.

Authors Introduction

Mr. Hanlin Cai

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 Internet of Things and machine learning.

Mr. Zheng 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 machine learning and path planning.

Mr. Jiaqi Hu

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 machine learning and signal processing.

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 control systems.

Wall Crack Detection based on Adaptive Double Threshold Greyscale Transform

Mingrui Lin[1], Xin Xu[1], Shilin Chen[1], Tengxiang Li[1], Yuhan Hong[1], Weiqin Wang[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: mingrui.lin.2021@mumail.ie, xin.xu.2021@mumail.ie, shilin.chen.2021@mumail.ie, tengxiang.li.2021@mumail.ie, yuhang.hong.2021@mumail.ie, weiqin.wang.2021@mumail.ie, limwh@ucsiuniversity.edu.my, chinhong.wong@mu.ie

Abstract

The construction industry is an important industry supporting social-economic development. Detecting cracks in walls could help to maintain the conditions of buildings. A wall crack detection algorithm based on adaptive double threshold greyscale transform is proposed. The MATLAB built-in '*graythresh*' function is modified in order to have a flag bit to preliminarily selection of the greyscale transformation threshold where the grey transform threshold and the binarization threshold can be automatically adjusted based on the image processing effect using a preliminary-selected threshold. The algorithm enhances the crack information and weakens the background information by multiple morphological operations of removing isolated small area pixels. The MATLAB simulation results show that the proposed algorithm has an accuracy of 96.16%, time from inputting the image to the completion of the labelling is 3-15 seconds, which depends on the complexity of the crack.

Keywords: wall crack detection; image processing; double threshold; adaptive greyscale transform; MATLAB

1. Introduction

In recent years, with the rapid development of the economy, the speed of urban construction is accelerating and the scale of old city renovation is also expanding. The construction industry has become an important industry to support social and economic development. Therefore, regular detection of wall cracks is an important task for building health inspection.

In the past, the inspection and maintenance of buildings mainly relied on manual inspection. Some wall cracks were located in an unpredictable dangerous areas, so it was very difficult to carry out inspection work. With the development of digital image processing and computer vision, morphological image processing based on computer vision technology is widely used in target detection and has been successfully applied to bridge and building crack detection. Compared with the manual inspection method, crack detection based on morphological image processing provides a lower cost and better real-time performance [1][2][3]. In this work, an adaptive double threshold greyscale transform and binarization algorithm for wall crack detection is proposed.

2. Related Works

At present, crack image detection methods mainly include spatial domain image processing and transform domain analysis such as frequency domain and wavelet domain processing.

S. Ogawa *et. al.* [4] proposed a wall crack image recognition based on crack feature extraction combining Gaussian mixture model and image filtering, and classification by support vector machine. This method is able to achieve an accuracy of up to 80%.

H. B. Yun *et. al.* [5] proposed crack recognition and segmentation using morphological image-processing techniques for flexible pavements. The algorithm consists of two subprocesses: (a) the grouping of fragments by using a morphological dilation transform and (b) the connection of fragments by using a morphological thinning transform. This method could improve crack detection accuracy.

L. Zhang *et. al.* [6] proposed an automatic detection method based on deep convolutional neural networks for road crack detection. Results show that the learned deep features with the proposed deep learning framework provide superior crack detection performance when compared with features extracted with existing hand-craft methods.

Q. Zou *et. al.* [7] proposed a DeepCrack-an end-to-end trainable deep convolutional neural network for automatic crack detection by learning high-level features for crack representation. In their work, multi-scale deep convolutional features learned at hierarchical convolutional stages are fused together to capture the line structures. Results show that DeepCrack achieves over 0.87 ODS *F* -measure value.

X. Yang *et. al.* [8] proposed a novel deep learning technique named fully convolutional network to detect cracks. Results show that the fully convolutional network is feasible and sufficient for crack identification and measurement.

3. The Proposed Method

In this paper, a wall crack detection algorithm based on adaptive double threshold greyscale transform is proposed. The MATLAB built-in function, '*graythresh*', is modified in order to select the greyscale transformation threshold which is the grey transform threshold and the binarization threshold. The algorithm enhances the crack information and weakens the background information through multiple morphological operations. Since the fixed threshold is not suitable for small-size image detection, the image size is transformed before image preprocessing to convert images of different sizes into standard sizes and resize them back to the original size after completing them to match the threshold value. The traditional edge detection is discarded, and the closed operation is used to increase image connectivity. A judgment method of an overlapping anchor is proposed to solve the problem of repeated marking that usually occurs in the traditional marking algorithm. Fig. 1 shows the flow chart of the algorithm.

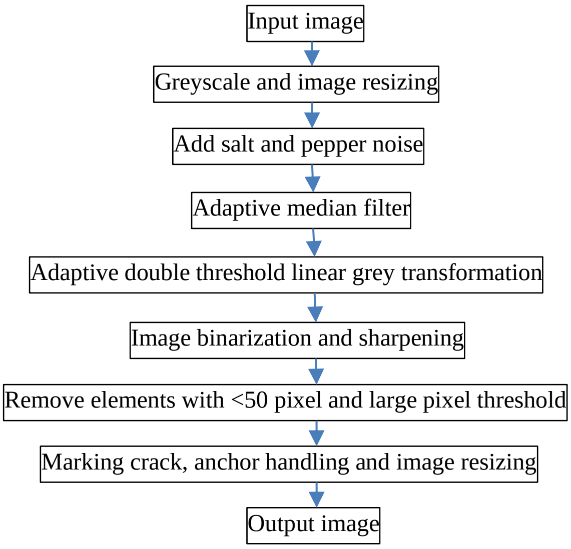

Input image
Greyscale and image resizing
Add salt and pepper noise
Adaptive median filter
Adaptive double threshold linear grey transformation
Image binarization and sharpening
Remove elements with <50 pixel and large pixel threshold
Marking crack, anchor handling and image resizing
Output image

Fig. 1. The flow chart of the algorithm

4. Results and Discussion

To increase the accuracy, the greyscale process is performed on the collected colour using a MATLAB built-in function *rgb2gray*. Fig. 2 shows the image before and after greyscale process.

(a) (b)

Fig. 2. Greyscale process (a) before and (b) after.

Since the fixed threshold is not suitable for small-size image detection, image size transformation is introduced to convert images of different sizes into standard sizes of 2500×2500. In this work, the nearest neighbour interpolation is used for reducing image size. Cubic spline interpolation and bilinear interpolation are used to enlarge images.

Due to the interference of noise, the quality of the image will be reduced and the crack information in the image will become fuzzy. In this work, an adaptive median filtering technique is applied to smooth and reduce the noise of images [9][10].

Through simulation experiments, the results show that the adaptive median filtering algorithm is significantly better than the traditional filtering algorithm in preserving the edge details of the image. Fig. 3 shows the difference between added salt and pepper noise to the image and the image after adaptive median filtering is applied.

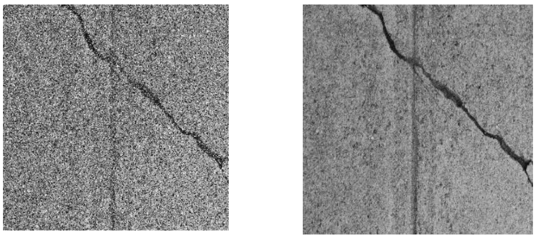

(a) (b)

Fig. 3. Comparison between (a) adds salt and pepper noise to the image with 0.3 of noise density (b) image after adaptive median filtering

Due to the complexity and variety of crack images, the fixed single grey transform threshold cannot well adapt to the complex image. A MATLAB built-in function *gradthresh* is modified by adding a flag bit to achieve the purpose of initial selection of the grey transform threshold. The grey transform threshold and the binarization threshold can be automatically adjusted based on the image processing effect using a preliminary-selected threshold.

The experimental results show that images with little difference between the crack information and background information need to be enhanced. Based on the value of each colour component count matrix from the histogram, the maximum value is larger and the variance is smaller for images with an obvious difference between the crack information and background information which required a smaller grey threshold. On the other hand, the maximum value of each colour component count matrix of the histogram counts has a smaller maximum value and a larger variance. Therefore, *graythresh* function is modified. If the maximum value of counts is larger than 10^5, the flag is set to 1, otherwise flag is set to 0.

From the test results, for images with a large distinction between crack information and background information, stretching the original greyscale interval from [0.25 0.35] to [0, 1] provides a good processing effect. However, for images with a small distinction between crack information and background information, stretching the original greyscale interval from [0.5 0.6] to [0, 1] gives a good processing effect.

However, a single threshold cannot satisfy the needs of all different images, the adaptive double threshold transformation method is needed to solve this problem. The selection of the greyscale transform interval is depended on the initial greyscale transform interval which was selected by the flag bit. Fig. 4 shows the difference between various transformation intervals.

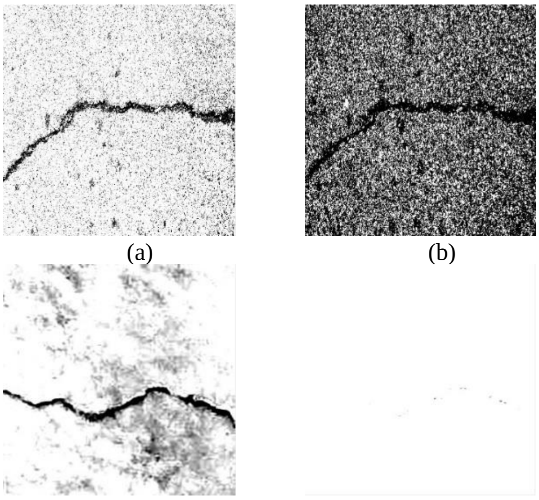

(a) (b)
(c) (d)

Fig. 4. (a) images with large distinction ([0.25 0.35] to [0, 1]), (b) images with large distinction ([0.5 0.6] to [0, 1]), (c) images with small distinction ([0.5 0.6] to [0, 1]) and (d) images with small distinction ([0.25 0.35] to [0, 1])

For the greyscale transformation interval of [0.25, 0.35], if the black pixel ratio of the processed image after binarization is less than 1%, the initial interval can be considered inappropriate and should be replaced with the greyscale transform interval of [0.5,0.6]. Whereas, for the greyscale transform interval of [0.5 0.6], if the black pixel ratio of the processed image, after binarization, is greater than 35%, the initial interval can be considered inappropriate and should be replaced with the greyscale transform interval of [0.2 0.35].

Combined with the selection of parameters in the previous steps, it was found that an acceptable result is achieved on the image with an obvious crack with a 0.2 threshold value. Whereas, for images without obvious crack information, a 0.7 threshold value is selected to achieve a better effect.

Considering the diversity of the image between the crack and background, in order to enhance the robustness of the algorithm, a threshold of 0.48 was selected. Fig. 5 shows the difference between images with and without obvious cracks.

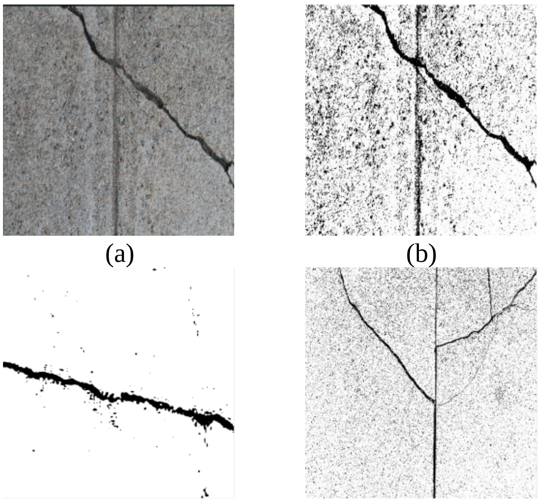

(a) (b) (c) (d)

Fig. 5. Illustration of the proposed approach:(a) original image (b) image after binarization with threshold 0.48 (c) not obvious crack image with threshold 0.7 (d) complex background image with threshold 0.2

Image noise can be observed even after binarization and Gaussian filtering enhancement. In this work, two small area removing operations were performed followed by a close operation between the two removing operations. Choosing only one large threshold-removing operation, not only the small area cracks will be removed but all the crack information will also be filtered. Therefore, two filtering processes are performed. The first filtering process which is a small threshold filtering operation is applied to remove most of the noise and protect the crack information followed by the second filtering process which is a close operation to compensate for the crack information that is filtered out in the first filtering operation. After these two filtering processes, the image only left most of the crack information, hence, a larger threshold can be chosen to completely filter out the noise. Since the depth and thickness of the cracks vary from image to image, automatic thresholding in the second filtering operation is used to fit most of the images.

Experimental results show that most of the noise areas are in the range of 0-30 pixel points, so, the threshold value of 50 was chosen for the first filtering operation. On the other hand, for most 2500×2500 images, the maximum connected domain area is between $10^3 - 10^6$, while the other areas are relatively small. Hence, 1% of the maximum connected domain area was chosen as the threshold for the second filtering operation. It was found that the selected threshold value for the second filtering operation achieves a better preprocessing effect for most images. Fig. 6 show the difference between the images before and after the double removing operation.

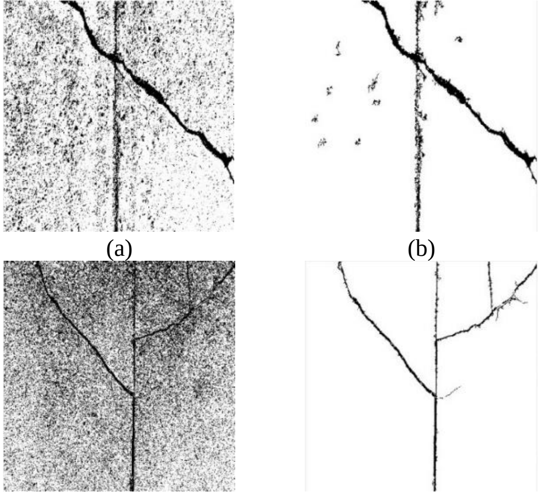

(a) (b) (c) (d)

Fig. 6. Image (a)&(c) before and (b)&(d) double removing operation

The selection and setting of the structural elements are very important for the closing operation. The selection of different structural elements leads to different segmentation and different filtering effects. A disk-shaped structuring element with a radius of 3 is chosen. Fig. 7 shows the difference between images before and after the closing operation.

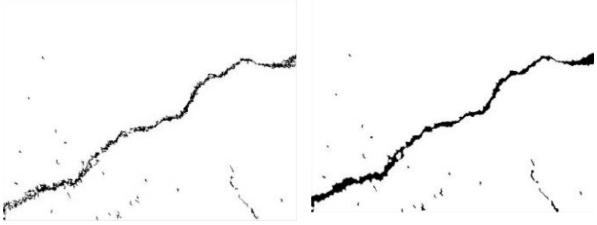

(a) (b)

Fig. 7. Image (a) before and (b) after closing operation

The crack on the preprocessed image is marked. Due to the size of the preprocessed image is not the same as the original image, a size transformation on the image to restore it to the original size is performed. In this work, connected domain analysis is used to extract the crack information.

Due to the complexity of the crack image, especially for some block cracks, using only the connected domain approach to calibrate the cracks is not enough to achieve the expected results.

Before marking, the parameters of all the connected domains are calculated, including the coordinates of the points and the area of the anchor. If the area of the anchor box is too small, it will not be drawn. From the experimental results, the threshold to $0.08 \times Area_{max}$ gives a better annotation effect and is suitable for most images.

In order to avoid double marking of anchor boxes, it is necessary to hide the overlapping anchor boxes. If the midpoint of the smaller rectangle lies inside the larger rectangle, it is assumed that at least 50% of the area of the smaller rectangle overlaps with the larger rectangle. If there is an overlap, the larger anchor box will be kept. Fig. 8 shows the marking for different crack images.

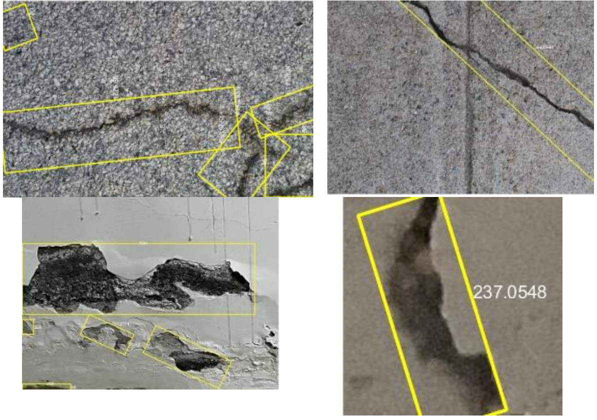


Fig. 8. Marking for different crack images

In order to verify the accuracy of the algorithm, two sets of data set were selected for testing namely the "Concrete Crack Images for Classification" dataset [11] and the "CRACK500" dataset [12][13]. The first 1000 images in the "Concrete Crack Images for Classification" dataset were selected for testing whereas the first 200 images in the "CRACK500" dataset were selected for testing, and the test results are shown in Table 1. Because the images in the "Concrete Crack Images for Classification" dataset are simple and the image background in the "CRACK500" dataset is complicated, in order to improve the scientific accuracy and the accuracy of the algorithm, the ratio of 3:1 is used between "Concrete Crack Images for Classification" and "CRACK500".

Table 1. Accuracy Results

Dataset	Total image	Fail to detect	Accuracy (%)
Concrete Crack Images for Classification	1000	6	99.40
CRACK500	200	28	86.32
Total (weighted)*	1200	104	96.16

* Concrete Crack Images for Classification:CRACK500 = 3:1

5. Conclusion

The experimental verification shows that the proposed algorithm has a detection accuracy of 96.16% for most of the images. However, the detection accuracy for some special cases or complex crack images still needs to be optimized

References

1. L. Yu, S. He, X. Liu, S. Jiang, and S. Xiang, "Intelligent Crack Detection and Quantification in the Concrete Bridge: A Deep Learning-Assisted Image Processing Approach," *Advances in Civil Engineering*, vol. 2022, 2022, doi: 10.1155/2022/1813821.
2. S. Sonkhiya, "MATLAB Image Processing Method using Histograms for Railway Track Crack Fault Detection," *Int J Res Appl Sci Eng Technol*, vol. 9, no. VI, 2021, doi: 10.22214/ijraset.2021.34888.
3. Q. Q. Li, Q. Zou, and Q. Z. Mao, "Pavement crack detection based on minimum cost path searching," *Zhongguo Gonglu Xuebao/China Journal of Highway and Transport*, vol. 23, no. 6, 2010.
4. S. Ogawa, K. Matsushima, and O. Takahashi, "Efficient Pavement Crack Area Classification Using Gaussian Mixture Model Based Features," in *Proceedings of the 2019 International Conference on Mechatronics, Robotics and Systems Engineering, MoRSE 2019*, 2019. doi: 10.1109/MoRSE48060.2019.8998713.
5. H. B. Yun, S. Mokhtari, and L. Wu, "Crack recognition and segmentation using morphological image-processing techniques for flexible pavements," *Transp Res Rec*, vol. 2523, 2015, doi: 10.3141/2523-13.
6. L. Zhang, F. Yang, Y. Daniel Zhang, and Y. J. Zhu, "Road crack detection using deep convolutional neural network," in *Proceedings - International Conference on Image Processing, ICIP*, 2016, vol. 2016-August. doi: 10.1109/ICIP.2016.7533052.
7. Q. Zou, Z. Zhang, Q. Li, X. Qi, Q. Wang, and S. Wang, "DeepCrack: Learning hierarchical convolutional features for crack detection," *IEEE Transactions on Image Processing*, vol. 28, no. 3, 2019, doi: 10.1109/TIP.2018.2878966.
8. X. Yang, H. Li, Y. Yu, X. Luo, T. Huang, and X. Yang, "Automatic Pixel-Level Crack Detection and Measurement Using Fully Convolutional Network," *Computer-Aided Civil and Infrastructure Engineering*, vol. 33, no. 12, 2018, doi: 10.1111/mice.12412.
9. C. Dai, K. Jiang, and Q. Wang, "Recognition of tunnel lining cracks based on digital image processing," *Math Probl Eng*, vol. 2020, 2020, doi: 10.1155/2020/5162583.
10. H. Hu, G. Dong, B. Peng, J. Xing, and W. Song, "Method for Detecting Micron Cracks on a Magnetic Rotor Surface Based on a Support Vector Machine," *IEEE Access*, vol. 6, 2018, doi: 10.1109/ACCESS.2018.2870435.
11. Çağlar Fırat Özgenel, "Concrete Crack Images for Classification," *15 Jan*, 2018. https://data.mendeley.com/datasets/5y9wdsg2zt/2

12. W. Wang and C. Su, "Convolutional Neural Network-Based Pavement Crack Segmentation Using Pyramid Attention Network," *IEEE Access*, vol. 8, 2020, doi: 10.1109/ACCESS.2020.3037667.
13. R. Augustauskas and A. Lipnickas, "Improved pixel-level pavement-defect segmentation using a deep autoencoder," *Sensors (Switzerland)*, vol. 20, no. 9, 2020, doi: 10.3390/s20092557.

Authors Introduction

Mr. Mingrui Lin

He is currently pursuing Bachelor of Engineering as 3rd year student in the Department of Maynooth International Engineering College, Fuzhou University, China. His research interests are image processing and maching learning.

Ms. Xin Xu

She is currently pursuing Bachelor of Engineering as 3rd year student in the Department of Maynooth International Engineering College, Fuzhou University, China. Her research interests are machine learning and telecommunication.

Mr. Shilin Chen

He is currently pursuing Bachelor of Bachelor of Engineering as 3rd year student in the Department of Maynooth International Engineering College, Fuzhou University, China. His research interests are Embedded system, FPGA, and DSP

Mr. Tengxiang Li

He is currently pursuing Bachelor of Engineering as 3rd year student in the Department of Maynooth International Engineering College, Fuzhou University, China. His research interests are signal processing and image processing.

Mr. Yuhang Hong

He is currently pursuing Bachelor of Engineering as 3rd year student in the Department of Maynooth International Engineering College, Fuzhou University, China. His research interests are image processing and embedded system.

Ms. Weiqin

She is currently pursuing Bachelor of Engineering as 3rd year student in the Department of Maynooth International Engineering College, Fuzhou University, China. Her research interests are robotics and embedded system 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.

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 control systems.

Noise filtering of Hyperspectral Data of Oil Palms by Median Mean Projection Filtering

Imanurfatiehah Ibrahim, Hamzah Arof, Mahazani Mohamad, Mohamad Sofian Abu Talip
Department of Electrical Engineering, Faculty of Engineering,
University of Malaya, 50603, Kuala Lumpur, Malaysia

Mohd Izzuddin Anuar
Malaysian Palm Oil Board, 43000 Kajang Selangor, Malaysia

E-mail: imanurfatiehah@gmail.com, ahamzah@um.edu.my, mahazani@um.edu.my, sofian_abutalip@um.edu.my,
mohamad.izzuddin@mpob.gov.my

Abstract

One of the many applications of hyperspectral imaging is in agriculture. However, hyperspectral data captured from airborne UAV sometimes contain speckle noise that make their spectral signatures different from those of field spectroscopy using similar wavelengths. Therefore, there exists a need to filter noisy hyperspectral data to improve their quality so that a strong correlation can be established between the airborne and field hyperspectral data for effective analysis. For oil palm hyperspectral data, an efficient and effective method is introduced to filter noise using median-mean projection filtering. This novel approach generates superior results compared to those produced by the conventional method of convolving the data with 2-D filters, in terms of output quality and signal to noise ratio. The resulting data also exhibit better luminosity and contrast. The proposed method was implemented using MATLAB R2021b running on Intel i7 processor and the average execution time was less than 10 seconds. The execution time can be further lessened if the codes are optimized and GPU is used. Overall, the proposed method removes the speckle noise and improves the luminosity and contrast of the data. This technique can be a useful tool to those working in the oil palm industry.

Keywords: Hyperspectral images, med-mean projection filtering, UAV, noise.

1. Introduction

Hyperspectral imaging has been used in many applications including agriculture, remote sensing, oil and gas exploration, quality control**,** forensic science, biotechnology, medicine and others [1][2]. Regular two dimensional (2-D) images captured by common cameras are either in color or grayscale. Usually, color images have three components, and they are red, green and blue. In terms of wavelengths (also known as bands or channels), the standard color space (sRGB) adopted by computer monitor manufacturers assigns red to 612nm, green to 549nm and blue to 464nm [3]. Airborne hyperspectral remote sensing provides data in the form of stacked 2-D images from the reflections of many bands of wavelengths captured by a hyperspectral camera. Each 2-D image corresponds to the reflection of a narrow channel or band of wavelength. Since all of the images are taken of the same spot (or location) using different bands of wavelength, it is expected that there will be overlaps of content among images. However, some images contain exclusive information not available in others. When collected and combined, this information becomes a powerful tool for data analysis. For example, in aerial hyperspectral data of oil palms, images of different bands are successfully used to detect diseases such as freckle, blast and Ganoderma infections [4][5].

Often, hyperspectral data obtained from long and short wavelengths contain more noise than those obtained using medium wavelengths. Noise contamination in the form of speckle noise (also known as salt and pepper noise) reduces the quality of the images. If the noise is severe, it must be filtered so that important information is not obscured or dampened. For instance, airborne hyperspectral data from the aerial view may contain noise that make them appear relatively different from the field spectroscopy data taken on the ground.

Consequently, it is hard to establish a clear correlation between them for ground truth verification. In this case, the noise in the airborne hyperspectral data must be removed so that spectral signatures in the data can be preserved, identified, and used for detection, identification, classification, and other forms of data analysis.

Normally, data that are badly affected by speckle noise are discarded. This is because, the values they contain are too volatile to be useful. Usually, only channels that are associated with very short or very long wavelengths are severely affected by speckle noise. That implies that most of the data are unaffected by speckle noise. In this study, we introduce a method to filter channels that are badly affected by speckle noise. This way, we can salvage noisy data associated with short and long wavelengths from being discarded.

2. Methodology

Noise in hyperspectral images mainly originate from changes in sun illumination intensity during the airborne image acquisition especially when part of the area is shaded by thick cloud. In addition, high amount of humidity in the field and air may absorb or scatter water-sensitive wavelengths used to capture the images. Furthermore, haze, floating particles, debris and strong reflectance from the background soil during sunny day may introduce noise or reduce contrast of the hyperspectral images [6]. Another factor that affects the quality of the hyperspectral data is the resolution of the image. It is the corresponding area on the ground represented by a single pixel in the image. The bigger the area, the lower the resolution is. Normally, the resolution is dictated by the quality of the camera used and the altitude of the UAV when the data is captured. Finally, the flight speed also has an impact on the data quality. Thus, it is important to maintain the altitude and speed of the UAV when capturing the hyperspectral data.

2.1. *Database*

The area under study is in Lekir, Perak, Malaysia. It is a flat coastal area which contains almost 20 thousand oil palms of Dura x Pisifera (DxP) species. The age of the oil palms is nearly ten years old by 2023. The area receives a good rainfall of approximately 2700 mm every year. The oil palms are free from any disease, stress or infestation except for *Ganoderma* infection that affects less than 2 percent of the oil palm population.

The hyperspectral camera used was Resonon Pika L (Resonon Inc., Bozemann, Montana, USA) mounted on a DJI M600 Pro hexacopter drone. The camera captured a set of hyperspectral images with 300 spectral bands from 350 nm – 900 nm with ±2 nm spectral resolution. The hexacopter-type drone has a maximum payload of 6 kg and a flight duration of 15 – 20 minutes. A calibration tarpaulin was used as image white-calibration reference on the ground. Data recording activities were performed under clear sky during sunny days.

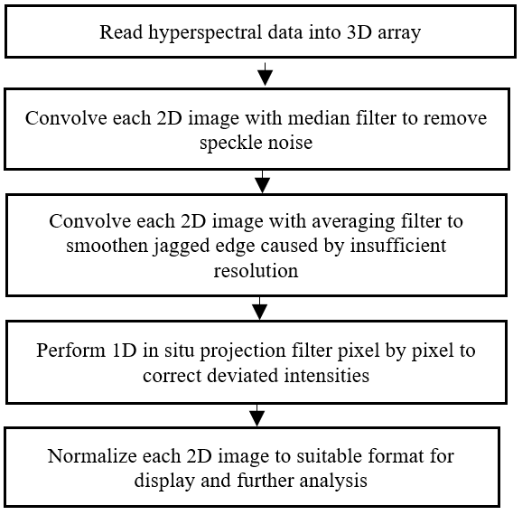

Read hyperspectral data into 3D array
Convolve each 2D image with median filter to remove speckle noise
Convolve each 2D image with averaging filter to smoothen jagged edge caused by insufficient resolution
Perform 1D in situ projection filter pixel by pixel to correct deviated intensities
Normalize each 2D image to suitable format for display and further analysis

Fig. 1 Flowchart of steps in median-mean projection filtering.

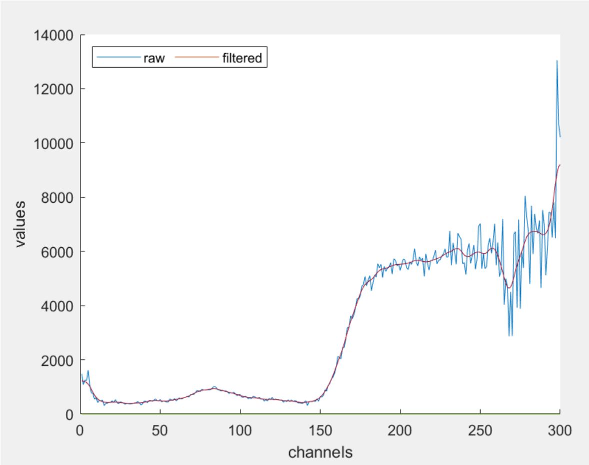


Fig. 2 The smoothing effect of applying the projection filter on 300 channels values of a sample pixel.

2.2. *Approach*

In this work, median-mean (med-µ) projection filtering is used to filter the noise. The main idea is the fact that a 2-D image obtained from a particular band shows an in situ positive correlation with images of neighboring bands of contiguous wavelengths. The flowchart of steps taken to filter noise from the hyperspectral data is shown in Fig. 1. The first step in the process is to read the hyperspectral data and store them into a 3D array. The size of the data is 1000x1000x300. Then each 2D image is filtered (convolved) with a median filter to remove speckle noise. In our work the size of the filter is limited to 3x3 since a larger median filter would eliminate thin leaves from the frond of the oil palms and take a longer time to execute.

Next, the image is subjected to average (or mean) filtering. Since the altitude of the UAV was 80m when the data were captured, the resolution of the images is good but not excellent. In fact, it is estimated that the best resolution attained is approximately 10 cm x 10 cm for each pixel. This makes some fronds of the canopy appear jagged. Average filtering smoothens the jagged edges caused by the insufficient resolution, but it also blurs sharp edges. Thus, the size of the averaging filter is also limited to 3x3 so that it doesn't reduce the contrast of the images too much.

Finally, the images undergo in-situ projection filtering, pixel by pixel, to capitalize on the strong correlation between contiguous bands (channels). The projection is performed one dimensionally from the middle band towards the low and high ends of the channels where pixel values become volatile as noise content increases. The projection filter employs Savitzky-Golay data smoothing technique to subdue volatility at the tail ends of the channels. In our experiments, we opted for Savitzky-Golay denoising filter of 3rd degree polynomial with 11 frame length. Fig. 2 shows the effect of performing projection filter on the values of 300 bands of one sample pixel. It is seen that the volatile values of the 300 bands are smoothened by projection filtering. It is especially obvious at the tail ends of the channels. The process is repeated for every pixel in the image so that the values of the pixels in bands that are severely affected by speckle noise can be replaced by the projected values from the filter. Finally, to display the data of each channel in the form of an image, they must be normalized so that their range will fall between 0 and 255. Then they can be displayed for further analysis.

3. Results and Discussion

The method is applied to hyperspectral data containing 300 spectral bands. The size of each band is 1000 x 1000 pixels. However, since there are 300 bands or images, each image is processed consecutively. The clarity of the images markedly improves after filtering as the noise is reduced. Fig. 3 shows the results of applying median filtering, projection filtering and the proposed method on a 400x400 area of one channel of the hyperspectral data that is affected by speckle noise. The size of the median filter used is 3x3. When the size of the median filter is increased to 5x5 or 7x7, it takes longer to implement. The bigger filter tends to remove the noise better but also blur the image more as it eliminates small objects that are bright or dark such as fronds and leaves. The mean filter is the moving average filter of 3x3. By itself the Savitzky-Golay projection filter is quite effective in filtering the speckles, but the best result is obtained from the combination of the median, mean and projection filters. The signal to noise ratios (SNR) and the average execution times of the filters are given in Table 1. As stated, the image becomes quite clean after filtering as the SNR value exceeds 40dB. Here SNR is estimated as the ratio of the standard deviation (sigma) of a reference image over the difference of the sigmas of a noisy image and the reference image. It is written mathematically as following Eq. (1)

$$\mathrm{SNR} = 10 \log_{10} (\sigma_r / (\sigma_n - \sigma_r)) \qquad (1)$$

where σ_r is the sigma of the reference image

σ_n is the sigma of the noisy image

This is based on the assumption that the strength of the signal in each band (image) is identical since they are normalized and the reference band is considered relatively free of noise. Normally, the reference image is one of the middle channels with the lowest speckle noise or the lowest standard deviation of uniform areas.

Furthermore, the standard deviation of each image is calculated only in selected areas that are considered of uniform intensity by manual inspection. Alternatively, the areas can be automatically identified as the ones with the lowest standard deviations in the reference channel. In our case, these areas are the dark spots between palm leaves. The proposed method was implemented on MATLAB platform running an Intel i7 multicore processor. The average execution time for processing a 400x400 subregion for one channel is approximately 0.5s as stated in Table 1. With further refinement of the codes and the use of parallel processing or GPU, the execution time can be decreased more.

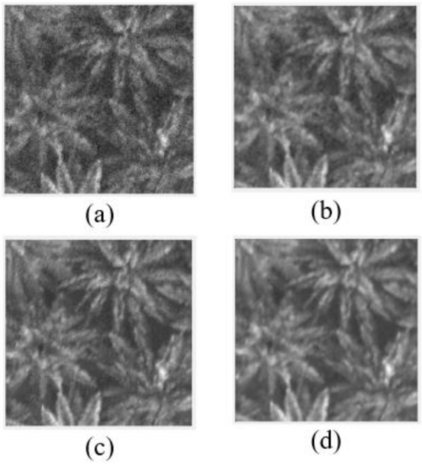

(a) (b) (c) (d)

Fig. 3 A noisy subregion of an image a) Original subregion b) Filtered by median filter c) Filtered by projection filter d) Filtered by median and projection filters.

Table 1. The performances of the median, projection and proposed method.

Methods	Median	Projection	Proposed
SNR	21.4 dB	34.8 dB	40.5 dB
Time/image	0.1s	0.4s	0.5s

4. Conclusion

In this paper, a new method to filter speckle noise from hyperspectral images is proposed. The proposed method employs projection filter which capitalizes on the positive correlation among contiguous channels. It removes speckle noise from the images, improves their visual quality and increases the SNR values.

References

1. K.A. Parveez, H. Elina, S.K. Loh; O. Meilina, M. S, Kamalrudin, Z. B. Mohamad Nor Izudin, S. Shamala, A.A.H. Zafarizal, I. Zainab. Oil Palm Economic Performance In Malaysia And R&D Progress In 2019. Journal Oil Palm Research, vol.32, no. 2, pp. 159-190, 2019.
2. A. Roslan and A. S. Idris. Economic Impact of Ganoderma Incidence on Malaysian Oil Palm Plantation – A Case Study in Johor. Oil Palm Industry Economic Journal, vol. 12, no.1. pp.24-30, 2012.
3. A.S. Idris, M.S., Mazliham, P. Loonis, and M. W. Basri. Ganosken for Early Detection of Ganoderma Infection in Oil Palm. MPOB Information Series, 442, pp 1-4. 2010
4. A.S., Idris, S. Rajinder, A.Z. Madihah and M.W. Basri. Multiplex PCR-DNA Kit for Early Detection and Identification of Ganoderma Species in Oil Palm. MPOB Information Series, 73, pp-1-4, 2010.
5. D. Ariffin and A.S. Idris. The Ganoderma Selective Medium (GSM). MPOB Information Series, 8, pp1-2, 1992.
6. M.A. Izzuddin, A.S. Idris, M.N.Nisfariza, A.A., Nordiana, H.Z.M. Shafri and B. Ezzati. The Development of Spectral Indices for Early Detection of Ganoderma Disease in Oil Palm Seedlings. International Journal of Remote Sensing, 38, pp. 6505-6527, 2017.

Authors Introduction

Imanurfatiehah Ibrahim

She graduated from International Islamic University of Malaysia and obtained her masters from University of Malaya. Currently she is a PhD student at the Electrical Engineering Department, Faculty of Engineering, University of Malaya. Her research interests are in the areas of big data processing and remote sensing.

Prof. Ir. Dr. Hamzah Arof

He received the B.Sc. degree from Michigan State University, USA, and the Ph.D. degree from the University of Wales, U.K. He is currently a Professor with the Electrical Engineering Department, Universiti Malaya, Malaysia. He is an established academician. He has published more than 80 articles. His research interests include image and signal processing, robotics, and photonics

Mohd Izzuddin Anuar

He is a is is a researcher at Malaysian Palm Oil Board and a PhD student at Electrical Engineering Department, Faculty of Engineering, University of Malaya. He graduated from Universiti Teknologi Malaysia and obtained his masters from Universiti Putra Malaysia. Currently he is a PhD student at the Electrical Engineering Department, Faculty of Engineering, University of Malaya. His research interests are in remote sensing and pattern recognition.

Mahazani Mohamad

He received the B.Eng degree from Kyoto University, Japan and the M.Eng. Degree from University of Queensland, Australia. He is currently lecturer at the Electrical Engineering Department, University of Malaya, Malaysia. His area of expertise is Computer Systems Engineering.

Mohamad Sofian Abu Talip

He received the B.Eng degree from Universiti Sains Malaysia, the M.Sc. Degree from International Islamic University Malaysia, and PhD degree at the Keio University, Japan. He is currently Senior Lecturer at the Electrical Engineering Department, University of Malaya, Malaysia. His area of expertise are Reconfigurable System, Embedded System, System on Programmable Chip and FPGA-Based Design.

Smart Telehealth Appointment System – WI Care

Siah Cheong Lin, Heshalini Rajagopal, Thong Chee Ling
Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia

E-mail: heshalini@ucsiuniversity.edu.my, chloethong@ucsiuniversity.edu.my
www.ucsiuniversity.edu.my

Abstract

Modern technology is all about speed and efficiency to connects you with your healthcare provider, ensures virtual communications, and gives you more access and control over the services you receive. Based on literature study, there is a lack of healthcare appointment system that enables patients making appointment and doctor confirming the appointment on the fly. This project aims to create a Smart Telehealth Appointment System -WI Care that enables user to make an appointment in ease. Compared to existing health care system, WI Care will be more prior to services instead of technology. Due to COVID-19 pandemics, constant lockdown had resulting social distancing measures, movement control order, and people avoid outdoor activities. These also led to poor service quality due to limitation of staff working in office with less working hours which directly affects the GPs and doctors' career at dead end. For the fundamentals of the system, WI Care will have an appointment system that can be managed by doctors meanwhile appointer can keep submitting form. WI Care ensure both parties can be conducting the same procedure as how every GPs and doctors able to work anywhere anytime. With dashboard and time scheduling, these are able to create chance for doctors and GPs having manageable workload as individual without the reliance of hiring staff members. This project suggests the healthy environment whereas every doctor can join WI Care as big community, share skills and tackles problem together.

Keywords: Telehealth; Appointment System; Web-Based Clinic System

1. Introduction

Health Care System is system that provides for services to be given to people in defined settings, which including residences, academic institutions, workplaces, public places, societies, hospitals, and health centers, in order to contribute to their wellness. It also specifically designed to meet the health needs of a population, and as a result, it is much broader than one would think upon first, which would include community care and programs, public health and protection, private and publicly funded health care, and so forth.

Health Care System can break down into many counterparts, ranking by most common used such as Electronic Health Record (EHR) Software that collects information of the patients and provide recommendations; Medical Database Software that stores medical historical records and treatment plans; Telemedicine that assist on handle appointments with patients online through web and mobile platform etc. [1].

However, Smart Telehealth Appointment System is a web application that consist similar feature to any appointment system. The system can be accessed by doctors, administrator and any other user. Tasks can be performed by current system in making appointment, scheduling and modifying the result through the system, and also enabled description and design editing. Despite lockdown has been lifted in our country long ago, people still rely on virtual meeting format. These activities are also applied to medical field as well. Currently, doctors and nurses also prefer using handheld devices to record patients' status, which has proven for better precision and efficiency for diagnoses and treatments, overall improving the quality of healthcare [2]. Consumers also start to value their health more during the crisis of COVID-19. The findings validate that virtual visit are being used by consumers higher than before. From 2019 till early 2020, the result shows that virtual visits rate ascended from 15% to 19%, and a huge increment to 28% in April 2020 [3]. The majority of customers were pleased with their visits and mentioned that they would utilize this form of therapy again.

One of the systems that applies for the current healthcare technology are the smart systems. It is a system integrated with functions of sensing, control, and actuation to examine real life scenario, hence make decisions based on available data that is predictive or adaptive and therefore carry out "smart" action. Smart system has been applied in many fields of work like consultation, delivery, smart home, or even devices such as SMART TV, Smart refrigerators and etc.

Most existing smart healthcare project are focus on IoT field, where the smart technologies are applied on wearable items and devices such as Biosensors, connected inhalers, Smart watches, FitBits, Blood pressure monitors and more [4]. However, telehealth also started gain popularity only when arrival of COVID-19. Telehealth cover services like video call meeting, healthcare portals to assist doctors and patients in keeping track of therapy, storing health records, and collecting and accessing data via wearable devices and clinic visits. In spite of that, most telehealth project are only targeting mobile platform. Although smartphone devices are far more convenient to track medical records and communications as it has high accessibility, most of the mobile system platform need to update manually, contains lack of IoT interaction and limitations for hardware. Telehealth project in app format also tend to have less accurate record and support services compare to other devices. Virtual meeting and notifications are two of the important services needed to offer within telehealth consulting system, so the best alternative to replace existing telehealth project is access through mobile website integrated with automation services.

2. Methods

The WI Care serves as a web application integrated with services that provides utility similar to every healthcare system. Table 1 has shown the functionality comparison between the WI Care with other existing healthcare system. The main focus of the WI Care is to enhance customer experience towards the appointment services. Ensuring every record is all-rounded standby with privacy protection and reduce the possibility of network congestion. WI Care has been developed with Visual Studio Code based on HTML, CSS and JavaScript. Front-end design includes the Main Portal which stores information, and the back-end design is developed under PHP syntax where the files is allocated under XAMPP folder that includes Login portal, Appointment Portal and Admin Portal. The main composition of WI Care is not just about the system. It prioritizes social media aspect that came along with automation that will came as the future development that will be included for later version release. For the main elements, this project came with a combination of four portals that make every user access to the system in different perspective. The system will also include different types of access levels for admins and doctors to create account and manage the dashboard, edit appointment list and also set schedule time. Appointer can have access information about WI Care services and also create appointment between doctor and appointer. Database will be stored in local server which contains admin and doctor accounts, and also capable of receive appointment form data.

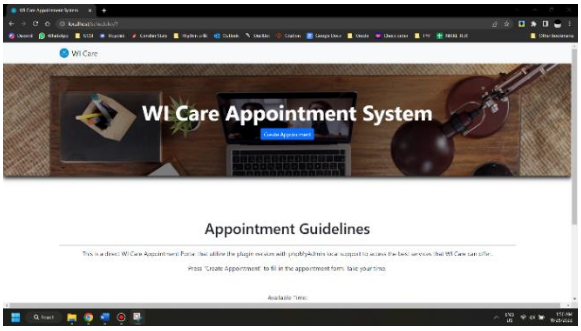
Table 1: Comparison of smart healthcare systems

Applications	Anura ™	SmartHealth (Android)	Sense Care	WI Care
IoMT sensor support			✓	
Online appointment			✓	✓
Account management	✓	✓	✓	✓
Store information and medical records	✓	✓	✓	✓
A.I. Processing	✓		✓	
Mobile Apps	✓	✓		

3. Results and Discussion

3.1. *Main Portal*

As for the back-end development of WI Care, two appointment portals were developed for admin, doctors and appointer access which can be found in Figs. 1 and 2.



This is a direct WI Care Appointment Portal that utilise the plugin version with phpMyAdmin local support to access the best services that WI Care can offer.

Press 'Create Appointment' to fill in the appointment form. Take your time.

Available Time:

Fig. 1: Appointment Portal

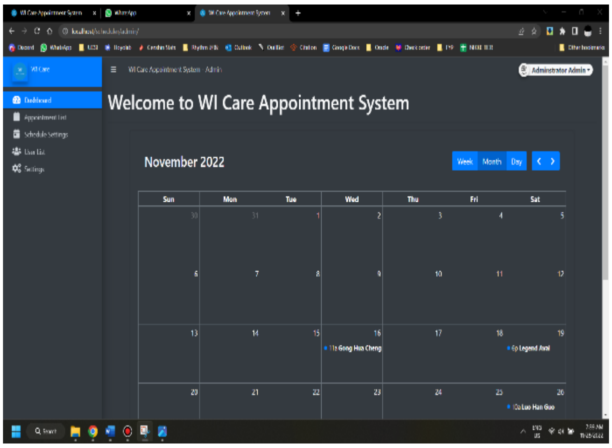

WI Care
WI Care Appointment System · Admin
Administrator Admin
Dashboard
Appointment List
Schedule Settings
User List
Settings
Welcome to WI Care Appointment System
November 2022
Week
Month
Day
Sun
Mon
Tue
Wed
Thu
Fri
Sat

Fig. 2: Admin Portal

As for the front-end development of WI Care, three webpage tabs were designed for WI Care main portal information display to let user understand clearly regarding the background of WI Care services and system. The sample are shown in Figs. 3, 4 and 5.

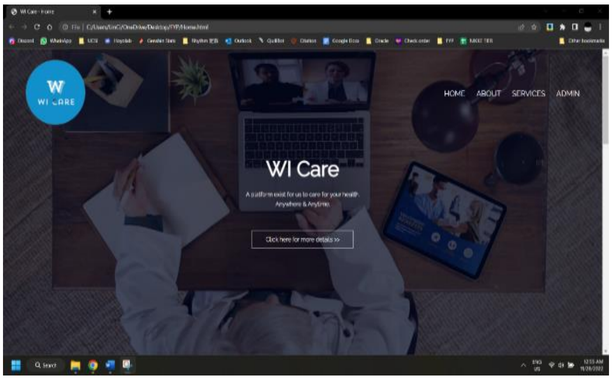

W
WI CARE
HOME
ABOUT
SERVICES
ADMIN
WI Care
A platform exist for us to care for your health Anywhere & Anytime
Click here for more details >>

Fig. 3: Home Tab

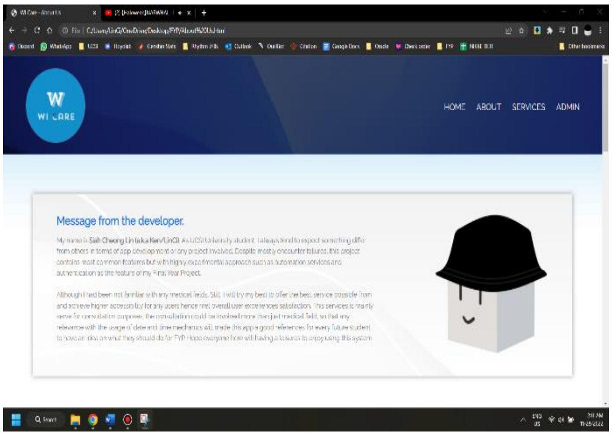

W
WI CARE
HOME
ABOUT
SERVICES
ADMIN
Message from the developer.

Fig. 4: About Tab

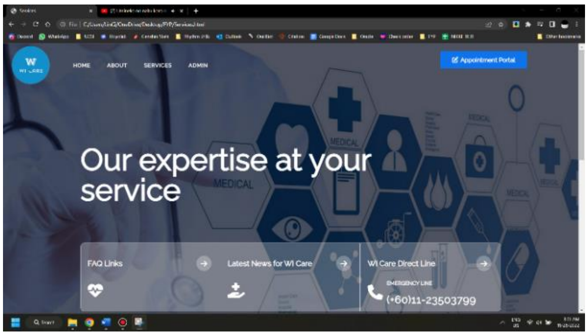
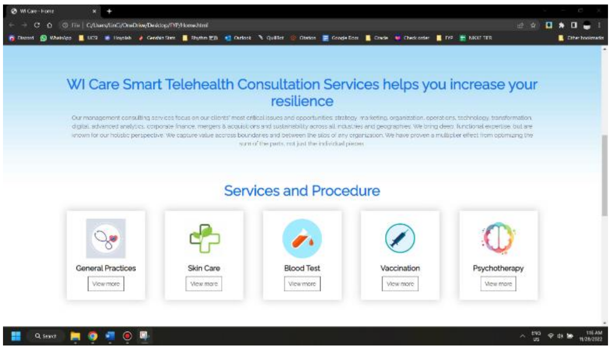

W
WI CARE
HOME
ABOUT
SERVICES
ADMIN
Appointment Portal
Our expertise at your service
FAQ Links
Latest News for WI Care
WI Care Direct Line
EMERGENCY LINE
(+60)11-23503799

Fig. 5: Service Tab


WI Care Smart Telehealth Consultation Services helps you increase your resilience

Services and Procedure
General Practices
Skin Care
Blood Test
Vaccination
Psychotherapy
View more

Fig. 6: Main Portal – Home Tab (Content: Introduction / Service and Procedure)

Fig. 6 shows the Home Tab contains the introduction of WI Care and the provided services included. The services such as General Practices, Skin Care, Blood Test, Vaccination and lastly Psychotherapy. “View more” is a button that will navigate to the About Tab which doesn’t allocate at Home Tab, as selecting “View More”, the page will be navigate to About Tab and scroll down to the Service Description section. This applies to other 4 services as well which can be referred from Fig. 7. Whereas after selecting “View More” on Skin Care section, it will automatically lead to the same section on Fig. 7 and automatically show the Skin Care Info on Fig. 8.

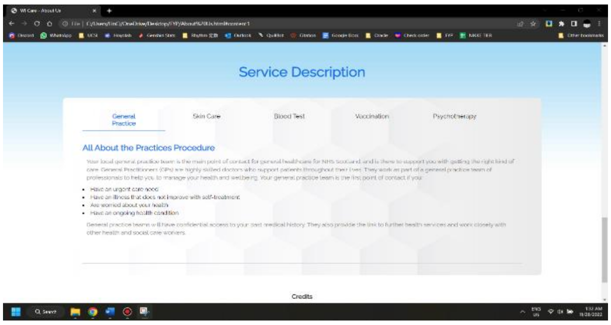
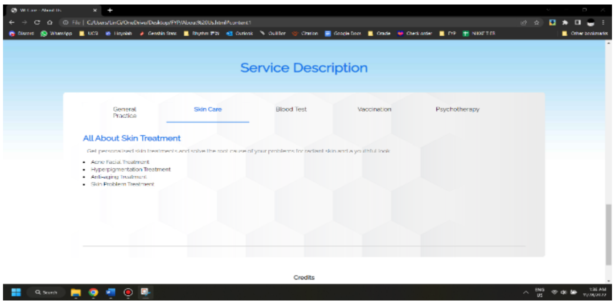

Service Description
General Practice
Skin Care
Blood Test
Vaccination
Psychotherapy
All About the Practices Procedure
Credits

Fig. 7: Main Portal – About Tab (Content: Service Description – General Practice)

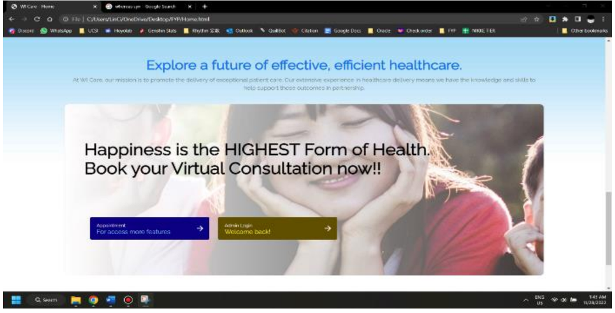

Service Description
General Practice
Skin Care
Blood Test
Vaccination
Psychotherapy
All About Skin Treatment
Credits

Get personalized skin treatments and solve the root cause of your problems for radiant skin and a youthful look

- Acne Facial Treatment
- Hyperpigmentation Treatment
- Anti-aging Treatment
- Skin Problem Treatment

Fig. 8: Main Portal – About Tab (Content: Service Description – Skin Care)

After Service and Procedure section, at the ending section for the home Tab display one white gradient background image which contains 2 buttons, the buttons are “Appointment” in blue color and “Admin Login” in yellow color which can be refer from Fig. 9.


Explore a future of effective, efficient healthcare.
Happiness is the HIGHEST Form of Health.
Book your Virtual Consultation now!!
Appointment
For access more features
Admin Login
Welcome back!

At WI Care, our mission is to promote the delivery of exceptional patient care. Our extensive experience in healthcare delivery means we have the knowledge and skills to help support these outcomes in partnership.

Fig. 9: Main Portal – Home Tab (Content: Ending section)

3.2. *Appointment Portal*

Appointment Portal and Main portal are the portals can be access from any user. Appointment Portal is also the only portal can be access from other 3 portals. The summarize webpage design of this portal can be referred back to Fig. 10.

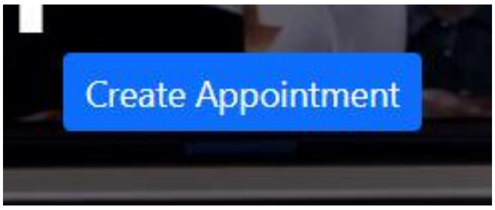

Create Appointment

Fig. 10: Appointment Portal – Create Appointment Button

“Create Appointment” is the only interactable button within the portal. After select the button, an appointment modal form will appear for the user, refer from Fig. 11.

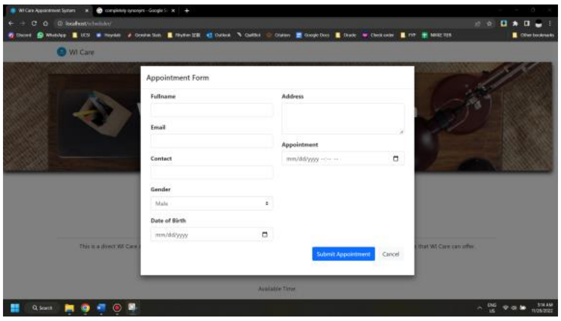

Appointment Form
Fullname
Address
Email
Appointment
mm/dd/yyyy --:-- --
Contact
Gender
Male
Date of Birth
mm/dd/yyyy
Submit Appointment
Cancel
Available Time

Fig. 11: Appointment Form displayed in modal view

Appointment Form consist of 7 field needed to be fill in. User must input all of the field to submit the appointment. Full name, email, contact and address will be input using the text-box provided. For the gender, it will appear a dropdown list that shows male or female, users are require to select one of them to complete the input, refer from Fig. 12.

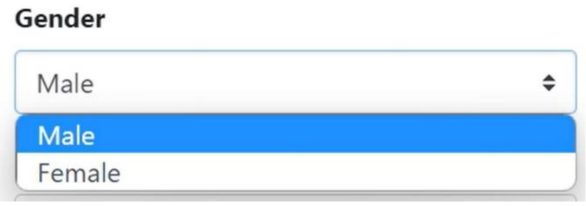
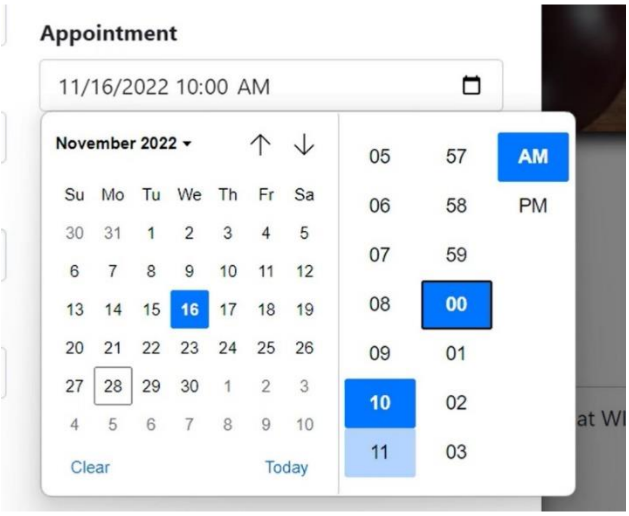

Gender
Male
Male
Female

Fig. 12: Appointment form – Gender

As for the date of birth and appointment, user can either type in as normal text-field or press the calendar logo to show the calendar display and select one of the dates which will automatically fill in the result after the selection, referring from Fig. 13. Additionally, the calendar can be toggle months using up and down arrows. User can select anywhere outside of calendar display to close the calendar.

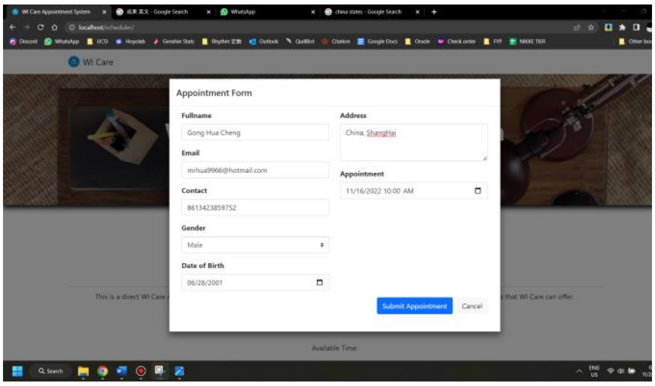


Appointment

11/16/2022 10:00 AM

Fig. 13: Appointment Form – Calendar View

After all of the input has been fill in, the result of the input should be shown as Fig. 14 displays.



Appointment Form

Fullname: Gong Hua Cheng

Email: mrhua9966@hotmail.com

Contact: 8613423859752

Gender: Male

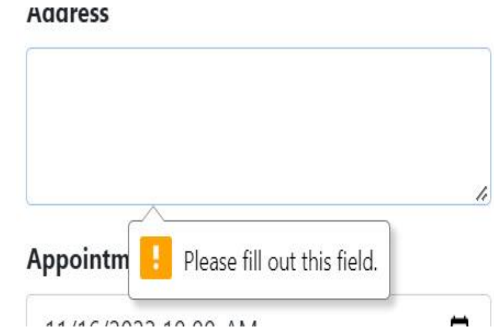
Date of Birth: 06/28/2001

Address: China, ShangHai

Appointment: 11/16/2022 10:00 AM

Fig. 14: Input result of Appointment Form

Selecting “Submit Appointment” button will have three types of conditions occurred, if users not complete fill in all of the require inputs, the system will pop-up message informed user to fill in the blanks, referring from Fig. 12. The second condition is if the schedule time does not meet with the time available, another error message will show on top of the form prompting the schedule time is invalid, referring from Fig. 13. And the last condition is if all of the input success, after select “Submit Appointment” will automatically close the form and display submit successful message on the top right corner of the webpage, referring from Figs. 14,15, 16 and 17. Selecting “Cancel” will close the Appointment Form.

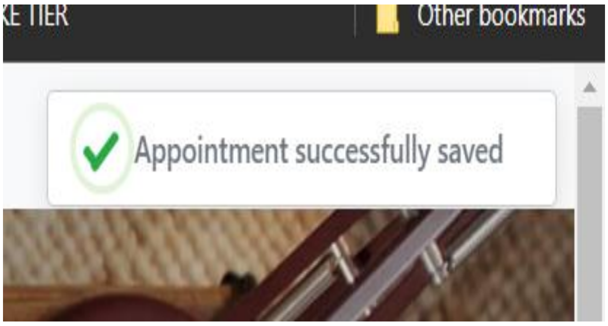


Address

Appointm

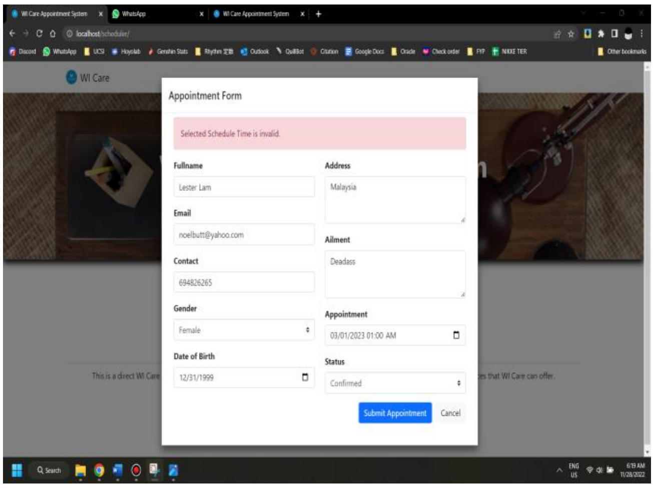
Please fill out this field.

Fig. 15: Submit Appointment form



Appointment successfully saved

Fig. 16: Appointment Form submission successful fill in the input field.



Appointment Form

Selected Schedule Time is invalid.

Fullname: Lester Lam

Email: noelbutt@yahoo.com

Contact: 694826265

Gender: Female

Date of Birth: 12/31/1999

Address: Malaysia

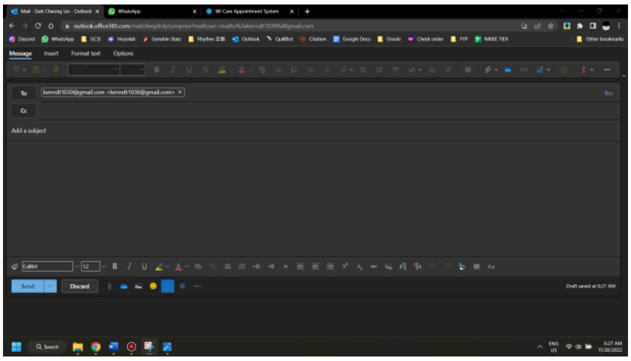
Ailment: Deadass

Appointment: 03/01/2023 01:00 AM

Status: Confirmed

Fig. 17: Submit Appointment with appointment date that is not active.

Scrolling down to the footer of the Appointment Portal will found a blue text which is a button that leads user to Microsoft Outlook and automatically fill in the developer’s email whereas user can reply for any system error occur, referring from Fig. 18 And so, this has concluded the Appointment Portal available interaction.



To: kenndt1030@gmail.com <kenndt1030@gmail.com>

Mail - Siah Cheong Lin - Outlook
WhatsApp
WI Care Appointment System
outlook.office365.com/mail/deeplink/compose?mailtouri=mailto%3akenndt1030%40gmail.com
Message
Insert
Format text
Options
Cc
Bcc
Add a subject
Send
Discard

Fig. 18: Microsoft Outlook with developers email as receiver.

WI Care appointment system are not limited to medical field. Any system that involves appointment can utilize our WI Care system and integrated as part of their main feature as well. WI Care also do benefit towards our target audience. With this system, most of the doctor should get significantly better management time on report access, IT usage quality and most importantly, the better health-care self-evaluation.

4. Conclusion

Most of the existing health-care system prioritize IoT modules for screening purpose or sensor [5] to extract the real-time data. By thinking back to square one, services are definitely one of the important elements in healthcare, and not everybody will have the financial ability to utilize those new technology due to absurd charges even just one time of screening. WI Care might not have the most advanced technology on hand, as trade off we had successfully developed a website that attracts most of the medical specialist. The website had released appointment portals is to ensure to reduce the workload especially for those GPs that prefer work individually.

As the appointment involved GPs, appointers are also the end-user that use the appointment system interact with GPs. Despite appointers only need to submit a form, and so besides submission, it is important to let appointers to join WI Care as part of community instead of just merely customers that come once and never return. So, within the website, I have established several information to let the users track our team progress on what our vision and mission is all about, which includes our medical services as well.

Lastly, only 57% of the adults using desktop or laptop compared to 97% of adults that uses smartphone as their main source of connection between other people [6]. It is important that WI Care system should enable phone user to have an access for all function, as I had implemented multiple framework that supports vertical view and buttons that able to hide certain views to fit the vertical screen, especially applies for the appointment system. Although overall system is not robust enough, but by achieving all of these objectives, which leaves the very first version of WI Care provides a good experience for accessing more opportunities to study, develop and learn from other developers as part of our milestone to become a better developer.

References

1. Dyzma, M. & Netguru. (2022, April 26). 13 Types of Healthcare Software. Netguru. Retrieved June 10, 2022.
2. Elezaj, R. (2018, October 22). How technology has changed the world of medicine. Geospatial World. Retrieved May 27, 2022.
3. Betts, D. (2020, August 12). Are consumers already living the future of health? Deloitte Insights. Retrieved May 27, 2022.
4. Gambino, E. (2022, March 8). 5 Examples of Smart Technology in Healthcare. Impact Networking. Retrieved May 30, 2022.
5. Al-Mejrad, A. S. (2021). Design and Implementation of in-Home Real-Time Wireless Fever Monitoring in Pediatrics during the COVID-19 Pandemic. Journal of Robotics, Networking and Artificial Life, 8(1), 37-40.
6. S.V. Balshetwar & Kiran V Shinde . (2021). Smart Health Consulting Android Applications. Journal of Emerging Technologies and Innovative Research, 8(7), 361-364.

Authors Introduction

Mr. Siah Cheong Lin

He is pursuing his Bachelor of Computer Science (Honours), UCSI University, Malaysia, currently.

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. Thong Chee Ling

Chee Ling Thong is an Associate Professor at Institute of Computer Science and Digital Innovation, UCSI University. Her research interests include IT Innovation in solving real-life problem. She has received three research grants from UCSI University Malaysia as Principal Investigator for projects entitled ‘Android-based Time Tracker for Shuttle Bus ‘, ‘iOS Time and Location Tracker for Shuttle Bus’ and ‘Mobile Car App for Travel Agency’. These projects are fully copyrighted by Intellectual Property Corporation of Malaysia (MyIPO). In 2021, Thong was awarded the "Product Innovation and Commercialization Award" by the university.

Web-based Stocktaking application in Businesses

Cheah Dei Xuan, Heshalini Rajagopal, Shayla Islam
Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia

E-mail: heshalini@ucsiuniversity.edu.my, shayla@ucsiuniversity.edu.my

www.ucsiuniversity.edu.my

Abstract

As the impact of the technology hits towards the business's aspect into digitalized, its direct things easily and manageable which it able to reduce the workload. The purpose of this report was to do research and develop a web-based application on business stock checking where workers able to track their quantity of their stock from running out of stock as well as to track the business statistic. This application was followed SDLC model to develop while Vue 2 framework, Java script language, html 5, and tailwind CSS to code the web-based application. In this study, a database was developed for shopkeeper to store all the data related to the shop on cloud-based that able to prevent from disaster or human errors. Furthermore, a web-based system was developed to perform stock checking and monitoring which allow shopkeeper to check on statistic of the product and also the notify shopkeeper which stock are running low. In addition, a softcopy receipt was implemented as a default for customer instant of hardcopy where customer tossed the receipt, and hardcopy receipt are not recyclable due to some chemical substance. The contribution of the web-based application where shopkeeper able to keep track on their business status, preventing product from lacking stock, as well as to reduce the usage of paper.

Keywords: Stocktaking; Web-based application; E-Receipt

1. Introduction

Technology has changed the way how we complete a certain task. Interacting with technology improves the effectiveness and efficiency of the workload, as the world is moving toward industry 4.0 technology many sectors are relied [1]. Because of the increase of innovation coming from different platforms and providers making all this possible. Businesses either big business or startup business needs technology to carry on the task. While in this research it studies and develop a web-based stock checking system that will bring an innovative way for shopkeepers, or the owners use to check their stock in the store. The importance of checking the stock was to meet the supply of the product and the demand of the customer. From here the term supply pertains to how much the producers of a product or service are willing to produce and can provide to the market with the limited number of resources available [2]. Whereas term demand is how much of that product or service the buyers desire to have from the market [2]. The other importance of checking the stock was to prevent losses or stealing of stock that might lead to the loss of business assets. Checking the stock able to prevent overstocking as well.

Stocktaking in a more understandable way is known as stock counting, basically stocktaking in the count and check all the stock shopkeeper or the owner of the shop had purchases and the total amount they sell the goods and the amount they earn [3]. Businesses who did not do frequently on their stock will be having a lot of issues like loss of stocks, issue of financial, etc. hence stocktaking plays an important role in businesses as it brings a lot of benefits towards the business growth and prevents unwanted problems like over-stocking, able to identify damaged goods, easy to track from theft.

Back in those days when technology was not that advanced, checking stock was a troublesome task for business owners or shopkeepers where they have to count one by one and record it down in a record book. Slowly the apparency of the computer was out shopkeeper or owner of the shop type out all goods information one by one and the quantity of it. Once the goods have been sold out, they have to when back to the computer and make the changes to it. After when the system application and barcode were introduced, everything changed toward a more convenient on finishing the task. The first barcode was designed to encode the product information into bars to quickly identify that product it is an identity for the product after the barcode has been introduced all the product packaging will be labeled with a barcode and it became the identity of the product itself. Ever wonder why barcodes need to be designed that way example Fig. 1 shows the representation of numbers, thus it combined the number and convert it into the barcode and that's how the barcode was designed. While it also has its own sequence of combining the barcode which comprises of the quiet zone, number system digit, manufacture code, product code, and check digit. Each of the components plays an important role [4].

Barcodes separate into two different dimensional which is the one dimensional (1D) and two-dimensional (2D). In barcode one dimensional (1D) is in a systematic manner represents data by varying the width and spacing of parallel lines such as universal product code, European article number, code 39, code 128, Codabar, etc. [5]. On the other hand, a two-dimensional barcode (2D) represents data using two-dimensional symbols and shapes such as the QR code, Data Matrix code, etc. [5].

The motivation of this study is in businesses like Retail business selling products most of the shopkeepers or owners of the business will set up or purchase a set of expensive software and hardware who are also known as POS system for them to work which bring a lot of benefits to the shopkeeper and for the customer. While to have this setup it needs a large amount of cost, constantly upgrades and updates of the system are needed, costly hardware, etc. Some start-up businesses, can't effort a big amount of expense on those systems [6]. Hence, it's hard for startup businesses or small businesses to track the quantity of the goods they have and because of that, they are unable to trace the supply and demands to cater to their customer as tracing the business statics is very important to avoid issues like over-stocking. Studies show that lot of the small business fails is because lack of capital, lack of technology, lack of financial planning, etc. [6]. Hence this study aimed to develop a web-based application that able to work on multiple devices as it is a web-based application to help businesses track the quantity of goods in the shops to cater the supply and demands of customer as well as to track their business status.

2. Methods

This application was followed SDLC model to develop while Vue 2 framework, Java script language, html 5, and tailwind CSS to code the web-based application. In this web-based application the database will be using google Firebase as the main database as a JSON database. Every user has its dedicated data to manage and do changes to it others user is unable to access or do changes to it as they have no access to the data. Once the user was created the application will straight created few subcollection into the database with some example data. The sub-collection consists of account, categories, restock, sales, and products.

Shopkeepers first, can add all the information regarding the product so that the system has a record of this product for future use. Fig. 1 shows the flowchart of shopkeeper adding products. As the shopkeeper adds the product to the system shopkeeper can check on the quantity left for the product. If the quantity of the product is left to a certain amount the system will notify the shopkeeper that this product only left a few quantities. This system is also able to provide shopkeepers to keep track of their business status as well.

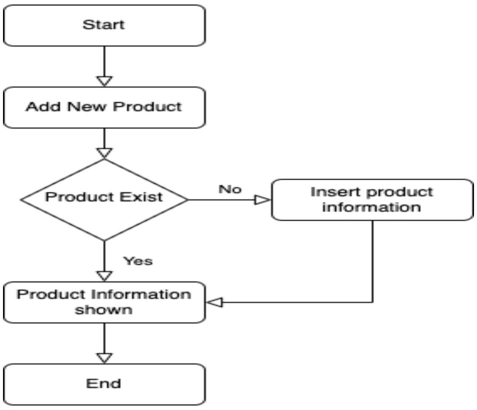


Fig. 1: Shopkeeper Add Product

The system provides shopkeepers to do cashier features like calculating the total amount of customer baskets and generating receipts for the customer. Fig. 2 shows that the flowchart of cashier. First the shopkeeper scan customer QR and the system would show the list of the customer shopping cart, follow by customer make payment if checkout denied it will straight end the process, if the check-out completed it will generate receipts and send the receipt to customer email if the customer provides an email address and the process end. Fig. 3 is the flowchart of shopkeeper do restock on their product. Where this feature was to provide shopkeeper to have an easy way to restock the stock instant of find the stock from the stock list and make edit on it. Hence, it runs with shopkeeper scan product barcode and the application will search for the product, if no product found shopkeeper must scan again, if the product found all the information regarding to the product will be listed out. after that shopkeeper fill up the details and application will update the database. If the update not success it will back to the input restock details, if the update success the process will end here and the database will be updated.

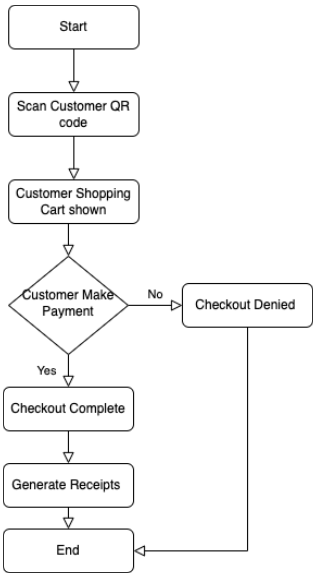

Start
Scan Customer QR code
Customer Shopping Cart shown
Customer Make Payment
No
Checkout Denied
Yes
Checkout Complete
Generate Receipts
End

Fig. 2: Shopkeeper Cashier Flowchart

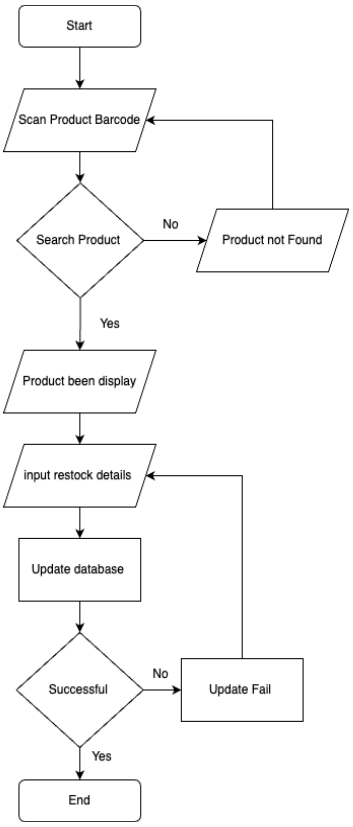

Start
Scan Product Barcode
Search Product
No
Product not Found
Yes
Product been display
input restock details
Update database
Successful
No
Update Fail
Yes
End

Fig. 3: Restock

The flowchart of Statistic is where the application generates a T-account for shopkeeper. As shown in Fig. 4, the first the application required shopkeeper to input the range of start date and end data to get data from the range and make calculation. Once the input is done all the details will be shown. There is a field that allow shopkeeper to insert other expenses during that period of time, if there are no expenses during the period of time the application will lead to the next process. At the following process is the whole t-account record saved to the database and a excel file is generate download form the page. The whole system will end once the record is saved and excel file are generated and downloaded to shopkeeper devices.

The customer can scan the product and the customer can get all the information regarding the product. Customers can add the product to the system cart once the customer is done with his or her shopping to checkout customer only need to show the generated QR by the system to the shopkeeper to scan and make payment as shown in Fig. 5.

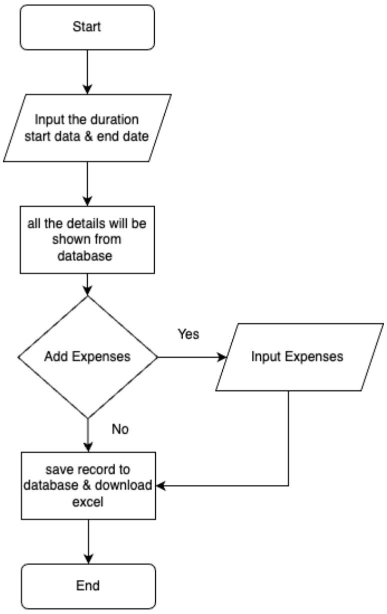

Start
Input the duration start data & end date
all the details will be shown from database
Add Expenses
Yes
Input Expenses
No
save record to database & download excel
End

Fig. 4: Statistic

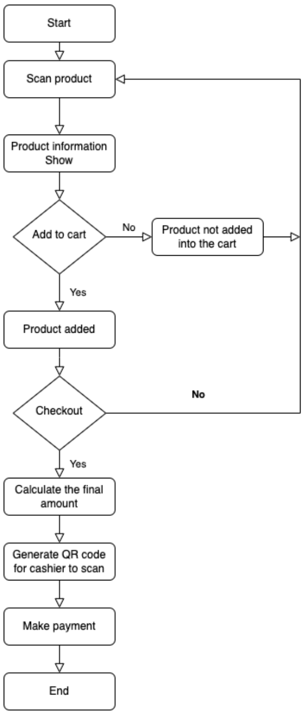

Start
Scan product
Product information Show
Add to cart
No
Product not added into the cart
Yes
Product added
Checkout
No
Yes
Calculate the final amount
Generate QR code for cashier to scan
Make payment
End

Fig. 5: Customer Flowchart

3. Results and Discussion

Inside the web-based application it came along with multiple features which make the whole application able to bring lots of benefits towards shopkeeper. Will be using the UI and descript more towards the application features. Fig. 6 is the main dashboard that shopkeeper will first view after signing in the application. In dashboard page, basically it separates into 4 row that shows everything regarding the business information all will be shown in a summary way. On the first row, there will be a summary of the shop name, shopkeeper while it can be change at the account page and it also have a "sale of the day" which it shows the total of sales that earn on that particular day. Shopkeeper able to download the list of the sales into excel files and the files will be name exactly the date as a record for future use on calculating the account. On the second row is the list of the products that shopkeeper added into the application. to add a product shopkeeper will need to go to the add product page. After the shopkeeper added the product into the application, all the products will be shown in the list. There is also a category list which help shopkeeper on quick filter on the list to search for the products. As product more and more added into the application this category list will be a good feature for shopkeeper. On the third row is the list of the sales of paid and unpaid, while on the right side is the chart to view the sales statistic from which date to which date and the application will generate into a bar-chart for shopkeeper to view and do compare.

Last row, which is a QR code been generated, and it is for the customer enter to the shop to do scanning of the product in the shop do shopping as well as to make payment through it. This generated QR code contain a link for customer to scan and it direct customer to the page which is the customer page follow by the shop identity key (ID). Fig. 7 is the example of generated QR code.

Fig. 8 shows the User interface of the Add product page, it is a form where shopkeeper required to key in all the details for the particular product include the product name, buying price, selling price, quantity, product image, and the barcode of the product. Application will provide a button for shopkeeper to use device camera to scan the product and the barcode will be shown on the form, Fig. 9 is the example of the camera scanning the product barcode and the result. Once all the fields were filled with the details, shopkeeper click on the submit button and the product will be added.

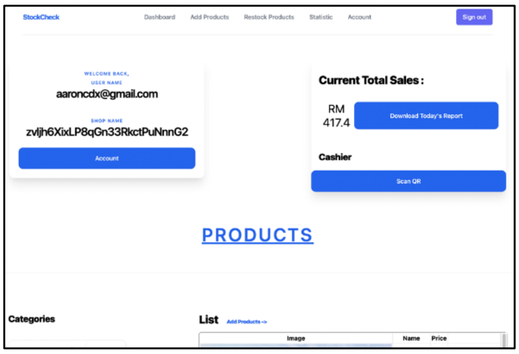

StockCheck
Dashboard
Add Products
Restock Products
Statistic
Account
Sign out
WELCOME BACK,
USER NAME
aaroncdx@gmail.com
SHOP NAME
zvljh6XixLP8qGn33RkctPuNnnG2
Account
Current Total Sales :
RM
417.4
Download Today's Report
Cashier
Scan QR
PRODUCTS
Categories
List
Add Products ->
Image
Name
Price

Fig. 6: Dashboard of the Stock-Checking application

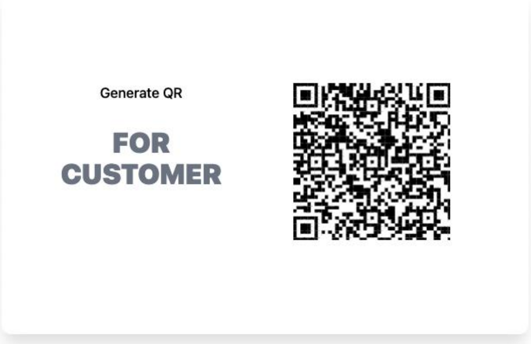

Generate QR
FOR
CUSTOMER

Fig. 7: Example of Generated QR Code

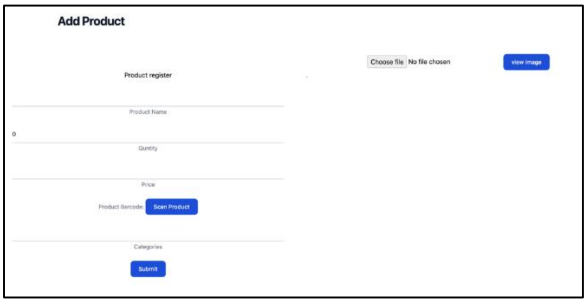

Add Product
Product register
Choose file
No file chosen
view image
Product Name
0
Qunity
Price
Product Barcode
Scan Product
Categories
Submit

Fig. 8: Add Product Page

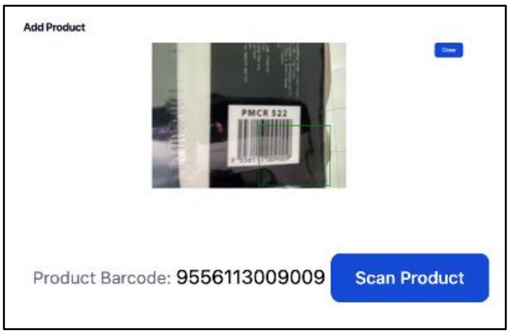

Add Product
PMCR 522
Clear
Product Barcode: 9556113009009
Scan Product

Fig. 9: Example of the Scanning process and the result

The restock page is when a particular product running out of quantity shopkeeper went. In this page shopkeeper just need to scan the product using the device camera and all the details regarding to the product pops up. This feature was to allow shopkeeper to have an easier way to update the stock quantity rather than find the particular product from the product list and made changes on to it. By using this features shopkeeper straight away can just update the quantity as well as all the restock data will be storage in a separate collection in database follow by all the details like buying price of the product. Hence, for the application to calculate the T account is more accurate as it sums up all the buying price onto the credit side. Fig. 10 example of the restock page user interface.

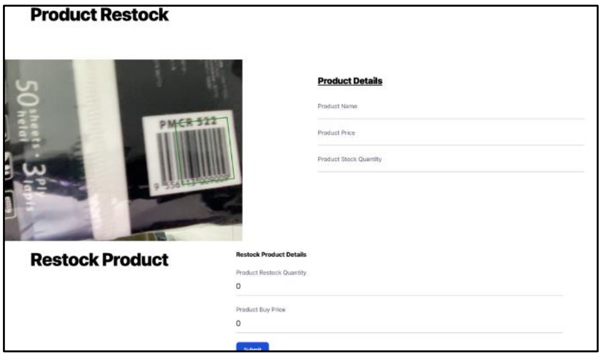
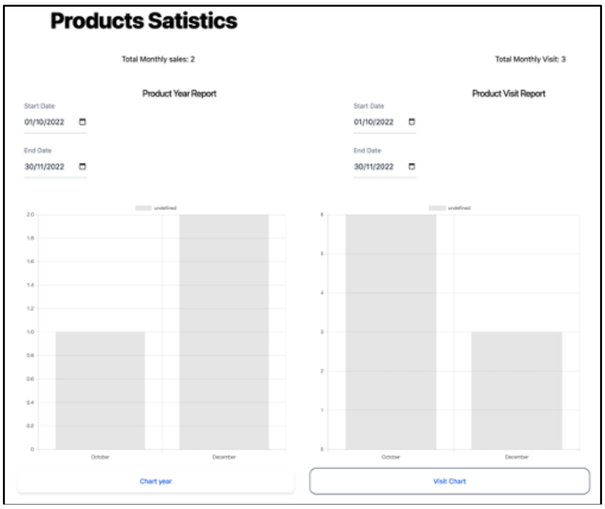

Product Restock
PMCR 522
Product Details
Product Name
Product Price
Product Stock Quantity
Restock Product
Restock Product Details
Product Restock Quantity
0
Product Buy Price
0

Fig. 10: Example of Restock Page UI

In this application, it allows shopkeeper to make fully customizable towards shopkeeper data. Hence product details page is where to do some little changes on the product details. Application allows it to delete the product as well. While below it shows the number of quantities that sell and the number of visit that this product has in the current month. Once the customer takes the product scan and view the details it increases the visit number, as for the number of quantity sell is where the product been sold by customer it increases.

There are two empty charts are ready with the start date and end date field for shopkeeper to track on the particular product from time to time. With this statistic features allow shopkeeper to make decision on the number of bringing the product should increase or to decrease. Fig. 11 shows the example result of the bar chart statistic for few months generate by the application.



Products Satistics

Total Monthly sales: 2

Total Monthly Visit: 3

Product Year Report

Start Date 01/10/2022

End Date 30/11/2022

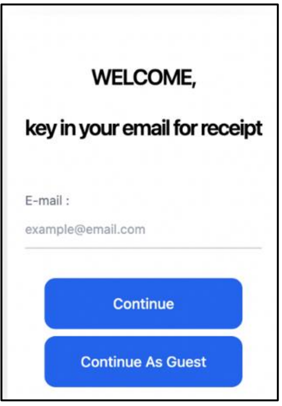
Product Visit Report

Start Date 01/10/2022

End Date 30/11/2022

Fig. 11: Example of generated Bar Chart Statistic

Stock Checking web-based Application it allows customer to interact with the application as well. As mentioned, shopkeeper dashboard has generated an QR code for customer to scan and enter to their shop to do shopping. Fig. 12 is the first view where customer enter, this is where customer has option to key in their email address or continue as guest. Customer who keys in their email address able to get an email receipt. After selecting the option, customer can do scanning on the product to view all the details regarding the product specially to look at the price for the product. It will pop out a data-toggle model to show all the product details as shown in Fig. 13.

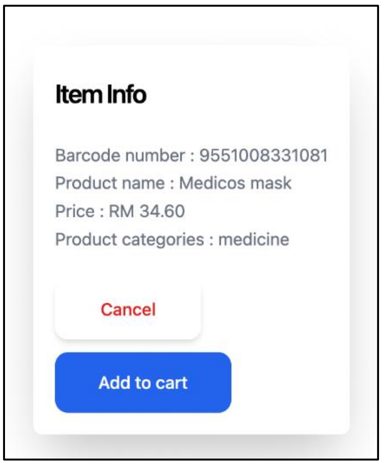


WELCOME,

key in your email for receipt

E-mail : example@email.com

Fig. 12: Customer page UI



Item Info

Barcode number : 9551008331081

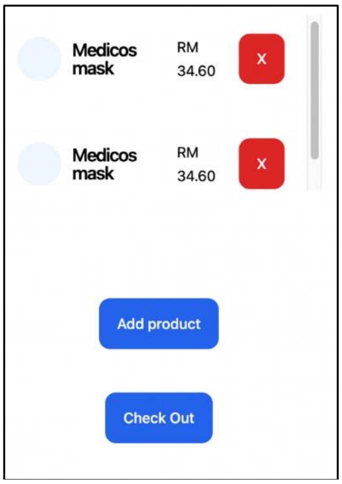
Product name : Medicos mask

Price : RM 34.60

Product categories : medicine

Fig. 7: Example of Product Details Data-Toggle

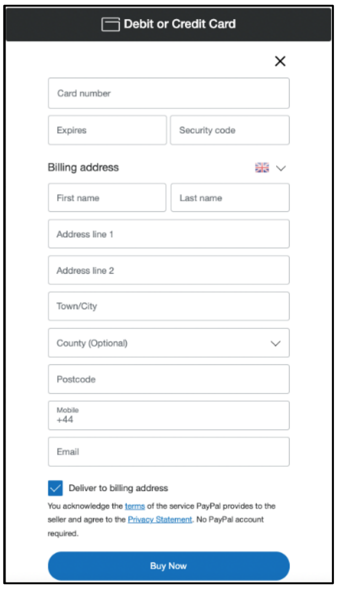
Customer are allowed to add the product to the shopping-cart list and will be shown in Fig. 14. After done shopping, customer can click on the "checkout" button and the application will summarize the shopping cart and continue to make payment. While in payment, customer have 3 option to make the payment which is the pay by cash, e-wallet, or online-banking. For cash payment and e-wallet a QR code were generate and customer have to go to the cashier and show the QR code to the shopkeeper to scan. Once the shopkeeper scanned the application will lead the shopkeeper to the cashier page which will be discuss later on. where for online-banking, application will lead customer to another page to do online-banking by Pay-pal as shown in Fig. 15.



Medicos mask RM 34.60

Medicos mask RM 34.60

Fig. 14: Example of the Shopping-Cart List

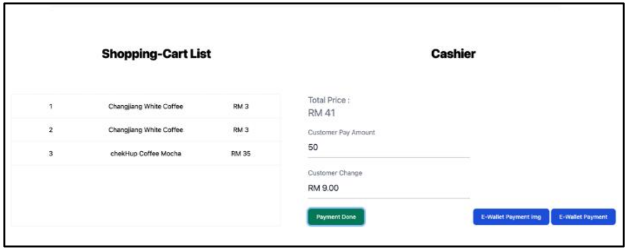


☑ Deliver to billing address

You acknowledge the terms of the service PayPal provides to the seller and agree to the Privacy Statement. No PayPal account required.

Fig. 15: Example of Online-Banking Payment

Customer who wants to pay by cash or e-wallet they are required to provide the QR code to the shopkeeper to scan. As for the shopkeeper, the button to scan the QR code is locate at the dashboard. Once shopkeeper click on the “scan QR” it opens the device camera to scan customer QR code and the application redirect the page to the cashier page as shown in Fig. 16.



1	Changjiang White Coffee	RM 3
2	Changjiang White Coffee	RM 3
3	chekHup Coffee Mocha	RM 35

Total Price :
RM 41

Customer Pay Amount
50

Customer Change
RM 9.00

Fig. 16: Cashier Page UI

Once all the payment has made and the transaction is complete the database will update the status to “paid”. For customer who insert their email address will immediately receive an Email receipt regarding the item that customer has purchased.

4. Conclusion

In this study, a database was developed for shopkeeper to store all the data related to the shop on cloud-based that able to prevent from disaster or human errors. Next, a web-based system was developed to perform stock checking and monitoring which allow shopkeeper to check on statistic of the product and also the notify shopkeeper which stock are running low. The developed system also implemented a softcopy receipt as a default for customer instant of hardcopy where customer tossed the receipt, and hardcopy receipt are not recyclable due to some chemical substance. The contribution of the web-based application where shopkeeper able to keep track on their business status, preventing product from lacking stock, as well as to reduce the usage of paper. This system is beneficial and able provide an innovative way to interact and bring benefits towards to community.

References

1. Kumar A, Rajagopal H. Automated Seeding and Irrigation System using Arduino. Journal of Robotics, Networking and Artificial Life, 2022, 8(4): 259–262.
2. J. Fernando, “Understanding The Law of Supply And Demand,” Investopedia, 08-Feb-2022.
3. G. M. Cidal, Y. A. Cimbek, G. Karahan, O. E. Boler, O. Ozkardesler, and H. Uvet, “A Study On The Development of Semi Automated Warehouse Stock Counting System,” 2019 6th International Conference on Electrical and Electronics Engineering (ICEEE), 2019.
4. Milind Amrutkar, Anup Palsokar, Pankaj Raibagkar, “QR Code Based Stock Management System” (IRJET).
5. R. Focardi, F. L. Luccio, and H. A. Wahsheh, “Security Threats And Solutions For Two-Dimensional Barcodes: A Comparative Study,” Computer and Network Security Essentials, pp. 207–219, 2017.
6. Ma’aji, M.M., Abdullah, N.A.H. and Khaw, K.L.H., 2018. Predicting financial distress among SMEs in Malaysia. *European Scientific Journal, ESJ*, *14*(7), pp.91-102.

Authors Introduction

Mr.Cheah Dei Xuan

He is pursuing his Bachelor of Computer Science (Hons) Mobile Computing and Networking, UCSI University, Malaysia, currently.

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. Shayla Islam

Currently, she is an Associate Professor in UCSI University, Malaysia. She has completed PhD degree in Engineering from Electrical and Computer Engineering (ECE) departmentat International Islamic University Malaysia (IIUM) in 2016 under Malaysian International Scholarship (MIS). She completed her MSc.at the department of Electrical and Computer Engineering in InternationalIslamic University Malaysia in 2012. She completed her BSc. in ComputerScience and Engineering from International Islamic UniversityChittagong, Bangladesh. Dr. Shayla has awarded a Silver medal for herresearch work at International Islamic University Malaysia. In consequences, she has also awarded a Young Scientist Award for the contribution of research paper at 2nd International Conference on Green Computing and Engineering Technologies, 2016 (ICGCET'16), Organized by the Department of Energy Technology, Aalborg University, Esbjerg, Denmark. Her current researchinterests include data communications and networking, computer networks and wirelesscommunication, network mobility in heterogeneous network, computer architecture & cloudcomputing. She published more than 70 papers in international journals and conferences(WoS/ISI/ESCI and Scopus indexed). She works as a reviewer for many international conferencesas well as journals.

Pharmacy Warehouse Management System

Gan Jhui Ken, Heshalini Rajagopal, Shaik Shabana Anjum
Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia

E-mail: heshalini@ucsiuniversity.edu.my, shabana@ucsiuniversity.edu.my

www.ucsiuniversity.edu.my

Abstract

Managing of the stock and supply of medication plays an essential part for the provisioning of health care, thus the development of the web-based warehouse system. This project attempts to help gain a better understanding of the issue that the warehouse system is facing. And through the multiple sources based on the current warehouse management system, inefficiencies, namely the incorrect inventory management, can be solved. With the use of modern technology, the warehouse management system can be made better to help improve the quality of work for both the employees and admins. Additionally, this system will include the basic functions of a warehouse management system that will allow users to utilize the system with ease. The project also aims to improve the current pharmacy warehouse management systems so that it will be viable for users to understand and utilize the system efficiently.

Keywords: Stocktaking; Web-based application; Warehouse Management System

1. Introduction

Private pharmacies have started to create medicine delivery services and they must store their medicines at their own warehouse. During covid, most people can’t physically go to the pharmacy because they might get infected, and pharmacies must adapt to the pandemic. Therefore, private pharmacies began to create delivery services from their own store with the stored medicines from their storeroom. The medicine that can be delivered is mostly without prescriptions such as Paracetamol, Zyrtec, and Antibiotics. Their own warehouse does not have a warehouse management system thus it will cause confusion when shipping in and out of the warehouse.

The warehouse's function is to enable the storage of products in one area and the delivery of those things to the company's consumers. The type of company will determine how things can be ordered and transported from the manufacturers. The entire distribution process is made up of a number of smaller procedures that must all be closely managed. As a result, the Warehouse Management System, a software solution for warehouses, was developed. Distribution and manufacturing organizations will save a lot of money if they install the Warehouse Management System correctly since it will improve their entire work operations and provide essential business analysis [1].

The Warehouse Management System's goal is to manage the movement and storage of raw materials and finished goods in a warehouse. The Warehouse Management System also manages the transactions, shipping, receiving, put-away, and picking processes. Warehouse Management System (WMS) is a database-driven computer tool designed to increase warehouse productivity by guiding cutaways and maintaining extremely precise inventory by recording warehouse operations. Based on real-time data, the Warehouse Management System will also steer and optimize stocks. In order to better monitor the movement of raw materials and products, the Warehouse Management System will incorporate Auto ID Data Capture (AIDC) technology such as barcode scanners, personal computers, wireless local area networks, and radio-frequency identification (RFID). The central database will provide reports that are related to the products in the warehouse. The warehouse control system functions as information receiver from the upper-level host system. Upper-level host system, which is most often being the warehouse management system, will translate the information for the daily operations. The Warehouse Management System will usually manage processes and equipment on the operational level so that fewer errors will occur [2].

The benefits that Warehouse Management System (WMS) can bring to a business and organizations are greater cost savings, reduced errors and improved customer services. The benefits of implementing a Warehouse Management System are Real-time inventory management, where it manages the inventory of the warehouse much more easily and in real time. Warehouse Management System can support Logistics service provider but also the company's customers to plan the resources and inventory. Warehouse Management System also helps to reduce mishandling of inventory. This will allow each product to be scanned for every process at the warehouse, which includes inbound receiving and outbound processing. Inventory control will reduce the mishandling of inventory greatly thus improving efficiency. Efficient Returns Process Managing and handling the customer returns became easier and efficient, since it tracks the returned inventory at the detailed level [3][4][5].

Covid 19 has affected a lot of businesses and the pharmaceutical sector was greatly affected during the pandemic. The pandemic has caused social distancing among people which has caused people to not go and buy medicines from pharmacy physically as the fear or getting the disease. This has caused the sales of pharmacy to drop significantly as most people would prefer to order online. Private pharmacies have resorted to setting up an online medicine delivery service, but their warehouse has no system to manage it.

With the implementation of warehouse management system, private pharmacies will be able to manage their warehouses more efficiently and it allows less labor constraints. Therefore, this study aims to design and develop a pharmacy-based warehouse management system in order to deliver a system that can benefit the employees and admins.

2. Methods

The prototype system project will be utilizing various programming language such as Java and JavaScript. MySQL will be used for the components of the database. The flowchart diagram as shown in Fig. 1 has both the shipping in of products where it will be added by the employee and admin into the database system once its accepted. If the product is denied, the admin or employee must input the correct details. The shipping out of products will indicate the number of items that customers want, and the admin and employee must input the correct amount of product that customers want, or it will be denied by the system.

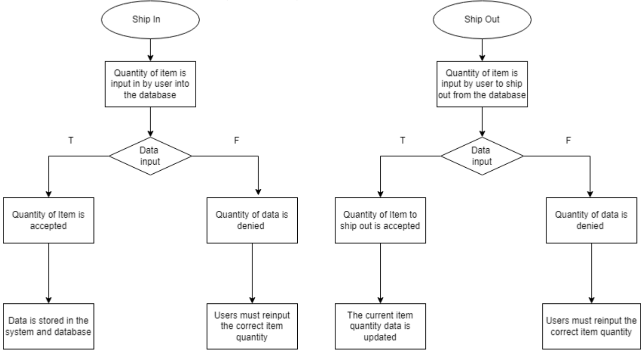
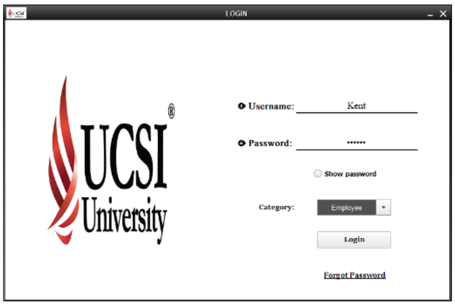

Ship In
Quantity of item is input in by user into the database
T
Data input
F
Quantity of Item is accepted
Quantity of data is denied
Data is stored in the system and database
Users must reinput the correct item quantity
Ship Out
Quantity of item is input by user to ship out from the database
T
Data input
F
Quantity of Item to ship out is accepted
Quantity of data is denied
The current item quantity data is updated
Users must reinput the correct item quantity

Fig. 1: Flowchart of the system

3. Results and Discussion

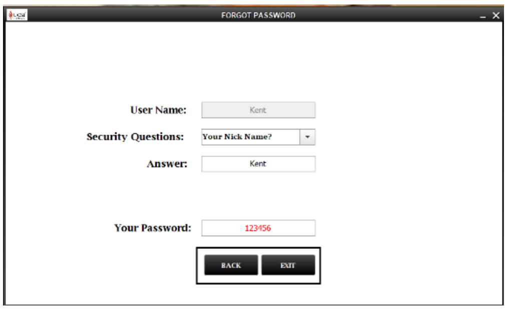
Fig. 2 shows the login page of the system which requires user to input their username and password. They must also choose the category that they belong to which is either employee or admin. After finishing inputting both username and password, they can proceed to click the login button to enter the system.

Fig. 3 shows the forgot password page which appears after user clicks the forgot password button. The user needs to fill in their username and choose the security question with the answer and input the new password to reset the account.


LOGIN
UCSI University
Show password
Login
Forgot Password

Username: Kent
Password: ••••••
Category: Employee

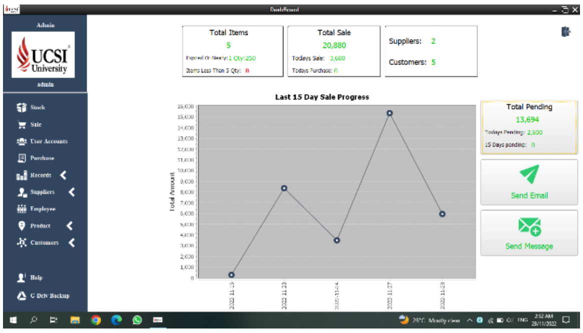
Fig. 2: Login Page


FORGOT PASSWORD
BACK
EXIT

User Name: Kent
Security Questions: Your Nick Name?
Answer: Kent
Your Password: 123456

Fig. 3: Forgot Password Page

Fig. 4 shows the dashboard of the overall system for the employees, the page contains a graph char to record the total amount of sales generated per day. It also shows the total items and total sales for the overall transactions. The page also shows the total number of suppliers and customers. The left panel of the system has stock button where employees can check the stock of items, sale button where employees can create sale for customer. Purchase button where employees can purchase products from suppliers or companies. Supplier button which allows employees to create suppliers to be added to the system. Product button which allows employee to add new product to the system and to check whether the product is expired or out of stock. Customer button where employees can add new customer into the system. And lastly, help button which the employee can contact the admin via their info.


DashBoard
Admin
UCSI University
admin
Stock
Sale
User Accounts
Purchase
Records
Suppliers
Employee
Product
Customers
Help
G Driv Backup
Last 15 Day Sale Progress
Total Amount
16,000
15,000
14,000
13,000
12,000
11,000
10,000
9,000
8,000
7,000
6,000
5,000
4,000
3,000
2,000
1,000
0
2022-11-15
2022-11-23
2022-11-24
2022-11-27
2022-11-28
Send Email
Send Message
25°C Mostly clear
ENG
2:52 AM
28/11/2022

Total Items: 5
Expired Or Nearly: 1 Qty:250
Items Less Than 5 Qty: 0

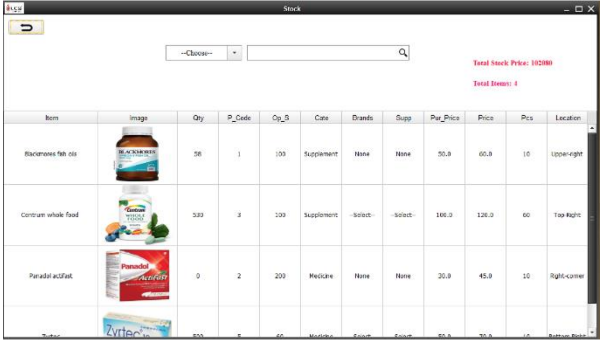
Total Sale: 20,880
Todays Sale: 3,600
Todays Purchase: 0

Suppliers: 2
Customers: 5

Total Pending: 13,694
Todays Pending: 2,600
15 Days pending: 0

Fig. 4: Dashboard Page

Fig. 5 shows the stock page of the system which both the employees and admin can check the stock or location of the products. The search bar on top allows both the employee and admin to type the product that they want to search for.

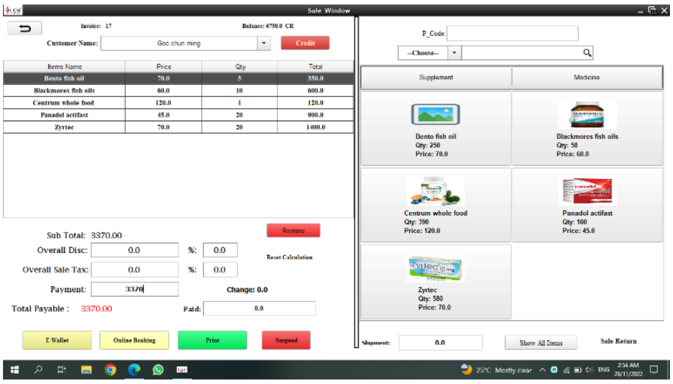

Stock
--Choose--

Total Stock Price: 102080
Total Items: 4

Item	Image	Qty	P_Code	Op_S	Cate	Brands	Supp	Pur_Price	Price	Pcs	Location
Blackmores fish oils		58	1	100	Supplement	None	None	50.0	60.0	10	Upper-right
Centrum whole food		530	3	100	Supplement	-Select-	-Select-	100.0	120.0	60	Top Right
Panadol actifast		0	2	200	Medicine	None	None	30.0	45.0	10	Right-corner

Fig. 5: Stock Page

The sale page (Fig. 6) shows the products that are sold, and both the employee and admin can select the customers that are registered in the system and proceed to sell them the item based on the quantity that the customer desires. Both employee and admin can opt to provided discount and charge additional sales tax to the customer. Once everything is inputted, the payment via E-Wallet or Online Banking, and it will be saved in the database as a transaction.


Sale_Window
Credit
Remove
Reset Calculation
E-Wallet
Online Banking
Print
Suspend
--Choose--
Supplement
Medicine
Show All Items
Sale Return
25°C Mostly clear
ENG
2:54 AM
28/11/2022

Invoice: 17
Balance: 6750.0 CR
Customer Name: Goo chun ming

Items Name	Price	Qty	Total
Bento fish oil	70.0	5	350.0
Blackmores fish oils	60.0	10	600.0
Centrum whole food	120.0	1	120.0
Panadol actifast	45.0	20	900.0
Zyrtec	70.0	20	1400.0

Sub Total: 3370.00
Overall Disc: 0.0 %: 0.0
Overall Sale Tax: 0.0 %: 0.0
Payment: 3370
Change: 0.0
Total Payable : 3370.00
Paid: 0.0

P_Code:

Bento fish oil
Qty: 250
Price: 70.0

Blackmores fish oils
Qty: 58
Price: 60.0

Centrum whole food
Qty: 390
Price: 120.0

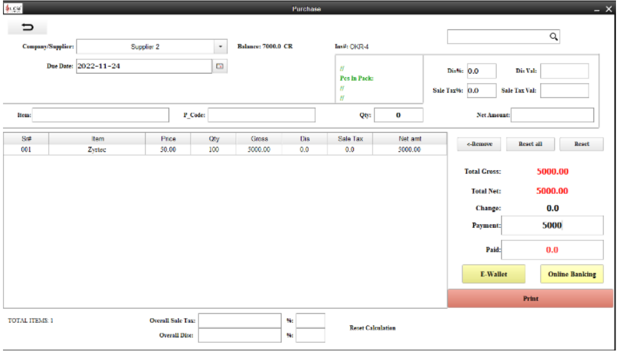
Panadol actifast
Qty: 100
Price: 45.0

Zyrtec
Qty: 580
Price: 70.0

Shipment: 0.0

Fig. 6: Sale page

For Fig. 7, the purchase page will allow the admin and employees to purchase products from the suppliers. They can choose what item to buy by typing on the item bar or input the product code. After that, they can select the quantity that they want, and the net amount will appear. Once the total net amount appears, they can pay the supplier via E-Wallet or Online Banking. After the payment is done, a notification will appear as seen in Fig. 8.

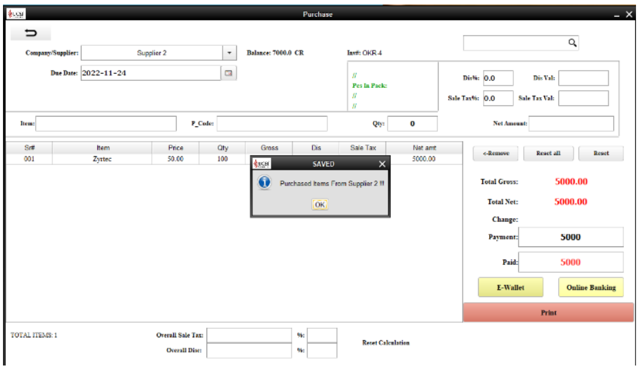

Purchase
<-Remove
Reset all
Reset
E-Wallet
Online Banking
Print
Reset Calculation

Company/Supplier: Supplier 2
Balance: 7000.0 CR
Inv#: OKR-4
Due Date: 2022-11-24
Pcs In Pack:
Dis%: 0.0
Dis Val:
Sale Tax%: 0.0
Sale Tax Val:
Item:
P_Code:
Qty: 0
Net Amount:

Sr#	Item	Price	Qty	Gross	Dis	Sale Tax	Net amt
001	Zyrtec	50.00	100	5000.00	0.0	0.0	5000.00

Total Gross: 5000.00
Total Net: 5000.00
Change: 0.0
Payment: 5000
Paid: 0.0
TOTAL ITEMS: 1
Overall Sale Tax: %:
Overall Disc: %:

Fig. 7: Purchase Page


Purchase
<-Remove
Reset all
Reset
E-Wallet
Online Banking
Print
Reset Calculation
SAVED
OK

Company/Supplier: Supplier 2
Balance: 7000.0 CR
Inv#: OKR-4
Due Date: 2022-11-24
Pcs In Pack:
Dis%: 0.0
Dis Val:
Sale Tax%: 0.0
Sale Tax Val:
Item:
P_Code:
Qty: 0
Net Amount:

Sr#	Item	Price	Qty	Gross	Dis	Sale Tax	Net amt
001	Zyrtec	50.00	100				5000.00

Purchased Items From Supplier 2 !!!

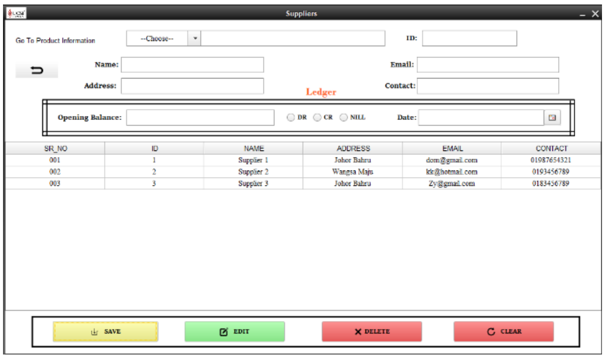
Total Gross: 5000.00
Total Net: 5000.00
Change:
Payment: 5000
Paid: 5000
TOTAL ITEMS: 1
Overall Sale Tax: %:
Overall Disc: %:

Fig. 8: Successful Purchase

The system also generates the purchase receipt of the transaction where both the employees and admin can print the receipt.

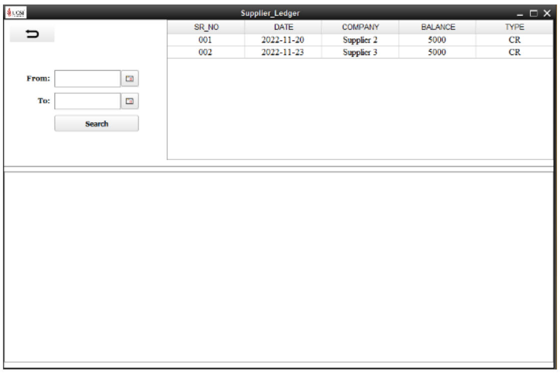
In this system, admin and employees can add new suppliers or company into the database. Firstly, they must input the name, address, email and contact and then save the info. A notification stating inserted successfully will appear after saving. Fig. 9 shows after the new supplier is added into the database and the info will be shown as well. The admin and employees can delete the suppliers from the database. A notification will inform them whether they really want to delete the supplier before being deleted. After confirming the deletion of the supplier, a notification stating deleted successfully will appear and the supplier is gone from the database.

Fig. 10 shows the supplier ledger where the admins and employees are able to check the existing suppliers' balance. Fig. 11 shows the product information page where the employees can add new products into the database. The employee must input the item name, product code, category, subcategory, brand, supplier, opening stock, purchase price, retail price, pieces in pack, quantity, location, purchase date and expiry date before saving into the database. The employee can also opt to clear the information to reinput the product details again.


Suppliers
Go To Product Information
--Choose--
ID:
Name:
Address:
Email:
Contact:
Ledger
Opening Balance:
○ DR ○ CR ○ NILL
Date:
SAVE
EDIT
DELETE
CLEAR

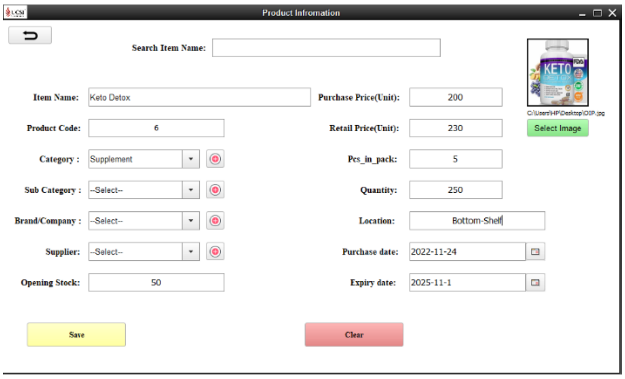
SR_NO	ID	NAME	ADDRESS	EMAIL	CONTACT
001	1	Supplier 1	Johor Bahru	dom@gmail.com	01987654321
002	2	Supplier 2	Wangsa Maju	kk@hotmail.com	0193456789
003	3	Supplier 3	Johor Bahru	Zy@gmail.com	0183456789

Fig. 9: Supplier Added to The Database


Supplier_Ledger
From:
To:
Search

SR_NO	DATE	COMPANY	BALANCE	TYPE
001	2022-11-20	Supplier 2	5000	CR
002	2022-11-23	Supplier 3	5000	CR

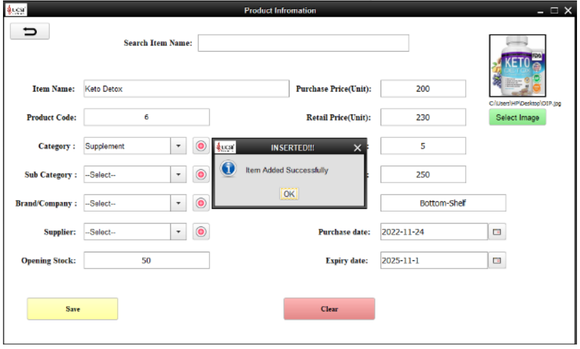
Fig. 10: Supplier Ledger


Product Infromation
Search Item Name:
C:\Users\HP\Desktop\OIP.jpg
Select Image
Save
Clear

Item Name: Keto Detox
Product Code: 6
Category : Supplement
Sub Category : --Select--
Brand/Company : --Select--
Supplier: --Select--
Opening Stock: 50
Purchase Price(Unit): 200
Retail Price(Unit): 230
Pcs_in_pack: 5
Quantity: 250
Location: Bottom-Shelf
Purchase date: 2022-11-24
Expiry date: 2025-11-1

Fig. 11: Add Product Page

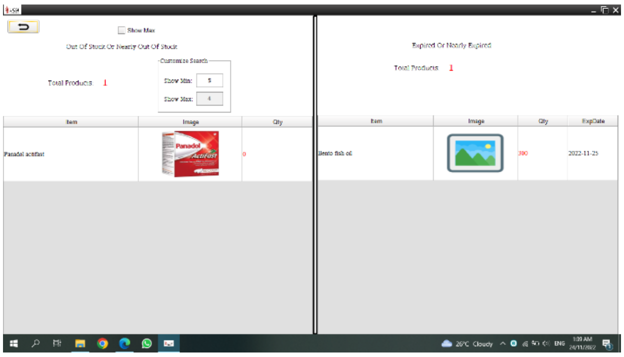
Fig. 12 shows the notification after the product information is inputted and saved successfully. Fig. 13 which is the Near Expiry or Out of Stock Page shows the products that are almost expiring or out of stock.


Product Infromation
Search Item Name:
C:\Users\HP\Desktop\OIP.jpg
Select Image
INSERTED!!!
OK
Save
Clear

Item Added Successfully

Item Name: Keto Detox
Product Code: 6
Category : Supplement
Sub Category : --Select--
Brand/Company : --Select--
Supplier: --Select--
Opening Stock: 50
Purchase Price(Unit): 200
Retail Price(Unit): 230
5
250
Bottom-Shelf
Purchase date: 2022-11-24
Expiry date: 2025-11-1

Fig. 12: Product Information Added Successfully


☐ Show Max
Out Of Stock Or Nearly Out Of Stock
Customize Search
Expired Or Nearly Expired

Total Products 1
Show Min: 5
Show Max: 4

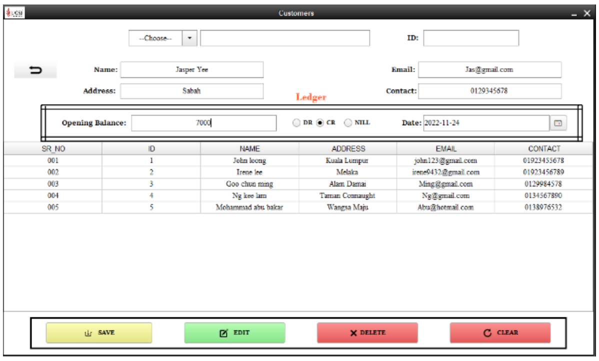
Item	Image	Qty
Panadol actifast		0

Total Products 1

Item	Image	Qty	ExpDate
Bento fish oil		300	2022-11-25

Fig. 13: Near Expiry or Out of Stock Page

The customer page allows both the admin and employee to add customers into the database system. They must input the name, address, email, contact, opening balance and date of joining before saving the details into the database system as shown in Fig. 14.

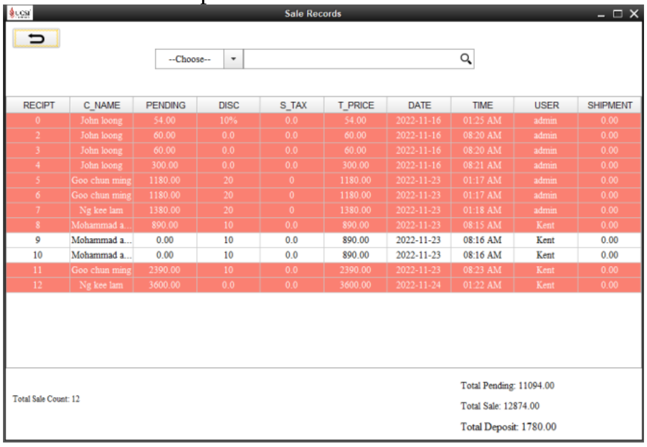


ID:
Name: Jasper Yee
Email: Jas@gmail.com
Address: Sabah
Contact: 0129345678
Opening Balance: 7000
○ DR ● CR ○ NILL
Date: 2022-11-24

SR_NO	ID	NAME	ADDRESS	EMAIL	CONTACT
001	1	John loong	Kuala Lumpur	john123@gmail.com	01923455678
002	2	Irene lee	Melaka	irene9432@gmail.com	01923456789
003	3	Goo chun ming	Alam Damai	Ming@gmail.com	0129984578
004	4	Ng kee lam	Taman Connaught	Ng@gmail.com	0134567890
005	5	Mohammad abu bakar	Wangsa Maju	Abu@hotmail.com	0138976532

Fig. 14: Customer Page

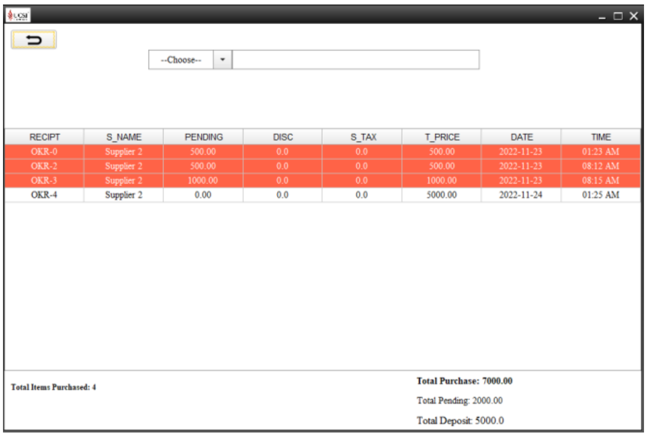
Figs. 15 and 16 shows the sales and purchase record pages in which the admin can check the transactions of both the sales and purchase.



RECIPT	C_NAME	PENDING	DISC	S_TAX	T_PRICE	DATE	TIME	USER	SHIPMENT
0	John loong	54.00	10%	0.0	54.00	2022-11-16	01:25 AM	admin	0.00
2	John loong	60.00	0.0	0.0	60.00	2022-11-16	08:20 AM	admin	0.00
3	John loong	60.00	0.0	0.0	60.00	2022-11-16	08:20 AM	admin	0.00
4	John loong	300.00	0.0	0.0	300.00	2022-11-16	08:21 AM	admin	0.00
5	Goo chun ming	1180.00	20	0	1180.00	2022-11-23	01:17 AM	admin	0.00
6	Goo chun ming	1180.00	20	0	1180.00	2022-11-23	01:17 AM	admin	0.00
7	Ng kee lam	1380.00	20	0	1380.00	2022-11-23	01:18 AM	admin	0.00
8	Mohammad a...	890.00	10	0.0	890.00	2022-11-23	08:15 AM	Kent	0.00
9	Mohammad a...	0.00	10	0.0	890.00	2022-11-23	08:16 AM	Kent	0.00
10	Mohammad a...	0.00	10	0.0	890.00	2022-11-23	08:16 AM	Kent	0.00
11	Goo chun ming	2390.00	10	0.0	2390.00	2022-11-23	08:23 AM	Kent	0.00
12	Ng kee lam	3600.00	0.0	0.0	3600.00	2022-11-24	01:22 AM	Kent	0.00

Total Sale Count: 12
Total Pending: 11094.00
Total Sale: 12874.00
Total Deposit: 1780.00

Fig. 15: Sales Record



RECIPT	S_NAME	PENDING	DISC	S_TAX	T_PRICE	DATE	TIME
OKR-0	Supplier 2	500.00	0.0	0.0	500.00	2022-11-23	01:23 AM
OKR-2	Supplier 2	500.00	0.0	0.0	500.00	2022-11-23	08:12 AM
OKR-3	Supplier 2	1000.00	0.0	0.0	1000.00	2022-11-23	08:15 AM
OKR-4	Supplier 2	0.00	0.0	0.0	5000.00	2022-11-24	01:25 AM

Total Items Purchased: 4
Total Purchase: 7000.00
Total Pending: 2000.00
Total Deposit: 5000.0

Fig. 16: Purchase Record

4. Conclusion

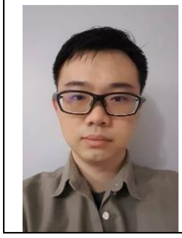
The development of technology allows warehouse management system to rapidly evolve with each day [6]. The technology allows warehouse management system to record data more efficiently compare to being done manually with people. The implementation of Pharmacy Warehouse Management System allows pharmacies to have their work handled with ease. After researching and going through multiple warehouse management system and to study their strengths and weaknesses, the proposed Pharmacy Management System was designed and developed. The Pharmacy Management System allows the pharmacies to check the stocks the stocks of products, generate sales transactions for the customers, purchasing products directly from the supplier or company. Besides that, the system also allows pharmacies to check the customers and suppliers that are registered in the system. Lastly, the system will also allow the users of the system to check the total amount of transaction per day.

References

1. Žunić, E., Delalić, S., Hodžić, K., Beširević, A. and Hindija, H., 2018, November. Smart warehouse management system concept with implementation. In 2018 14th Symposium on Neural Networks and Applications (NEUREL) (pp. 1-5). IEEE.
2. Ten Hompel, M. and Schmidt, T., 2008. Warehouse management. Springer Berlin Heidelberg.
3. Andiyappillai, N. (2020). Factors Influencing the Successful Implementation of the Warehouse Management System (WMS). International Journal of Computer Applications, 177(32), 21-25.
4. Lee, C.K., Lv, Y., Ng, K.K.H., Ho, W. and Choy, K.L., 2018. Design And Application of Internet of Things-Based Warehouse Management System for Smart Logistics. International Journal of Production Research, 56(8), pp.2753-2768.
5. Chen, C., Mao, J. and Gan, X., 2018. Design of Automated Warehouse Management System. In MATEC Web of Conferences (Vol. 232, p. 03049). EDP Sciences.
6. Hsia, K. H., Wu, M. G., Lin, J. N., Zhong, H. J., & Zhuang, Z. Y. (2018). Development of Auto-Stacking Warehouse Truck. J. Robotics Netw. Artif. Life, 4(4), 334-337.

Authors Introduction

Mr.Gan Jhui Ken

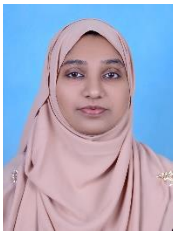


He is pursuing his B. Sc. (Hons) Business Information Systems, UCSI University, Malaysia, currently.

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. Shaik Shabana Anjum



She received a B.Eng. degree in Computer Science and engineering and M.Eng. degree (Hons.) in Software Engineering from Anna University, Chennai, India, in 2010 and 2012, respectively, and a Ph.D. degree in Computer Science from the University of Malaya, Malaysia, in 2018. She is an IEEE member. She has also served as a Research Assistant and Post-doctoral research fellow for projects involving traffic congestion and the Internet of Things (IoT) with the Centre for Mobile Cloud Computing Research, Faculty of Computer Science and Information Technology, University of Malaya. She is currently a Lecturer at UCSI University, Malaysia. Her research interests include the IoT, wireless sensor networks, radio frequency identification, ad hoc networks, cognitive radio, and energy harvesting. She has been accoladed with many awards at international competitions and appreciation at the faculty level for her research projects.

Development of an Automatic Allocation Parking System

Ng Wai Lam, Neesha Jothi, Shayla Islam
Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia

E-mail: neesha@ucsiuniversity.edu.my
www.ucsiuniversity.edu.my

Abstract

Today's people drive cars when they go out. As the number of cars increases, the demand for parking spaces in the parking lot also increases. Due to the lack of parking spaces in the parking lot, people spend more time looking for parking spaces after entering the parking lot, and also consume a lot of car oil.There is no any other parking system that has the ability to automatic allocation parking space for user. This is the reason for making this system. Learning from other similar systems, it is decided that this project is to develop a new parking system that will encompass some functions from the existing systems, but the system will automatic allocation parking spaces for user. To gather the required information, the project uses literature review to determine how this type of system functions and how it is made. Then to make a questionnaire cantered on the users' previous experience on existing systems. Lastly, to make a comparison of existing systems and to analyze the systems strengths and weaknesses to determine what functions to add to this new system. A total of 54 respondents answered the survey of the project. The development process will begin and be completed. Then extensive testing will begin to test all the functions to see if the system has any bugs and errors.

Keywords: automatic, parking lot, system, functions

1. Introduction

A parking lot is a designated area where cars are parked for a certain amount of time. Since cars are the main form of transportation in the world today, in fact, there are even different types of parking lots, ranging from public, private, multi-storey, and even automatic parking lots [1] being developed in urban areas. Some may also ask to pay for parking services. Technology and infrastructure developments have also led to different types of parking systems. This includes a parking guidance system, automatic parking system, vehicle built-in parking assistance system, etc. The purpose of this kind of parking is to reduce searching and parking time and petrol [2].

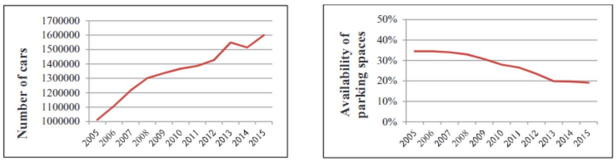
In the 21st century finding a free car parking slot has become a mind-numbing process, especially for people who travel in the morning to work, attend classes or are following their daily routine, they find it highly difficult and challenging to get a parking space to them [3]. Even paying parking lots have problems finding parking spaces. Nowadays, the demand for parking spaces is gradually increasing [4], resulting in a shortage of parking spaces in some parking lots. Many drivers have to spend time looking for an empty parking space after entering the parking lot [4], which also makes the parking lot crowded. In addition to wasting the driver's time looking for it, it also makes the driver feel irritable. One example is a shopping mall. The shopping mall which is the heart and soul of the city is one of the focal points of attraction for the public. The parking bay of the shopping mall is congested with vehicles causing a big traffic jam as a usual scenario during the weekends. The shopping mall personnel find it difficult to manage the deadlock situation. Due to the inefficiency of the parking system, vehicles are parked even at the roadside nearby the shopping mall. This hectic problem is addressed in a few of the national newspaper yearly [5]. But the situation remains the same even after applying different strategies to alleviate the problem. The root cause of the problem analyzed is the poor parking system employed in the shopping mall. When the problem is closely inspected, the reason for the traffic congestion is not because of lack of parking space but due to the lack of efficient scheduling of vehicles to the parking space. To alleviate the problem, the parking system should implement proper scheduling measures to schedule the vehicle effectively to the appropriate parking region without compromising the precious waiting time of the customer. This paper proposes a prototype Web-based automatic parking allocation system to solve this problem and save users more time when looking for parking spaces. Let the user feel that the parking system is not only convenient but also easy to operate. It also allows parking lot owners to use my system to reduce unnecessary costs such as maintenance costs, machine costs, machine operator costs, etc. For example, when the driver arrives at the entrance of the parking lot, he only needs to scan the QR code at the entrance and fill in the license plate number and some other information. Then the system will automatically arrange a parking space for the driver. The driver only needs to park according to the parking space provided by the system and does not need to drive around in the parking lot in order to find an empty parking space.

2. Literature Review

In public or private structures, an automobile parking system is a solution that consists of barrier gates, an access control system, and an automated parking system. With the use of the vehicle detecting sensors system, automobile owners may quickly find and reserve a parking place at any lot they feel is practical. The introduction of a simple payment method also improves the convenience of vehicle entry and exit. To allow a vehicle to enter and exit, the system uses access control, tickets, and tokens. The barrier gate will lift after the visitor car has a ticket or token from the entrance ticket or token dispenser, allowing it to enter the parking area. When a visitor wants to leave the premises, they can pay the parking fees at the parking payment machine or directly to a staff member in the business's reception or till area. The barrier gate will raise when the bought ticket or token is inserted into the ticket or token collecting device, allowing the tourist vehicle to leave the structure. Most of the parking lot systems are applied to shopping plazas, commercial buildings, high-rise residential buildings, etc [6].

Car parking is one of the primary challenges for transportation and traffic management around the world due to the rising rate of private car usage in metropolitan areas as a result of the quickly expanding economy, negligent regulations, and subsidies. Parking is becoming a barrier to through-traffic operation, as shown by the synchronisation between parking policies and traffic management.

Even if judgments are made on an ad hoc basis while developing policy, it is still accountable for the inefficient use of resources. Therefore, it is important to comprehend parking choice behaviour and actual parking space demand. Numerous studies have been conducted over the past three decades to assess parking characteristics, gauge parking demand, and examine driver behaviour when selecting a parking place.



a) b)

Fig.1. Dynamics of a) the number of private vehicles and b) the availability of parking spaces in St. Petersburg

Due to the rapid global rise of automobile auto mobilization, the issue of organizing permanent storage of cars is currently important. In St. Petersburg, the number of private vehicles has nearly tripled during the last ten years [7]. In the meantime, the number of parking lots and garages gradually declines as they are viewed as transient structures and are slated for demolition, freeing up significant areas for future construction of other structures such as roads, factories, corporate centers, etc. In addition, there are infill developments, primarily residential structures, and public institutions, all around St. Petersburg, and the majority of them lack parking spaces or have fewer spaces than the actual number of automobile owners.

This article [4] incorporates all of these factors and offers a cutting-edge analysis of models and studies on the parking system. Parking-related issues, parking features and their applications, driver choice behaviour, demand model development taking into account various elements, and policy evaluation as an essential component of the urban transportation system are all covered in detail. Despite being underdeveloped, the authors discovered that the research suggests that during all phases of planning and policy formulation, more focus should be placed on metrics such as ease of access, walk time, parking costs, parking guidance and information systems, management, etc. When seen collectively, the studies presented present insightful data about the overall parking infrastructure. It also gives planners and policymakers vital data for building, planning, and assessing parking systems.

According to [8][9][10] will talk about the challenges of locating parking spaces in privately owned parking lots (such as those at malls, offices, or apartment complexes) and the cutting-edge solutions used in the market. Most persons who have ever operated or ridden in a vehicle have unavoidably run into these problems.

First, the unnecessary time spent merely searching for parking spaces, particularly during busy times like weekends or holidays [9]. Without finding that one available parking space, one could waste a lot of time burning expensive fuel. Occupancy sensors detect the number of parking spaces that are currently available in most mall parking lots and display it at the entrance to each floor of the parking lot, making it easier for people to find a vacant place. Is this, however, truly enough? Even though we are aware that there is parking space available, we do not yet know where it is or how to get there. These issues can be summed up as a consumer's lack of knowledge regarding parking availability, detection, and wayfinding.

Second, there are not many ways to pay for parking [8]. Today, the majority of parking lots still only accept cash payments made manually at payment counters or automatically at machines. In this area, where more convenient payment methods are being deployed, progress is being made. You can avoid searching your wallets for spare change by using cashless purchases like TouchNGo cards and Visa Debit cards, which are instances of this. However, if you still need to physically reload your payment cards, there would be a problem with running out of funds.

Third, parking systems that rely on tickets are unreliable and ineffective [8]. When cars line up outside parking lot entrances while drivers manually take admission tickets or scan their credit cards, traffic congestion can result. This all goes against the concept of a "smart city," because time is lost by adding together these minor inefficiencies.

3. Results

Automated parking systems help to speed up the process of paying for parking tickets. This is very prominent among many shopping malls and universities across the world. People in the survey heavily rely on the system to be efficient and fast to make the process of parking much quicker than before the automated parking system. In the system that the project is developing, it has some features that will attract many users including automated allocation of parking spaces, a rewards system for user’s continuous use of the system and an easy-to-use payment system.

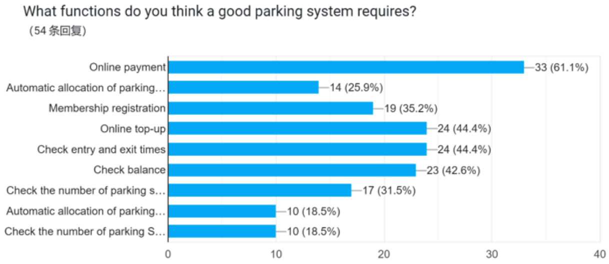


Fig.2. Bar chart of the possible functions to include

Table 1. Functions Respondent

Response	Frequency	Percentage
Online payment	33	61.1%
Automatic allocation of parking spaces	24	44.4%
Membership registration	19	35.2%
Online top-up	24	44.4%
Check entry and exit times	24	44.4%
Check balance	23	42.6%
Check the number of parking spaces available	27	50%

Fig.2 and Table 1 shows what functions all 54 respondents think a good parking system requires. Most of the respondents think that online payment is needed, there are 33 respondents choose online payment, which covers 61.1% of the total respondents. There 24 over 54 respondents think the automatic allocation of parking spaces, online top-up, and check entry and exit times is required in a good parking system. There are only 19 respondents who think membership registration is a good function. 23 respondents choose to check their balance, and 27 respondents choose to check the number of parking spaces available, which covers 50% of all respondents.

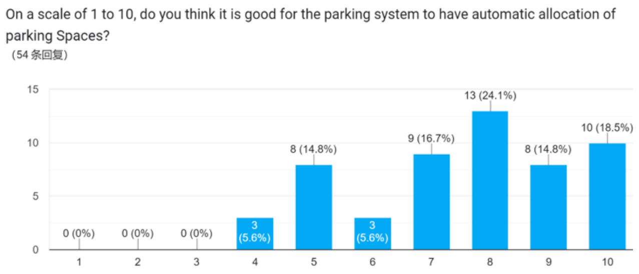


Fig.3. Likert Scale of Respondent

Table 2. Likert Scale Respondent in Percentage (%)

Response	**Frequency**	**Percentage**
1	0	0%
2	0	0%
3	0	0%
4	3	5.6%
5	8	14.8%
6	3	5.6%
7	9	16.7%
8	13	24.1%
9	8	14.8%
10	10	18.5%

Fig.3 and Table 2 shows that all 54 respondents agree or disagree that a good parking system requires automatic allocation of parking spaces. The response is from 1 (Strongly disagree) to 10 (Strongly agree). There is no respondent choose 1 – 3 points on the scale. 3 respondents have chosen 4 points and 6 points on the scale. 5 points and 9 points on the scale are chosen by 8 respondents, which covers 14.8% of all respondents. Total of 9 respondents choose 7 points on the scale. Most of the respondents choose 8 points on the scale, which covers 13 respondents, a total of 24.1% of the total respondents. 10 over 54 respondents which cover 18.5% of all respondents strongly agree that a good parking system required automatic allocation of parking spaces.

4. Conclusion

Today's people drive cars when they go out. As the number of cars increases, the demand for parking spaces in the parking lot also increases. Due to the lack of parking spaces in the parking lot, people spend more time looking for parking spaces after entering the parking lot, and also consume a lot of car oil. My system can help drivers solve the time-consuming problem of finding a parking space, so that the driver knows that the parking lot is full, and does not need to enter the parking lot and circle in order to find a parking space. Furthermore, my system also helps owners reduce unnecessary expenses, such as system maintenance and other expenses. Because my system only needs to arrange one or two staff members and able to manage the entire parking lot.

References

1. “The Different Types Of Parking Facilities,” *Parking Ottawa*, Sep. 09, 2015.
2. Y. Geng and C. G. Cassandras, “A new ‘Smart Parking’ System Infrastructure and Implementation,” *Procedia - Social and Behavioral Sciences*, vol. 54, pp. 1278–1287, Oct. 2012.
3. A. Kianpisheh, N. Mustaffa, P. Limtrairut, and P. Keikhosrokiani, “Smart Parking System (SPS) Architecture Using Ultrasonic Detector,” *https://www.researchgate.net*, vol. 6, pp. 51–58, Jul. 2012.
4. J. Parmar, P. Das, and S. M. Dave, “Study on demand and characteristics of parking system in urban areas: A review,” *Journal of Traffic and Transportation Engineering (English Edition)*, vol. 7, no. 1, pp. 111–124, Feb. 2020.
5. D. Thomas and B. C. Kovoor, “A Genetic Algorithm Approach to Autonomous Smart Vehicle Parking system,” *Procedia Computer Science*, vol. 125, pp. 68–76, 2018.
6. “Car Parking Management System in Malaysia | MSICT,” *msict.com.my*.
7. I. Duvanova, T. Simankina, A. Shevchenko, T. Musorina, and A. Yufereva, “Optimize the Use of a Parking Space in a Residential Area,” *Procedia Engineering*, vol. 165, pp. 1784–1793, Jan. 2016.
8. admin, “Smart Technology of Private Parking Lots in Malaysia,” *Asia Mobiliti*, Jul. 19, 2021.
9. I. Duvanova, T. Bubnova, and M. Romanovich, “Efficiency of Use Underground (Dike) Multilevel Parking in Conditions of Cramped Housing Development,” *Procedia Engineering*, vol. 165, pp. 1794–1800, 2016.
10. Sato, N., Yokotani, T. and Morita, Y. (2020) “Intention classification of a user of a walking assist cart by using support vector machine,” Journal of Robotics, Networking and Artificial Life, 7(1), p. 27. Available at: https://doi.org/10.2991/jrnal.k.200

Authors Introduction

Mr. Ng Wai Lam

He received his Bachelors from the Institute of Computer Science and Digital Innovation (ICSDI), UCSI University, Malaysia.

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. Shayla Islam

Currently, Dr. Shayla Islam is an Associate Professor in UCSI University, Malaysia. She has completed PhD degree in Engineering from Electrical and Computer Engineering (ECE) departmentat International Islamic University Malaysia (IIUM) in 2016 under Malaysian International Scholarship (MIS). She completed her MSc.at the department of Electrical and Computer Engineering in InternationalIslamic University Malaysia in 2012. She completed her BSc. in ComputerScience and Engineering from International Islamic UniversityChittagong, Bangladesh. Dr. Shayla has awarded a Silver medal for her research work at International Islamic University Malaysia. In consequences, she has also awarded a Young Scientist Award for the contribution of research paper at 2nd International Conference on Green Computing and Engineering Technologies, 2016 (ICGCET'16), Organized by the Department of Energy Technology, Aalborg University, Esbjerg, Denmark. Her current research interests include data communications and networking, computer networks and wirelesscommunication, network mobility in heterogeneous network, computer architecture & cloudcomputing. She published more than 70 papers in international journals and conferences(WoS/ISI/ESCI and Scopus indexed). She works as a reviewer for many international conferencesas well as journals.

Healthcare Mobile Application

Yee Chee Hong, Neesha Jothi, Javid Iqbal
Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia

E-mail: neesha@ucsiuniversity.edu.my
www.ucsiuniversity.edu.my

Abstract

The reason behind this project is being carried out because I wanted to help those people who live in a busy city to take care their health. They suffered a lot from working and other stress that caused themselves always to keep in unhealthy status. Although they know they are lives such unhealthy, still they have no time to manage it. Small health problem will always ignore by them and in the end fell ill seriously. Also, the unfriendly user interface of current healthcare mobile application that caused inconvenience to use them. This project will improve those weaknesses that allocated from current existing system. The healthcare mobile will open use to public users. It will come with online appointment, online pharmacy purchase, chat support and check information or history. New function added will be e wallet which can make payment or top up using serial number or smart wallet such as TNG pay and Shopee Pay.

Keywords: healthcare, mobile app, mobile application

1. Introduction

Health is always important to people who lives in this world, no matter in which generation. A good health will lead to a productive life [1]. In this generation, many people do not able to manage their health issues because of their busy working life in such bustling society. Overworking and exhausted will cause the increment of stress and decrement of self-esteem. It will cause more mental health problem such as anxiety, depression, and substance use disorders [2]. Therefore, with the high developed technology, healthcare system application comes out in front of people. They are developed to help these people to take care their healthy anytime and anywhere. In 2020, COVID-19 breaks out into the whole world. In the period of Covid-19, clinic and hospital had entered a state of emergency. Clinic and hospital all around the world become very busy on helping people to cure from illness. Therefore, healthcare application plays an important role here. Healthcare application can help those clinic and hospital to manage their system by provide online appointment for treatment. Healthcare application is rapid developed to help this critical situation. Although the critical situation has passed, people are still using the healthcare application because they realize the convenience of using healthcare application. Nowadays, mobile health app market has improved steadily [3].

2. Methodology

Research methodology allow project to gather the requirement information from public that needed in this project with several techniques or methodologies, then measuring and analyzing the requirement information collected. Qualitative research methodology is non-quantitative type of analysis which is aimed at finding out the quality of phenomenon. Quantitative research methodology is used to measuring the quantity or number of phenomena by use of statistical analysis. Qualitative gives deeper understanding on research that are not clear which quantitative gives clear generalized fact on research. In this paper, the authors will choose quantitative research methodology and it will be survey and questionnaire. It is easy to implement while it can gather public's opinion on this healthcare mobile application and make improvements. Google Form will be the platform used to make the survey and questionnaire. Question may be easy for the purpose of easier to collect simple information. Target respondent will no limit to any population because the application is made for public user. To gather the information requirement, the separated sections will focus on different information. Section A will collect information about the understanding of users towards healthcare mobile application. The questionnaire had been done by 50 persons. Each sector shows out all result in pie chart in Google Analysis as Fig. 1 and Fig. 2.

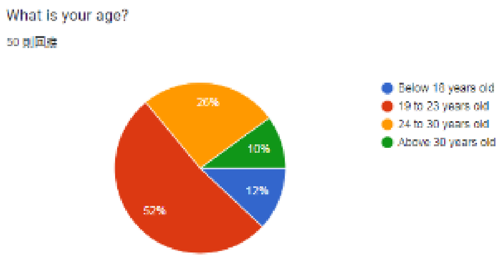


Fig.1.Sector A

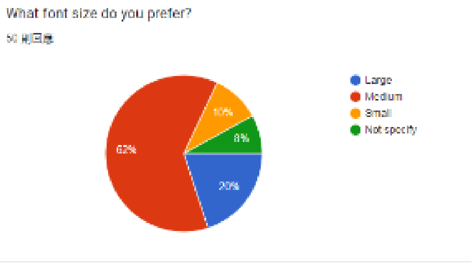


Fig.2.Sector C

In Sector A, results are all the information from participants such as age, gender, and habit on using application. In Sector B, results show their opinion on user interface of application they prefer, example low English understanding level, medium font size, bright theme application. In Sector C, results show requirement of application from participant. This affect to the development to ensure that make those function required works in the application. The functions such as allow to multi people make appointment, online purchase from pharmacy, preference of searching engine, usage of in-build e wallet and so on.

3. Results

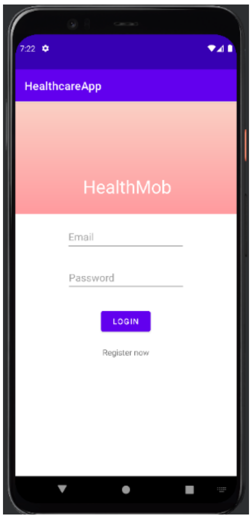


Fig.3.Login Page

Fig.3 show to login page of HealthMob application, user need to enter existing account with email and password, either one incorrect will not login. User can access to register page by click "Register now" text.

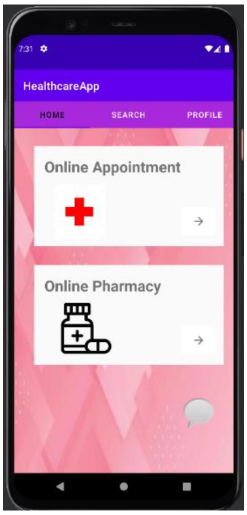


Fig.4.Home Menu

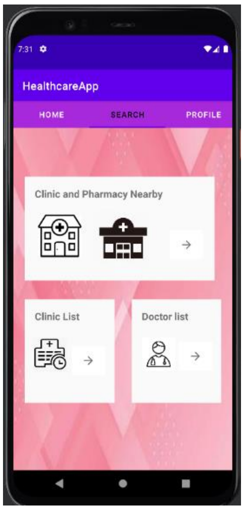

7:31
HealthcareApp
HOME
SEARCH
PROFILE
Clinic and Pharmacy Nearby
Clinic List
Doctor list

Fig.5.Home Search

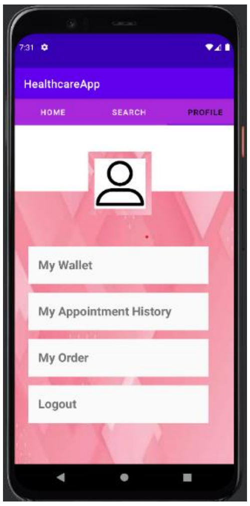

7:31
HealthcareApp
HOME
SEARCH
PROFILE
My Wallet
My Appointment History
My Order
Logout

Fig.6.Home profile

Fig.4 is menu home page of HealthMob application, user can perform action to make online appointment or make online pharmacy purchase item. Fig.5 is menu search page, user can perform action to search and view nearby clinic and pharmacy location, clinic and doctor list of HealthMob. Fig.6 is menu profile page, user can perform action to enter wallet page, view appointment and purchase item history, and logout from system.

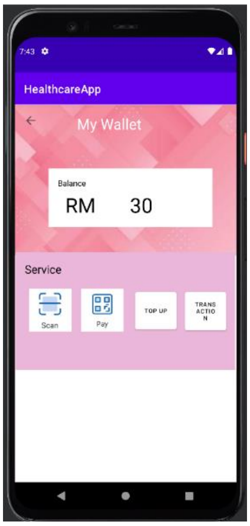

7:43
HealthcareApp
My Wallet
Balance
RM 30
Service
Scan
Pay
TOP UP
TRANSACTION

Fig.7.Wallet page

Fig.7 is wallet page can access from profile menu page. User can see their wallet and act such as scan or show QR, top up wallet and view transaction.

4. Conclusion

This healthcare mobile application can allow public user can use this application to manage their healthy. It allows user to choose their nearby clinic to make online appointment. At the same time, they can also purchase some medicine to keep themselves in health. With the in-build e-wallet system, user can easily make payment with this application as it can reload from others smart e-wallet too.

References

1. Dai, F. *et al.* (2020) “Design of an instrument for measuring heart rate and blood oxygen based on Arduino,” *Journal of Robotics, Networking and Artificial Life*, 7(4), p. 275. Available at: https://doi.org/10.2991/jrnal.k.201215.014.
2. V. B. S. H. Doctors and Verified By Star Health Doctors https://www.starhealth.in/blog/author/star-health-doctor, “Health and its importance,” Star Health, 22-Apr-2022. [Online]. Available: https://www.starhealth.in/blog/health-and-its-importance#:~:text=
3. “Understanding the agile software development lifecycle and process workflow,” *Smartsheet*, 18-Aug-2016. [Online]. Available: https://www.smartsheet.com/understanding-agile-software-development-lifecycle-and-process-workflow. [Accessed: 10-Jun-2022].

Authors Introduction

Mr. Yee Chee Hong

He received his bachelor's from the Institute of Computer Science and Digital Innovation (ICSDI), UCSI University, Malaysia.

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. Javid Iqbal

Javid Iqbal is currently working as a Lecturer at UCSI University, Malaysia. He holds a PhD in Information and Communication Technology from the National Energy University Malaysia. He holds a bachelor's degree in Computer Science and Engineering from Dr. M.G.R. Educational and Research Institute and Master's degree in Communication Systems from Dr. M.G.R. Educational and Research Institute. His research interest includes Multimedia, Augmented Reality and Virtual Reality.

Solar Powered Seed Sprayer Machine

Ammar A.M. Al- Talib[1], Yap Chee Xian[2], Ain Atiqa[3], Nor Fazilah Abdullah[4]
Department of Mechanical and Mechatronics, Faculty of Engineering,
*Technology and Built Environment, UCSI University, 56000 Kuala Lumpur, Malaysia**

E-mail: ammart@ucsiuniversity.edu.my, 1001954697@student.ucsiuniversity.edu.my, ainatiqa@ucsiuniversity.edu.my,
norfa@ucsiuniversity.edu.my

Abstract

The main objective of this project is to fabricate a complete functional seed sprayer machine which is fully powered by solar energy. The solar seed sprayer machine should be able to spray different types of vegetable seeds. Further analysis about the performance has been conducted through the seed amount sprayed over time and area covered by the machine. Solar energy is used as the power supply for the machine. Wireless communication is used to remotely control the machine, and the 3D printing technology is used to assist in the fabrication of required components. The solar seed sprayer machine under research is composed of four main systems. The remote driving system, solar charging system, seed storage dispenser system, and impeller spreader system. Different experiments have been conducted to assess the performance of the machine. The performance of machine is indicated through the capability of machine on spreading different types of seeds with various size and shape. The spread seed count has been also tested as well with the area covered by the machine.

Keywords: Agriculture, Solar Powered, Seed Spreader Machine, Remote Controlled

1. Introduction

Agriculture is the practice of cultivating plants and livestock [1]. Growing crops is one of the most important activities through centuries. To increase efficiency of growing crops, tools were developed to save time and efforts in agriculture [2]. From the combination of tools and animals to modern machinery to fulfil the growing food demand by huge population of human society. In the twentieth century, industrial agriculture was developed to dominate agriculture output to society [3]. The technology of industrial agriculture including the agriculture machinery, genetic technology, and synthetic fertilizers. Till today, developing on agriculture machinery continue to fulfil the future demand of the world.

In this era, most common agriculture machinery are tractors powered by fuel. However, fuel powered machinery is not sustainable as fuel is one of the finite resources which is causing pollution to the environment. During the combustion of fuel, air pollutants such as carbon dioxide and nitrogen dioxide are emitted to the air and being harmful to the environment and life beings on earth. To solve this issue, this project aims to develop a seed sprayer machine that is purely powered by solar energy.

2. Methodology and Experimental Setup

This project focus on the broadcasting as it possesses the highest efficiency of seed sowing for those crop that does not require singular space [4]. Crops that are suitable with broadcasting method including wheat, corn, lettuce, carrot, and more. Further research was done on the principle of broadcasting, working system of ordinary broadcast spreader, and every part of broadcast spreader. Broadcast spreader is an equipment that spread the seed over the field using centrifugal force [5]. Seeds are stored in a hopper positioned above a spinning disk. The spinning disk rotates at high rotation speed. When seed drops from the hopper and fall on the spinning disk, the centrifugal force generated by the spinning disk will throw the seed out for a distance [6].

The seed sprayer machine is purely powered by solar energy. Photovoltaic is the conversion of light into electricity using semi conducting materials [7]. With an additional solar charge controller, photovoltaic system could be used to generate electricity and stored in rechargeable battery. However, the voltage and current generated by the solar panel are floating depending on the sunlight. Thus, solar charge controller is used to avoid battery damage due to excessive charging or discharging [8].

To achieve remote control feature of solar seed sprayer machine, research is done on the wireless communication. Some common wireless communication method includes radio frequency, infrared, WiFi, and Bluetooth. WiFi and radio frequency are implemented in the project to control the solar seed sprayer machine remotely [9].

Fig. 1 is showing the circuit used for the solar energy charging control system.

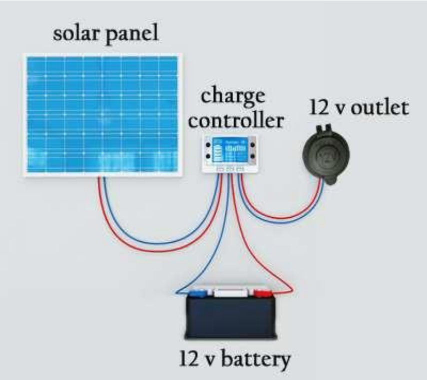


Fig. 1 The circuit used for the solar energy charging control system

The fabrication work of the solar seed sprayer machine has started with 3D modelling design using CAD software. The popular CAD software, Auto-desk Inventor, has been used to complete all the 3D modelling for all parts needed in the proposed solar seed sprayer machine. The fabrication process has strictly followed the model designed in CAD software. Calculations are done to analyze the expected performance of the solar seed sprayer machine.

The solar charging system has been tested before installation. The solar power system is operating at 12V constant voltage to power the whole system of solar seed sprayer machine. The solar panel has a maximum power of 20W with 18V working voltage. Calculations are showing that a 12V-7Ah lead acid battery discharged to 50%, will take about 4 hours to be fully recharged.

After the solar charging system is installed, seed storage container has been built using 12mm plywood sprayed with waterproof layer, the dispenser mechanism is designed based on a cereal dispenser design [10]. By rotating a spinning paddle, the material in the storage container will be dispensed. The dispenser mechanism was 3D modeled and printed. Lastly, the impeller spreader system has also been modeled and printed. The electric motor used in dispenser system has a rotation speed of 40 rpm while the impeller spreader system has a rotation speed of 280 rpm. By attaching a spur gear set, the rotation speed of dispenser mechanism has been increased to 60 rpm and the rotation speed of impeller spreader system is increased to 700 rpm.

Fig. 2 and 3 are showing the schematic diagram of the system and the fabricated prototype. Experiments has been conducted to test the performance of the proposed solar seed sprayer machine. Three different sets of experiments are conducted to test the functionality, the efficiency, and the area that could be covered by machine. Different types of seeds and beans with varies size and shape are prepared and dispensed by the machine.

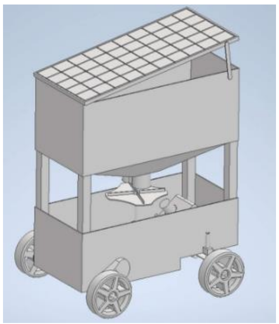


Fig. 2 The schematic diagram of the machine

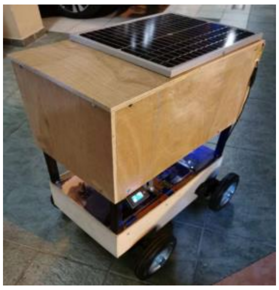


Fig. 3 The testing prototype

3. Results and discussion

To test the seed spreading, ten different types of seeds and beans are prepared with same weight and dispensed by the machine. All the test using constant weight, 100g. The performance of the machine has been observed and recorded in Table1 shown.

From Table 1, it can be observed that most of the seeds and beans are able to be dispensed but some are occasionally stuck. This can be attributed to the design flaw of dispenser mechanism; the seed will occasionally stick in between the edge of spinning paddle and the wall. As the torque of electric motor is 4.5kg.cm, the torque is not enough to overcome the resistance. However, with flow rate controlled, most of the seeds can be dispensed except the large ones such as the pumpkin seeds and the chickpeas. Further modification may be done to overcome these issues.

The next set of experiments are conducted to evaluate the seed dispense rate. The dispenser mechanism is switched on for 10 second and the seeds dispensed are collected. The collected seeds are weighted and recorded. Fig. 4 is showing the results.

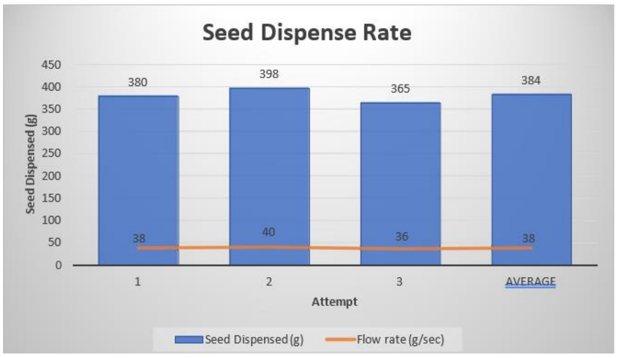


Fig. 4 Seed dispense rate experiment data

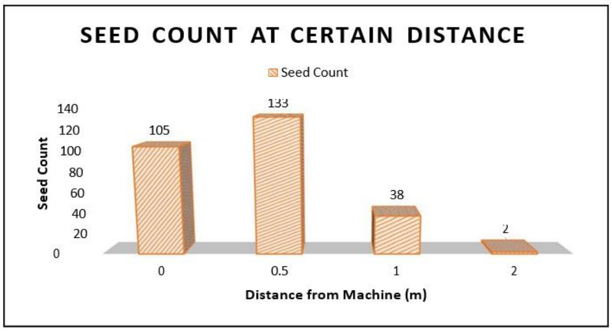


Fig. 5 Average Outcome of seed spreading area

Table 1 Capability of Seed Sprayer Machine on Different Types of Seeds

Types of seeds	**Size & Shape**	**Potential**
Sesame Seed		Dispense smoothly and seeds leaked when the machine was not operated.
Wheat Seed		Working but occasionally stuck
Rice		Dispense smoothly
Cracked Corn		Dispense smoothly
Pumpkin Seed		Stuck and may crush the seed
Green Bean		Dispense smoothly
Red Bean		Working but occasionally stuck
Soybean		Working but occasionally stuck
White Bean		Working but occasionally stuck
Chickpeas		Totally stuck in mechanism

The results have indicated that the average dispense rate of the machine is approximately 38g/sec at the maximum motor speed. The last set of experiments have been conducted to test the spreading area that can be covered by the machine. Different distances of 0.5, 1.0, 1.5, 2 meters from the machine are chosen. The machine has run for 10 seconds, and the seed count is recorded. The experiment has been repeated at three different attempts and the average outcome is considered.

The most effective distance is 0.5 metre from the machine and the maximum distance is 2.5 metre from the machine. Since the seed spraying is taking a circular pattern with the machine at the center, the effective area covered by the machine is approximately 3.14 m^2. Fig. 5 is showing the average outcome of seed spreading coverage area.

4. Conclusion

The proposed seed spreading machine has shown successful outcome and achieving the objectives of this project. The solar seed sprayer machine has been purely powered by solar energy. The machine is capable to dispense most of the seeds and beans with small and medium size, the maximum spreading rate achieved is 1353 seed count per second, and the area covered by the machine is 3.14 m^2 along its running way.

The machine can be considered as a successful implementation of the sustainable development goals (SDG’s) in applying renewable energy for agriculture and fighting against poverty.

Acknowledgment

The authors would like to express the gratitude towards the school of engineering at UCSI University for the support throughout the research. Special thanks would be extended to CERVIE office at UCSI University for the endless support to research.

References

1. Agriculture: definition and overview. (2014). Harris, D.R. & Fuller, D.Q.
2. Food and Agriculture Organization of the United Nations.
3. Lichtenberg, E. (2002). Agriculture and Environment. Handbook of Agriculture Economics,2,1249-1313.
4. Luna, T., Wilkinson, K. M., & Dumroese, R. K. (1949). Seed germination and sowing options. Nursery manual for native plants: A guide for tribal nurseries, 1, 133-151.
5. Balfour, J. R., & Shaw, M. (2011). Introduction to photovoltaic system design. Jones & Bartlett Publishers. Peter Grandics (2013).
6. New Concepts for the Reversal of Gravity and Time and Some Practical Derivations,
7. Solar Energy Perspective: Executive Summary. (13 January 2012). International Energy Agency, 2011.
8. Kannan, N., & Vakeesan, D. (2016). Solar energy for future world: -A review. Renewable and Sustainable Energy Reviews, 62, 1092-1105.
9. What is wireless communications? Everything you need to know, Chuck Moozakis & Sean Michael Kerner, (2022).
10. An automated cereal/grain dispenser, S.M. Karmusel, Afrin Momin, Jui Paranjape , Snehal Patil , Ankita Sawant, International Journal of Research in Engineering and Science (IJRES) ISSN (Online): 2320-9364, ISSN (Print): 2320-9356, Volume 9 Issue 5 ǁ 2021 ǁ PP. 38-43.

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.

Mr. Yap Chee Xian

Yap Chee Xian is currently further his Bachelor’s Degree in Mechanical Engineering at UCSI University, Kuala Lumpur. His field of research is renewable energy.

Ms. Ain Atiqa

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.

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.

Solar Powered Outdoor Air Purifier with Air Quality Monitoring

Ammar A. M. Al-Talib, Ik Chu Aung, Noor Idayu M.Tahir
Department of Mechanical and Mechatronics, Faculty of Engineering,
Technology and Built Environment, UCSI University, 56000 Kuala Lumpur, Malaysia

Sarah 'Atifah Saruchi
Faculty of Mechanical and Mechatronics, University Malaysia Pahang, 26600 Pekan,
Pahang, Malaysia
E-mail: ammart@ucsiuniversity.edu.my, 1001852392@ucsiuniversity.edu.my, NoorIdayu@ucsiuniversity.edu.my,
saratifah@yahoo.com

Abstract

This paper has documented the detailed design, fabrication and test of a solar powered air purifier prototypee with a High Efficiency Particulate Air Filter (HEPA) and Carbon Filters which can achieve air purification with self-sustainable ability. Besides. Several tests have been conducted to assess the performance of the proposed solar operated air purifier. In the first test, a 67.37% efficiency is achieved for the solar panel. Second test of air purifier test has shown the efficiency of cleaning ammonia pollutant in the air as 43.55% for burning cigarettes and 35.33% for floor detergent using the equipped two MQ135 sensors. The findings are showing that the floor detergent might have higher rate of diffusion than ammonia molecule found in cigarette smoke.

Keywords: Solar Power, Air purifier, HEPA, Air Quality

1. Introduction

Air pollution has plagued the world with millions of diseases and deaths to humanity. The effects of air pollution are of more influence at urbanized areas compared to rural areas. As technology advances, richer countries use active technology and efficient designs to improve their air quality over time. In contrast, lower income and developing countries have none or little emphasis in constructing strict guidelines and regulations. Progress of improving air quality is often downplayed and ignored as the economic growth of these countries is far more important than the environment [1].

All countries have their own standards of measuring air quality. Since this design project is based in Malaysia, the project is designed according to the Malaysian Standards. According to Malaysia ambient air quality standards, it incorporates six air pollutants criterion, including particulate matter with a size less than PM 10. SO2, CO, NO2, and O3 ground level ozone, as well as another extra parameter, particulate matter with a size less than PM 2.5 micron [2]. Fine particulate matters are contaminants found in the air that have been linked to both acute and chronic health problems. Even though most particulate matter research is focused on outdoor exposures, most of the individuals are spending most of their time indoors, where particulate matter from outdoors can penetrate to indoor with ease. Particulate matter is proven to affect respiratory, cardiovascular, and nervous systems, and linked to higher mortality risk [3]. For example, respiratory diseases like COVID-19 in recent years, the virus SARS-CoV-2 is largely spread via droplets and particulate matter in the air, thus transmitting it into the human respiratory system. It also prolongs the recovery phase from the disease’s symptoms [4].

An air purifier with HEPA Filter is being used widely to prevent harmful substances like particulate matter and chemical from entering a space in exchange of clean air. HEPA stands for high efficiency particulate air filter. It is a type of mechanical filter. Substances in the air like dust, pollen, mound, germs, and other airborne particles larger than 0.3 μ m can potentially be removed by a HEPA filter. The 0.3 μ m diameter standard expresses the worst-case situation, the particle size with the highest penetration. Larger or smaller particles are captured with considerably greater efficiency with a HEPA filter. There are tons of different air purifiers available in the market with HEPA filters. To differentiate the performance of different air purifiers, the Minimum Efficiency Reporting Value (MERV) was introduced by American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) for the rating of filter efficiency. For air purifier with filter that can filter out dust and particulate matter need at least MERV rating 13 [5].

2. Methodology and Experimental Setup

2.1 The prototype design details

The conceptual design was carried out using the SOLIDWORKS programme for planning, visual concept generation, modelling, feasibility evaluation, prototyping, and project management. The system under consideration is having an air purifier unit and the solar unit for the solar power supply. The air purifier unit consists of the core components which are the *HEPA* filter, *DC* fan, and car battery. While the solar unit having fold-able and flexible solar panel, solar charge controller and car battery. The working principle is based on the solar unit to be placed at an outdoor location during sunshine hours of the day. The car battery will be charged by the connection to the solar panel via charge controller. After full charging, the charged car battery will get disconnected from the solar unit and to be transferred to the air purifier unit and connected to the *DC* fan. The running *DC* fan will blow the air to the *HEPA* filter where it can filter up the air. The separation of solar charging and air filtration units has been decided to give better mobility and flexibility in using the system. With the portability of fold-able and flexible solar panel, it can be placed at locations of less shading and more solar radiation.

The final design has been modified by adding MQ-135 Sensor and Arduino Uno for testing the air quality of the solar powered air purifier. The MQ-135 Sensor could show reading of air quality while connected to the Arduino Uno board. It can send the data back to the computer to show the performance of the air purifier unit. Two MQ-135 sensors are placed in front of the air duct outlet and back of the air duct inlet to ensure the sensor could detect accurate data. Figure 1 showing the fabricated testing Solar-Air Filter Purifier

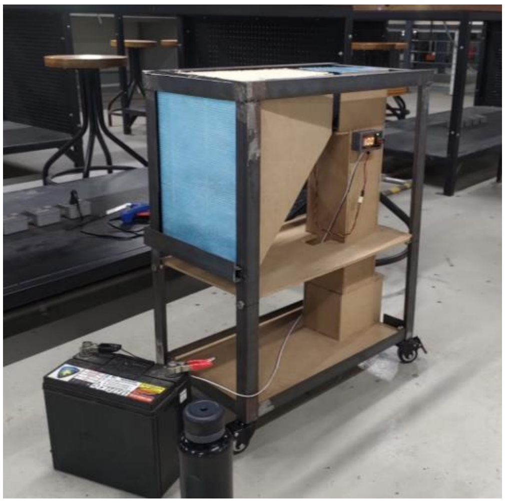


Fig. 1 Final prototype testing equipment

3. Result and Discussion

To achieve all the objectives of the project, the following two different experiments have been conducted.

3.1 Solar Panel Efficiency Test

The major measurement of this test was to measure the operating current from the solar panel which is a photovoltaic current. The test required the solar panel, solar charge controller, car battery, 100W yellow light source and a multi-meter.

This test is to measure the current output from the solar panel. A multimeter has been choosen. The measured readings are recorded in two decimal points for ease of comparison and calculations when compared to solar panel's performance under lab test conditions.

3.2 Air Purifier Test

The air purifier test is conducted to test the ammonia level of the inlet and outlet of the air purifier unit and assess the filtration efficiency. The test has been conducted by using 2 MQ135 Sensors placed at the inlet and outlet of the air purifier. The ammonia content readings would be compared between outlet and inlet to show the differences.

The tests conducted have followed Chen et al. experimental procedure [6]. The room should be clear from any object like furniture and to be clean and left untouched for 24 hours before initiating the filtration test.

The room chosen for the test is having 33.5 m^3 of volume. with trapezium shape. The air purifier is placed at the middle of the room and a tray is placed 60 cm away from the inlet of the air purifier. The test is run with sealed room to ensured there's no external particle movement from outside of the room.

The following detailed procedure has been followed to run the air purifier test.

1. Set testing tray 60cm away from the inlet of the air purifier unit.
2. Set up Arduino Uno by connecting to the laptop.
3. Ensure connection between battery and PWM motor controller.
4. Set up burning cigarette on testing tray for 10 minutes of injection period.
5. Remove test subject and wait for 10 minutes of static period.
6. Wait for 6 hours for dynamic period.
7. Set 100 percent power on PWM motor controller and activate Air purifier unit.
8. Start collect data to the laptop.
9. Wait for 6 hours for Data collection period.
10. Stop collect data to the laptop.
11. Repeat Steps 4 to 10 using multipurpose floor cleaner.

3.3 Battery Discharging Test

The battery discharging test is used to test the car battery depth of discharge and fan running time under fully charged car battery. The test results were used to compare with theoretical results and design results relate to fan running time and car battery charge level. The control variable has been set for fan running with 36W and 12V connected to a fully charged 45 AH car battery for 6 hours. The test has run simultaneously with air purifier test. The voltage of the car battery is measured with multi-meter before connected to the fan. After 6 hours, the car battery charge is measured again using the multi-meter. The following detailed procedure has been followed to run the battery discharging test.

1. Measure and record car battery voltage using multi-meter.
2. Connect the car battery to the fan using clamp terminals.
3. Start the on/off button on PWM fan controller.
4. Wait for 6 hours.
5. Measure and record car battery voltage using multi-meter.

3.4 Solar panel efficiency test result.

Results of the solar panel verification test are compared with the provided manual solar panel specifications under lab test conditions. The comparison results of the measured and lab test solar panel current is found to be 77.7%, Voltage at 86.77%, and wattage at 67.37%. The efficiency of the solar panel can be calculated by using the following equation (1).

$$n = \frac{Power\ out\ of\ solar\ cell\ (watt)}{Solar\ Power\ incident\ on\ solar\ cell\ (watt)} \qquad (1)$$

The power output of the solar cell is measured at 53.90w while the solar power incident on solar cell was 80w as the solar panel have received more than maximum of 100w yellow light. The calculation of the efficiency was 0.6737 which was 67.37 percent. Therefore, the solar panel could only produce 53.90 watt under 100w yellow light. Following equation (2) and (3) are percentage error equations to calculate the percentage error between theoretical and experimental wattage from solar panel.

$$Percentage\ error = \frac{theoretical - experimental}{theoretical} \times 100\% \quad (2)$$

$$= \frac{80 - 67.37}{80} \times 100\% = 15.79\% \quad (3)$$

3.5 *Air purifier test results.*

Figure 2 shows the test results for the air purifier performance when the cigarette is used to cause the air pollution.

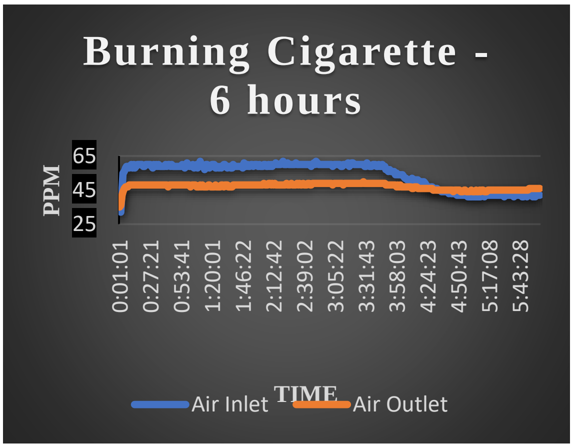


Fig. 2 PPM vs times of 6 hours burning cigarettes.

Figure 3 is showing the test results for the air purifier performance when the floor detergent is used to cause the air pollution.

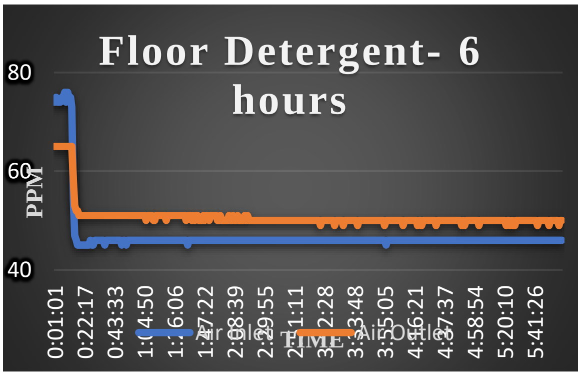


Fig. 3 PPM vs times of 6 hours floor detergent.

Control set of the air purifier test was recorded without any ammonia test subject has shown constant of 0 to 1 ppm fluctuation with majority of 0 ppm from air outlet and inlet sensor during the 6 hours testing. According to Li et al. [7] study found that typical indoor air ammonia content range around 10 to 70 ppb, which is 0.01 ppm to 0.07 ppm. The data has proven the accuracy of the sensor and let the user calibrate the sensor according to the data received. In this case, the MQ135 sensor detected majority of 0 ppm in both inlet and outlet which show little or close to none of ammonia content in the room air.

The floor detergent used in the test contains quaternary ammonium with 11.5 % (W/V). The test sample taken from the floor detergent is 100ml which contains 11.5ml of quaternary ammonium salt. The 11.5 % are converted to 115000ppm of quaternary ammonium salt in 100ml. The ammonium chloride in the quaternary ammoniums salts would travel at the rate of 0.228 ± 0.012 $cm^2 \cdot s^{-1}$ at 1 atm and 25℃. No dependence on relative humidity was observed over the range 10–92% [8].

As for cigarette, Cigarette ammonia content is varying from brands to brands. According to a study from Inaba et al. [9], the ammonia in cigarette is 870 ± 400 ppm per stick with 0.071 ± 0.006 $cm^2 \cdot s^{-1}$ diffusion rate [10].

The results are measured with Air inlet and Air Outlet of the air purifier. Both of the test pollutants chosen are containing significant amounts of ammonia. The test results have shown that the floor detergent has diffused more in the room as the highest detected amount of ammonia is 76 ppm at air inlet while burning cigarette recorded highest amount of ammonia is 60 ppm. It can be concluded that the floor detergent is having higher content amounts of ammonia than the cigarettes. The cigarette graph has shown that air inlet ammonia ppm was relatively high when compared to the air outlet ppm at the first 3 hours and 45 minutes of the test. The cigarettes smoke inside the room has been diffused into mid-air. The hot smoke released by the burning cigarettes get cooled down rapidly during the static period and dynamic period of the test. The cooled down smoke was descended to lower part of the room as smoke is heavier than the air. Both air inlet and air outlet ppm has started declined at 3 hours and 45 minutes. The decline of both air inlet and air outlet ppm meets a crossing point of 45 ppm at 4 hour and 30 minutes marked at the graph. The air outlet has remained constant of 45 ppm while air inlet has declined to 41 ppm at 5 hours mark. The cigarette smoke at lower level of the room has been cleared out while at the upper level of the room remains some residue of smoke. The in-balance of the smoke ppm in the room might been caused by minimal circulation of air in room.

Floor detergent graph is showing sudden drop from 73 ppm to 45 ppm at air inlet while 65 ppm to 51 ppm in 5 after 20 minutes of the test started. The ammonia molecule in floor detergent might have higher rate of diffusion than ammonia molecule found in cigarette smoke. Air inlet and air outlet of floor detergent test has shown constant 46 ppm and 50 ppm respectively for the rest of the test. The Figure 2 graph pattern is found similar to the pattern in Figure 3. The ammonia at lower level of the room has been cleared out while at the upper level of the room remains some of ammonia residue. The in-balance of the ammonia ppm in the room might been caused by minimal circulation of air in the room.

The following equation (4) can calculate the efficiency for both ammonia test with cigarette and floor detergent in 6 hours.

$$n = \frac{C_{Hin} - C_{Lout}}{C_{Hin}} \times 100\% \qquad (4)$$

Where C_{Hin} is highest ammonia ppm reading at the inlet of air purifier and C_{Lout} is lowest ammonia ppm reading at the outlet of air purifier. Table 1 is showing the efficiency of both ammonia test with cigarette and floor detergent in 6 hours.

Air purifier ammonia test	Efficiency (Percentage, %)
Burning Cigarette	43.55
Floor Detergent	35.53

Table 1 Efficiency of both ammonia test with cigarette and floor detergent in 6 hours.

The floor detergent might have a higher rate of diffusion than ammonia molecule found in cigarette smoke.

4. Conclusion

4.1 Conclusion

The solar powered air purifier has been designed, fabricated, and tested. It has been proven that the solar power could produce enough energy to run the low-cost air purifier that could be built by using locally available materials and resources.

The test results have shown 67.37% efficiency from solar panel in charging the used battery.

The MQ135 sensor used in air purifier test has shown the efficiency of cleaning ammonia pollutant in the air as 43.55% for burning cigarette and 35.33% for floor detergent.

It also found that floor detergent might have higher rate of diffusion than ammonia molecule found in cigarette smoke.

The proposed solar powered air purifier prototype is found useful and of low cost. It provides affordable and economical energy to power nearby applications for the developing country. Not only that, but the electricity power generation process does also not pollute the environment due to zero usage of fossil fuels and can be a good implementation of the sustainable development goals (SDG's).

Nomenclature

Symbol	Description	Dimension / Unit
PM 10	Particulate Matter 10 Micron	-
S02	Sulfur Dioxide	-
CO2	Carbon Monoxide	-
NO2	Nitrogen Dioxide	-
PM 10	Particulate Matter 2.5 Micron	-
HEPA	High Efficiency Particulate Air Filter	-
MERV	Minimum Efficiency Reporting Value	-
DC	Direct Current	
ppm	Parts Per Million	ppm
µ	Micro Meter	µm
IDE	Integrated Development Environment	-

Acknowledgment

The authors would like to express the gratitude towards the school of engineering at UCSI University for the support throughout the research. Special thanks would be extended to CERVIE office at UCSI University for the endless support to research.

References

1. WHO global air quality guidelines. Particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. Executive summary. (2021). Geneva: World Health Organization. Licence: CC BY-NC-SA 3.0 IGO
2. Malaysia Ambient Air Quality Standard. (2021). Ministry of Environment And Water: Air Quality Standards.
3. Malaysia Ambient Air Quality Standard. (2021). Ministry of Environment And Water: Air Quality Standards.
4. Prinz, A. L., & Richter, D. J. (2022). Long-term exposure to fine particulate matter air pollution: An ecological study of its effect on COVID-19 cases and fatality in Germany. Environmental Research, 204, 111948.
5. ANSI/ASHRAE Standard 52.2-2017. (2017). ASHRAE
6. Chen, W., Zhang, J. S., & Zhang, Z. (2005). Performance of air cleaners for removing multiple volatile organic compounds in indoor air. ASHRAE transactions, 111(1), 1101-1114.
7. Li, M., Weschler, C., Bekö, G., Wargocki, P., Lucic, G., & Williams, J. (2020). Human Ammonia Emission Rates under Various Indoor Environmental Conditions. Environmental Science &Amp; Technology, 54(9), 5419-5428.
8. Spiller, L. (1989). Determination of Ammonia/Air Diffusion Coefficient Using Nafion Lined Tube. Analytical Letters, 22(11-12), 2561-2573.
9. Inaba, Y., Uchiyama, S., & Kunugita, N. (2018). Spectrophotometric determination of ammonia levels in tobacco fillers of and sidestream smoke from different cigarette brands in Japan. Environmental Health And Preventive Medicine, 23(1).
10. John, E., Coburn, S., Liu, C., McAughey, J., Mariner, D., & McAdam, K. et al. (2018). Effect of temperature and humidity on the gas–particle partitioning of nicotine in mainstream cigarette smoke: A diffusion denuder study. Journal Of Aerosol Science, 117, 100-117.

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.

Ik Chu Aung

Ik Chu Aung received his Bachelor’s Degree in Mechanical Engineering in 2022 from UCSI University, Kuala Lumpur. His field of research is renewable energy.

Sarah ‘Atifah Saruchi

Sarah ‘Atifah Saruchi 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 works as a lecturer under the Department of Mechanical and Mechatronics, Faculty of Technology and Built Environment, UCSI Kuala Lumpur. Her research interests include control, vehicle and artificial intelligence.

Noor Idayu Mohd Tahir

Noor Idayu Mohd Tahir 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.

Design and Fabrication of a Mutual Control Electronic Circuit for Solar and Electrical Water Heating

Ammar A.M. Al-Talib
Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment, UCSI University, 56000 Cheras, Kuala Lumpur, Malaysia[*]

Sarah ‘Atifah Saruchi[*]
Faculty of Manufacturing and Mechatronic Engineering Technology, Universiti Malaysia Pahang, 26600 Pekan, Pahang, Malaysia

Cik Suhana Hassan
Department of Mechanical and Mechatronics, Faculty of Engineering, Technology and Built Environment, UCSI University, 56000 Cheras, Kuala Lumpur, Malaysia

Nor Aziyatul Izni
Centre of Foundation Studies, Universiti Teknologi MARA, Cawangan Selangor, Kampus Dengkil, 43800 Dengkil, Selangor, Malaysia
E-mail: ammart@ucsiuniversity.edu.my, sarahatifah@ump.edu.my, suhana@ucsiuniversity.edu.my, naizni@uitm.edu.my

Abstract

This research is a temperature controller that will be implemented to ensure that the water temperature of the solar water heating unit is maintained at the desirable level at all times of use. This control circuit is designed to control the On/Off action of the immersed electrical heater according to specific temperature range. A temperature sensor will sense the water temperature constantly and send signal to a micro-controller unit. The micro-controller will process the data according to a written program and control the actions of electrical heater. At the same time, temperature reading will be displayed through LCD and real-time data can be viewed from a computer via serial port. During times of sufficient sunlight, solar energy will be the main source used for heating water; otherwise, there will be an automatic switching to the electrical operated immersion heater. This controller will give reliability to users of solar water heating systems.

Keywords: Solar Water Heater (SWH), Temperature control, energy, PIC.

1. Introduction

Energy is the basis for the industrial revolution and modern civilization, and without it, modern life will almost cease to exist. Its availability is of great importance to humans where is required for domestic, commercial, as well as industrial applications. Oil prices have increased abundantly over the years. There is also a major disruption of oil production. These have caused major contributions for nations to review their energy policies for the incorporation of energy efficiency measures to combat not only the uncertainties of oil prices but also to conserve the fast depletion of global fossil fuel reserves. Solar energy is an unlimited, cheap, and clean source of energy which is utilized to replace conventional energy [1]. Not only does the cost of fuel drive the researchers to explore an alternative source of energy, but the pollution and ozone layer problems are other factors that should be considered in searching for cheap and clean sources.

---

[*] Corresponding Author

Heat from the sun is used to heat water in the solar water heater. Utilizing solar energy for domestic use is a well-established technique. Solar water heaters harness the sun's energy and use it directly for heating water. Solar water heaters have been commercially available since the 1800s. As early as 1860, air blowing over a sun-heated iron was used to heat homes. In the 1870s, an Englishman, Maughan, has invented the first instant water heater. His invention, however, influenced the designs of Edwin Ruud who was the inventor of the automatic storage water heater in 1889 [2]. Solar water heating is effective in places like Australia, California, and Malaysia where there is lots of sunshine [3]. Many designs of solar systems have become popular and competitive with the conventional heating systems due to the developments and improvements that made them economic and effective. The technical feasibility for many solar systems in domestic applications has been established. The need for energy in developing countries is increasing and depending on fossil fuels for the only energy source is not enough and they are getting depleted fast. In Malaysia, fossil fuels such as coal, can only supply energy up to the year 2030 [4]. Malaysia, being a developing country, increases its demand for energy from year to year. In the year 1995, the total amount of energy needed was 22,164 kilotons and increased to 25,558 in 1998. The main energy source in Malaysia is crude oil and petroleum. Although the oil produced is sufficient for local needs, in the future, the country will require more energy due to the growing economy. Being a tropical country, the solar radiation received in Malaysia is very sufficient. The average solar energy radiation yearly is approximately 500-700 Wm-2 and the total amount of solar radiation is between 4-8 hours [3].

Solar water heater is one of the most common applications of solar energy. The reason of its wide application is because of its viability and economics advantages [5]. In Malaysia, there are extensive research conducted to improve the performance of the solar water heater system [6][7]. There are two wide categories of solar water heaters which are active and passive heaters. The existence of these two methos make it possible to reduce the building energy consumption cost [8]. The collector absorbs the sun’s energy and changes it into heat energy. Over the years, various models have been studied as far as the collector type of the water heater is concerned [9].

2. Methodology

To ensure the stability of temperature for users at all times, an electronic circuit has been designed, which was composed of a temperature sensor and a microcontroller as the controlling unit. The system will start operation by sensing the temperature of water inside the water tank. The temperature sensor used is DS18S20 from Dallas semiconductor. This silicon temperature sensor read as a 9-bit digital value and sends reading to micro-controller for processing. A PIC16F84A is the micro-controller used in this control circuit. The PIC is programmed to control the ON-OFF of heater, display temperature reading on the LCD and provide an interface between control unit and computer via serial port. The software used to program the microcontroller is PicBasic Pro Compiler. This silicon temperature sensor required one data line (and ground) for communication with a central microprocessor. It has an operation temperature range of -55ºC to +125ºC and is accurate to ±0.5ºC over the range of -10ºC to +85ºC. In addition, the DS18S20 can derive power directly from data line, eliminate the need for an external power supply. Temperature is continuously being measured and can be read at any time. The temperature sensor output is represented by a 9-bit digital value (1 sign bit and 8 magnitude bits). The PIC16F84A is used as the main processing unit. The software tool used in-line with this microcontroller is PicBasic Pro Compiler. This microcontroller accepts the input from sensing element and displays the reading in both Celsius and Fahrenheit through LCD and provide interface to computer via serial port. A transistor is set HIGH/LOW according to program written to determine the action of heating system. PIC16F84A contains two set of I/O ports; Port A and Port B. Port A provided five I/O pins (RA0-RA4) and eight I/O pins from Port B (RB0-RB7). Some special features of this type of microcontroller are 10,000 erase/write cycles Enhanced FLASH Program memory typical, 10,000,000 typical erase/write cycles

EEROM Data memory typical, and In-Circuit Serial Programming (ICSP). RB6 is input to receive signal from sensor. RA0-RA4, RB3 are output pins for LCD and RB1 provide output to serial port. RB4 is output for electronic switch while RB5 will control the LED to indicate the On / Off of heater.

The display unit of this system is JHD162A SERIES 16x2 LCD. Microcontroller controlled LCDs are widely used in many applications, due to their low power consumption and flexible graphics displays. They are a compact device and can be interfaced to microcontroller easily. LCD can show both numerical and characters and display many messages simultaneously. Some functions like clearing the display, shifting the cursor, and displaying the ON/OFF functions are available on it.

This intelligent LCD module is supplied with a 5V power supply. RS, R/W, and E are used for control purposes and DB0-DB7 is data lines. A relay acts as key component to control the ON/OFF of electrical heater. Fig. 1 is showing the hardware block diagram.

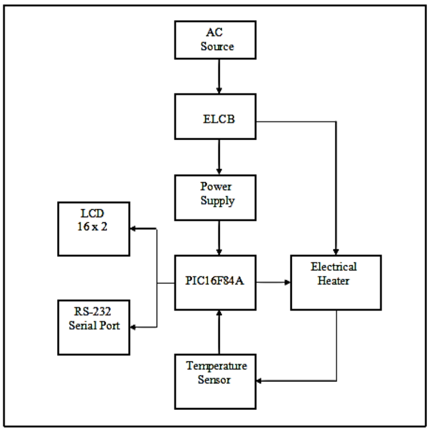

AC Source
ELCB
Power Supply
LCD 16 x 2
PIC16F84A
Electrical Heater
RS-232 Serial Port
Temperature Sensor

Fig. 1. The hardware block diagram.

The main power source for this system and electrical heater is AC supply. An ELCB (Earth Leakage Circuit Breaker) has been installed in this system for surge protector. The power supply part consists of a 6V center-tapped step-down transformer and 5V voltage regulator as supply for microcontroller and LCD. Temperature sensor derives power directly from data line. Fig. 2 is showing the Control Circuit schematic diagram.

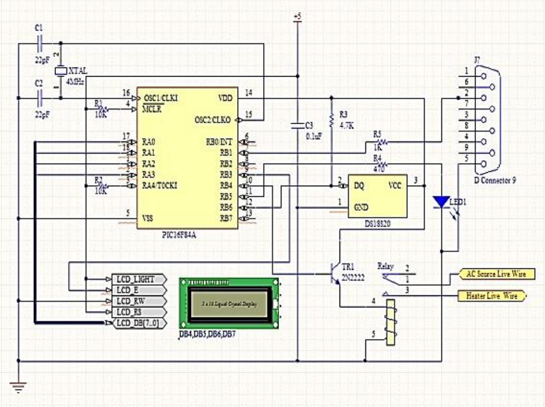

C1
22pF
XTAL
4MHz
C2
22pF
R1
10K
R2
10K
+5
OSC1/CLKI
MCLR
VDD
OSC2/CLKO
RA0
RA1
RA2
RA3
RA4/TOCKI
VSS
RB0/INT
RB1
RB2
RB3
RB4
RB5
RB6
RB7
PIC16F84A
C3
0.1uF
R3
4.7K
R5
1K
R4
470
DQ
VCC
GND
DS18S20
J7
D Connector 9
LED1
TR1
2N2222
Relay
AC Source Live Wire
Heater Live Wire
LCD_LIGHT
LCD_E
LCD_RW
LCD_RS
LCD_DB[7..0]
DB4,DB5,DB6,DB7

Fig. 2. The control circuit schematic diagram.

3. Results and Discussions

A PicBasic Pro Compiler (PBP) is used. It is a powerful programming language to program a PIC microcontroller. This English-like BASIC language is easier to read and write than assembly language. PicBasic Pro Compiler produces codes that may be programmed into a wide variety of PIC microcontrollers and various on-chip features likes A/D converters, hardware timers and serial ports. For this system, the program is written according to specific temperature range (40ºC - 60ºC) for display the reading and control the ON/OFF of heater. A .HEX file will be created after compile the written source code. The .HEX file is next programmed into PIC16F84A by IC-Prog1.05C software. An Integrated Thermosyphon Solar Water Heater prototype has been fabricated to implement the designed controller. Fig. 3 is showing a schematic diagram of the prototype tested.

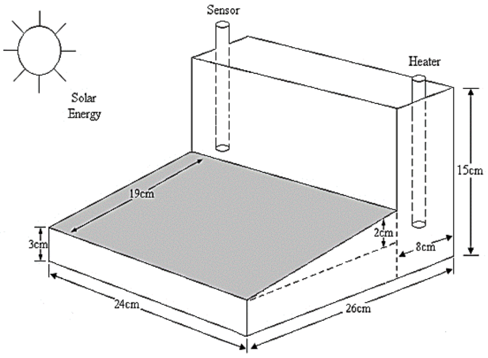

Sensor
Heater
Solar Energy
15cm
19cm
3cm
2cm
8cm
24cm
26cm

Fig. 3. A schematic diagram of the prototype used.

Fig. 4 and Fig. 5 are showing some of the output results.

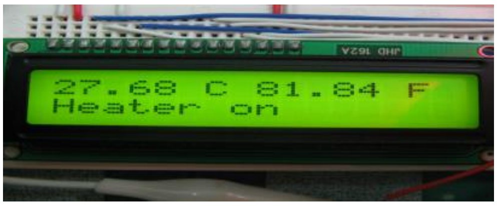

27.68 C 81.84 F
Heater on

Fig. 4. Temperature reading display 1.

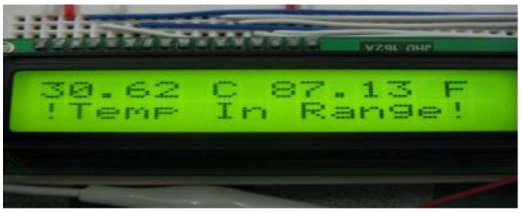

30.62 C 87.13 F
!Temp In Range!

Fig. 5. Temperature reading display 2.

Fig. 6 is showing the MATLAB Simulink output.

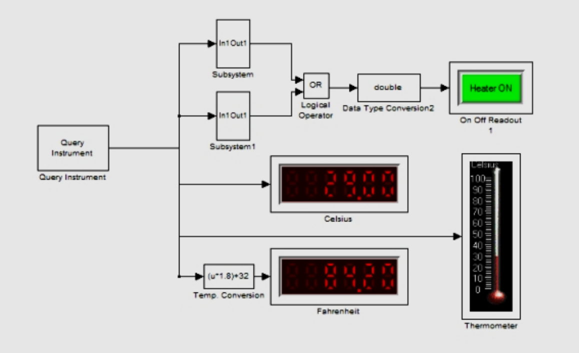

In1 Out1
Subsystem
OR
Logical Operator
double
Data Type Conversion2
Heater ON
On Off Readout 1
Query Instrument
Query Instrument
In1 Out1
Subsystem1
29.00
Celsius
Celsius
100
90
80
70
60
50
40
30
20
10
0
(u*1.8)+32
Temp. Conversion
84.20
Fahrenheit
Thermometer

Fig. 6. Control circuit interface in MATLAB Simulink.

A reliability test has been conducted to ensure the proper functionality of the designed control circuit. Fig. 7 is showing the reliability test results.

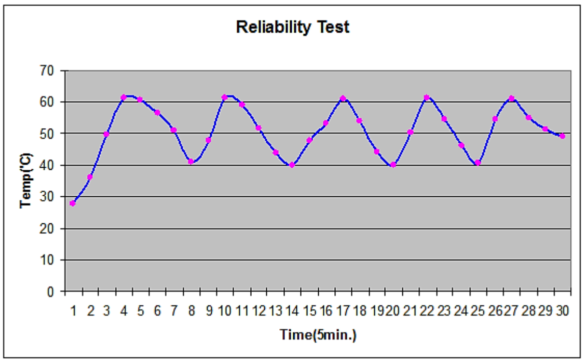


Fig. 7. The control circuit reliability graph.

4. Conclusion

The mutual control circuit for solar and electrical water heater was complete successfully. This temperature control is fully functioning and reliable to control the actions of electrical heater automatically according to specific temperature ranges if the water temperature cannot be maintained by the solar heater. With the mutual function between solar and electrical water heater, hot water is available and ready to be used at all times. This product is very practical and convenient to ensure that people can enjoy their hot bath after a busy working day without worrying about the weather condition. The circuit can be considered as an implementation of the sustainable development goals (SDG’s) in encouraging people to use the renewable energy sources.

Acknowledgement

The authors would like to express the gratitude towards the support from the UCSI University, Universiti Malaysia Pahang and Universiti Teknologi MARA and UCSI CERVIE who are supporting the research by all means.

5. References

[1]. F. Keith and J. Kreider, Principles of solar engineering, Washington DC: Hemisphere, 1978.

[2]. W. Larry, "Solar water heaters: What are the benefits?," 30 January 2020. [Online]. Available: https://www.thoughtco.com/solar-water-heaters-benefits-1204179. [Accessed 24 January 2023].

[3]. K. Sopian, M. Syahri, S. Abdullah, M. Y. Othman and B. Yatim, "Unglazed Fiberglass Reinforced Polyester Solar Water Heater with Integrated Storage System," *Journal of Energy Engineering*, Vol. 133, no. 1, p. 25, 2007.

[4]. T. N. Veziroglu, "Alternative energy sources VII," in *Proceedings of the 7th Miami International Conference on Alternative Energy Sources*, Miami Beach, Florida, USA, 1987.

[5]. N. I. S. A. Azha, H. Hussin, M. S. Nasif and T. Hussain, "Thermal Performance Enhancement in Flat Plate Solar Collector Solar Water Heater: A Review," *Processes*, Vol. 8, No. 7, p. 756, 2020.

[6]. M. S. Naghavi, B. C. Ang, B. Rahmanian, S. Naghavi, S. Bazri, R. Mahmoodian and H. S. C. Metselaar, "On-demand dynamic performance of a thermal battery in tankless domestic solar water heating in the tropical region," Applied Thermal Engineering, Vol. 167

[7]. M. S. Hossain, L. Kumar and A. Nahar, "A Comparative Performance Analysis between Serpentine-Flow Solar Water Heater and Photovoltaic Thermal Collector under Malaysian Climate Conditions," *International Journal of Photoenergy*, pp. 1-9, 2021.

[8]. A. M. Manokar and A. Karthick, "Review on progress in concrete solar water collectors," *Environmental Science and Pollution Research*, Vol. 28, pp. 22296-22309, 2021.

[9]. A. Shafieian, M. Khiadani and A. Nosrati, "Strategies to improve the thermal performance of heat pipe solar collectors in solar systems: A review," *Energy Conversion and Management*, Vol. 183, pp. 307-331, 2019.

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.

Sarah ‘Atifah Saruchi

Ts. Dr. Sarah ‘Atifah Saruchi 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 works as a lecturer under the Department of Mechanical and Mechatronics, Faculty of Technology and Built Environment, UCSI Kuala Lumpur. Her research interests include control, vehicle and artificial intelligence.

Cik Suhana Hassan

Dr. Cik Suhana Binti Hassan 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.

Nor Aziyatul Izni

Ts. Dr. Nor Aziyatul Izni received her Doctor of Philosophy at Malaysia-Japan International Institute of Technology (MJIIT), Universiti Teknologi Malaysia (UTM), Malaysia. Her research areas includes signal processing, applied statistical methods, applied mathematics, fusion algorithm, Artificial Intelligence as well as prediction and forecasting methods.

Optimization of the Major Factors Affecting the Recycling of Disposed (LDP) Plastics

Ammar A.M.Al Talib, Ain Atiqa, Chua Ray Sern Grayson
Department of Mechanical & Mechatronics, Faculty of Engineering Technology & Built Environment, UCSI University, Kuala Lumpur, Malaysia
E-mail: ammart@ucsiuniversity.edu.my, AinAtiqa@ucsiuniversity.edu.my, 1001438575@ucsiuniversity.edu.my

Abstract

Low density polyethylene (LDPE), which is widely adopted in many daily life products, known take hundreds of years to decompose in the landfill. As LDPE continues to expand its limitless use over all fields and industries, it also brings tremendous damages onto land and marine ecosystems. Although various efforts were carried out to recycle LDPE, but the operations were proved to be troublesome due to low demand of recycled LDPE pallets, obstacles to seek demanding markets and difficulties to process contaminated and dyed LDPE wastes. Existing recycling methods are still in research-stage to be able to effectively counter the ever-increasing problems of LDPE. Hence, this study aimed to investigate optimum recycling conditions of disposed LDPE plastics and explore potential green applications for the recycled LDPE. As proposed in this study, disposed LDPE bags are shredded, washed, dried and then inserted into oven for melting at heating temperature of 170^0 to 200^o C for different durations inside the oven to seek the optimum heat recycling conditions. Mechanical tests are conducted onto the recycled LDPE samples, which include hardness test. compression test and tensile test. The recycled LDPE has achieved 90.2 Shore-A hardness point, 9.2599 MPa of compressive strength and tensile strength of 9.0705 MPa. After comparison with similar available products in the market, this recycled LDPE can be utilized in the manufacturing of skateboard, shopping cartwheels, tabletops, garden paving tiles, and substitutes for bricks and woods in construction works.

Keywords: Recycling of LDPE; Shore Hardness; Compressive Strength; Tensile Strength

1. Introduction

Over the past decades, industries worldwide have been widely adopting plastics. Plastics are one of the most versatile materials in the world where the global production of plastic-made products surpassed 335 million tonnes in year 2017 [1]. Study from American Chemistry Council shows that replacing existing plastic packaging in the U.S. with non-plastic alternatives requires 4.5 times raise in packaging materials by weight and may result in 130 percent increase in global warming potential. Large-scale eliminations for use of plastics are basically impossible. To worsen the situation, only a mere one percent of LDPE products is recycled [2], as compared to 27% of recycled rate of HDPE which is the most recycled plastic [3]. Although LDPE products were collected frequently for mechanical recycling [4], only rigid, clean and soft LDPE wastes are collected to be recycled in most of the countries [5]. The contaminated and dyed LDPE must be washed and sorted before undergoing mechanical recycling [6], and hence most companies avoid the collection and recycling of LDPE.

Consequently, the market for the recycled LDPE is infrequent and not consistent with regards to time. Biodegradable bags also caused recycled LDPE to be unwanted. Most governments and company policies encouraged the usage of biodegradable plastic bags, instead of the conventional LDPE plastic bags. Laws and rules are implemented to greatly reduce the application of the conventional LDPE plastic bags in certain states in Malaysia [7].

There are three main available methods to dispose LDPE plastics: mechanical recycling, energy recovery and landfill [8]. Energy recovery operations are hard to maintain since there are multiple technological limitations and high operating costs [9]. Since most of the plastic products are not biodegradable [10], plastics in the landfill will not degrade for hundreds of years. Mechanical recycling is the most environmental-friendly approach in reducing the number of plastics, but it may cause soil contamination [11] and emission of plastic chemicals to the surrounding [12],[13].

Efforts to seek effective recycling procedures for LDPE was done by a few researchers. an investigation on the feasibility of reprocessing low-density polyethylene (LDPE) waste materials was done by Bassey et.al [14]. The targeted product of the research was PVC-liked ceiling tiles by addition of different grain size of sawdust into the LDPE wastes. Researcher Singh et. al has conducted research on the recycling of HDPE solid waste for additive manufacturing applications [5],[15]. Ragaert et.al investigated the topic of upcycling of contaminated post-industrial polypropylene waste by turning plastic waste into FDM filaments [16].

The objective of this paper is to investigate optimum recycling conditions of disposed LDPE plastics by using simple and available equipment and to assess some mechanical properties of the recycled LDPE to find useful green applications.

2. Methodology and Experimental Setup

2.1. *Materials*

Low-density Polyethylene (LDPE) plastics packaging bags are used in this study. The disposed LDPE packaging bags are obtained from a local packaging-bag manufacturing company. The LDPE packaging bags are available in red and white colors, both with thickness of 0.1 mm.

2.2. *Apparatus and Equipment*

The apparatus and equipment used for the research work are electrical oven, metal mould, baking paper, aluminium pans, weighing balance and G-clamps. Test machines utilized for mechanical tests include Teclock durometer, Kenco analog compression machine and Tinius Olsen tensile test machine.

2.3. *Specimen and Preparation*

The recycling procedure practiced in this study has started by shredding the ddisposed LDPE bag into small pieces, washed, and then dried to be used in further steps. To produce cylindrical shaped samples for compression tests, metal cylinder molds are used, whereas aluminium pans are used for the fabrication of board samples. The process has been continued by inserting the molds containing 100 grams of LDPE flakes into the oven and heated under the temperatures and time duration's shown in Table 1. Thorough compression has been used to ensure perfect bonding and homogenizing of samples. The specimens retrieved from the oven are cooled using natural air-cooling treatment at room temperature. Table 1 is showing the recycling conditions practiced in this study.

Table 1 Heat Treatment for Specimen Fabrication

Duration in Oven (min) / ***Oven Heating Temperature (℃)***	*40*	*50*	*60*
170	*X*	✓	✓
180	✓	✓	✓
190	✓	✓	✓
200	✓	✓	*X*

** ✓ indicates that samples are manufactured under the prescribed condition.*

**X indicates that the condition is excluded from sample preparation.*

2.4. *Mechanical Tests*

Mechanical tests are conducted on the produced samples to determine the mechanical properties of recycled LDPE, namely hardness test, compression test and tensile test, and to find the optimum recycling parameters.

2.4.1 Hardness Test

Hardness tests are conducted using Teclock Durometer GS-706G with Shore-A scale in comply to ASTM D2240 standard. The specimens to be tested are measured to ensure a thickness of at least 6mm. The samples are placed in parallel on a well-flattened test surface and ten readings are taken on the top and bottom surface of the sample. The mean hardness value of each condition is used for comparisons.

2.4.2 Compression Test

Kenco Compression Test Machine has been used to determinee the compressive strength of the recycled LDPE samples. All cylindrical samples have been cut and sanded to same height before the compression test to ensure minimum error. In the compression machine, constant speed loading is applied by compression of top and bottom platens. As comply to ASTM D695, the compressing loads are stopped once fracture, breakage or any sign of shearing occurred. Readings of maximum force applied has taken and the test is continued with other samples. The calculation of compressive strength for the samples is stated in ASTM D695, as shown in Eq. (1).

$$Compressive\ Strength, F = \frac{P}{A} \tag{1}$$

where;

P= *Maximum load*
A= *Original minimum cross-sectional area*

2.4.3 Tensile Test

Tinius Olsen Tensile Test Machine has been utilized to conduct tensile strength test of the dumbbell LDPE samples. Force against extension graphs were displayed in QMat Software from the connected computer. From the graphs in QMat, the maximum sustainable load for each condition is retrieved to determine the tensile strength of each specimen. Tensile strength is calculated according to formula stated in ASTM D638 standard, which is Eq. (2).

$$Tensile\ Strength, \sigma = \frac{P_{max}}{A_0} \tag{2}$$

where;

P_{max}= *Maximum load sustained*
A_0= *Average original cross-sectional area (gauge length)*

3. Result and Discussion

3.1. *Fabricated Specimen*

Figure 1 shows the fabricated cylindrical specimens fabricated under heating temperatures of 170℃ to 200℃ with time durations of 40 to 60 minutes. The cooling treatment performed in this study is natural air-cooling at room temperature. All samples prepared in this study possessed smooth outer surfaces and fully took the shape of the metal cylinder mould. Since red and white colour LDPE flakes are used in this study, all samples are looking marmoreal or marble-like patterns.

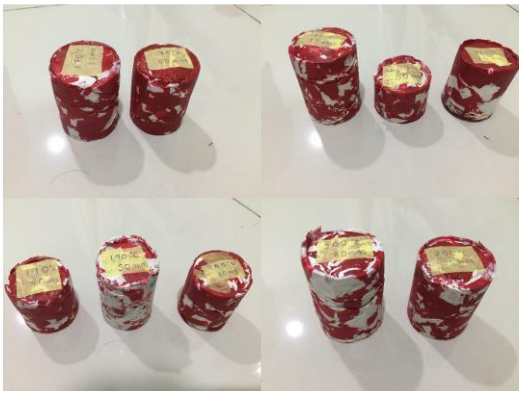


Fig. 1 Specimens fabricated under different heating temperatures and duration

Conducted experiments have shown that LDPE packaging bags can only start melting at a temperature of 170℃, with heating duration over 40 minutes. Similarly, it's found that the LDPE plastic is unable to be fully melted in the core part at heating temperature of 160ºC and heating duration of 50 minutes. The recycled LDPE has shown shrinkage after heating for 60 minutes at temperature of 200ºC.

Figure 2 shows the flat board shaped samples fabricated using aluminium pans and compression force. Part of the board samples are used to produce dumbbell-shaped test samples for tensile strength tests. The temperature range for the fabrication followed the conditions shown in Table 1.

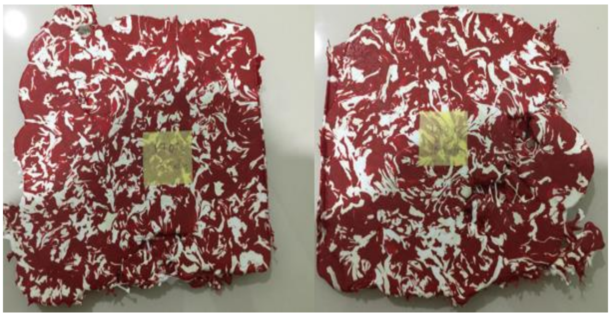


Fig. 2 Board samples fabricated

3.2. *Mechanical Test Result*

3.2.1 Hardness Test

Hardness values have been taken using Teclock Durometer GS-706G with Shore-A scale in comply to ASTM D2240 standard. The average hardness values are obtained from the mean of ten hardness measurements on different points on each of the specimens.

Figure 3 is presenting the relationship between hardness values, heating temperature and time duration. From the figure, the highest hardness value is 90.2 point, and it's achieved by the sample prepared under 190 ℃ and 50 mins. A general trend can be observed from the figure that the sspecimens fabricated using the heating temperature of 190 ℃ generally have higher Shore-A hardness values as compared to other temperatures. The lower hardness values for samples prepared with temperature of 170℃ is due to incomplete homogeneity of melted LDPE flakes during the heating process. The decrease in hardness values of samples manufactured under 200℃ can be attributed to thermal degradation.

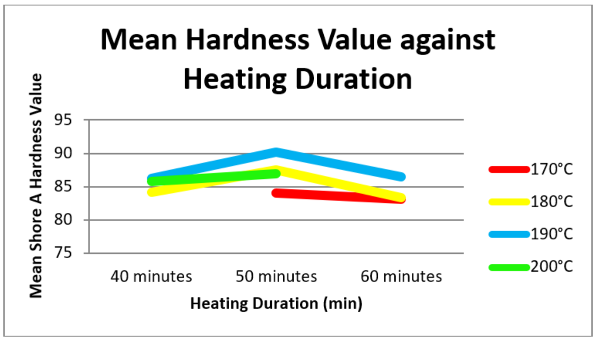


Fig. 3 Relationship between hardness value, temperature and time duration

3.2.2 *Compression Test*

Kenco Compression Test Machine has been used to determine the compressive strength of the recycled LDPE samples. The diameter of all specimens produced is 55 mm and thus the cross-sectional area of recycled LDPE samples is 2375.83 mm^2. The Compressive strength has been calculated using Eq. (1) and following the ASTM D695 standards.

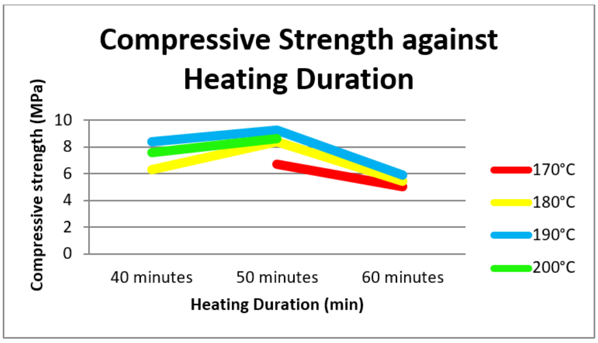


Fig. 4 Relationship between compressive strength, temperature and time duration

Figure 4 shows the relationship between compressive strength, temperature and time duration. It can be noticed that heating temperature of 190ºC and heating duration of 50 mins produced the sample with highest compressive strength of 9.26 MPa. Generally, samples fabricated with the duration of 50 min. have higher compressive strength when compared to time duration of 40 min. and 60 min. of the same heating temperature. Heating duration of 60 minutes and above may cause polymer degradation on the samples, thus breaking the long chain branching of LDPE and reduce its mechanical properties.

3.2.3 *Tensile Test*

The tensile strength of recycled LDPE for each recycling condition is determined by using Eq. (2) and is shown in Figure 5.

Figure 5 is showing the results of tensile strength, heating temperature and time duration. The highest attained tensile strength is 9.071 MPa. It is achieved by sample prepared under 190°C and 50 mins. The trend shown in tensile test is roughly like trend illustrated in Shore A hardness test and compression test. This phenomenon may be explained by the correlations between indentation hardness value with tensile and compressive strength [1].

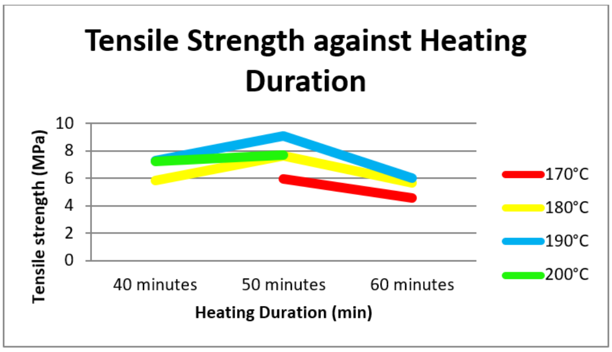


Fig. 5 Relationship between tensile strength, temperature and time duration

3.3. *Potential Applications*

With regards to the results obtained from the mechanical tests conducted, thereupon the recycled LDPE samples are found comparable with materials such as virgin LDPE, virgin polyurethane (PU), ultra-high-molecular-weight polyethylene (UHMW-PE) and ethylene propylene rubber (EPR). Although the obtained hardness value is found lower than virgin LDPE, the recycled LDPE is still comparable to certain classes of PU and EPR. The proposed applications in terms of hardness value includes idler rollers, shopping cartwheels, skateboard's wheels and O-rings.

The recycled LDPE specimens are comparable with C62 grade bricks, light bricks, brickwork's and D70-grade Hardwood. The obtained compressive strength of 9.26 MPa for the recycled LDPE is higher than the materials mentioned. Since all the mentioned materials are commonly utilized as construction materials, as partitioning and manufacturing of doors and door frames. Thus, recycled LDPE has the potential to be utilized for the same applications.

The recycled LDPE samples have shown an optimum tensile strength of 9.071 MPa, which is comparable to glass, wood, medium density fiber-wood, marble, limestone and alumina boards. All the mentioned materials are commonly used in the manufacturing of furniture and fixtures, where strength and long shelf-life are the most considered parameters.

4. Conclusion

By interpreting the data obtained and the graphs plotted for the mechanical tests conducted, it can be concluded that recycling at heating temperature of 190°C with heating duration of 50 minutes and natural air-cooling treatment are the best conditions for recycling of disposed LDPE plastic using the method introduced in this study in accordance with ASTM standards. Specimens fabricated using the heating temperature of 190°C in 50 minutes achieved 90.2A in Shore hardness test, 9.26 MPa for compressive strength and 9.071 MPa for tensile strength.

Potential green applications which can be suggested according to the mechanical properties obtained, are manufacturing of skateboard, shopping cartwheels, tabletops, garden paving tiles, and substitutes for bricks and woods in construction works.

It can be concluded that the promising results obtained are a good implementation towards achieving the Sustainable Development Goals (SDG's).

Acknowledgment

The authors would like to express the gratitude towards the school of engineering at UCSI University for the support throughout the research. Special thanks would be extended to CERVIE office at UCSI University for the endless support to research.

References

1. Koch, T., & Seidler, S. (2009). "Correlations of Indentation Hardness and Yield Stress in Polymers", 1–6.
2. Gu, F., Guo, J., Zhang, W., Summers, P. A., & Hall, P. (2017). "From waste plastics to industrial raw materials: A life cycle assessment of mechanical plastic recycling practice based on a real-world case study." Science of the Total Environment, 601–602, 1192–1207.
3. Meran, C., Ozturk, O., & Yuksel, M. (2008). "Examination of the possibility of recycling and utilizing recycled polyethylene and polypropylene." *Materials and Design*, 29(3), 701–705.
4. Metin, E., Eröztürk, A., & Neyim, C. (2003). "Solid waste management practices and review of recovery and recycling operations in Turkey." *Waste Management*, 23(5), 425–432.
5. Singh, N., Hui, D., Singh, R., Ahuja, I. P. S., Feo, L., & Fraternali, F. (2017). "Recycling of plastic solid waste: A state of art review and future applications." *Composites Part B: Engineering*, 115, 409–422.
6. Lazarevic, D., Aoustin, E., Buclet, N., & Brandt, N. (2010). "Plastic waste management in the context of a European recycling society: Comparing results and uncertainties in a life cycle perspective." *Resources, Conservation and Recycling*, 55(2), 246–259.
7. Agamuthu, P., & Victor, D. (2010). "Policy Evolution of Solid Waste Management in Malaysia", 916–924.
8. Al-Salem, S. M., Lettieri, P., & Baeyens, J. (2009). "Recycling and recovery routes of plastic solid waste (PSW): A review." *Waste Management*, 29(10), 2625–2643.
9. Lopez, G., Artetxe, M., Amutio, M., Bilbao, J., & Olazar, M. (2017). "Thermochemical routes for the valorization of waste polyolefinic plastics to produce fuels and chemicals. A review." *Renewable and Sustainable Energy Reviews*, 73(October 2016), 346–368.
10. Andrady, A. L. (1994). "Assessment of environmental biodegradation of synthetic polymers." Journal of Macromolecular Science, Part C, 34(1), 25–76.
11. Tang, Z., Zhang, L., Huang, Q., Yang, Y., Nie, Z., Cheng, J., Chai, M. (2015). "Contamination and risk of heavy metals in soils and sediments from a typical plastic waste recycling area in North China." Ecotoxicology and Environmental Safety, 122, 343–351.
12. Grigore, M. (2017). "Methods of Recycling, Properties and Applications of Recycled Thermoplastic Polymers." *Recycling*, 2(4), 24.
13. Tsai, C. J., Chen, M. L., Chang, K. F., Chang, F. K., & Mao, I. F. (2009). "The pollution characteristics of odor, volatile organochlorinated compounds and polycyclic aromatic hydrocarbons emitted from plastic waste recycling plants." *Chemosphere*, *74*(8), 1104–1110.

14. Bassey, G. I., Egbe, J. G., Ewa, D. E., Ettah, E. B., & Antigha, R. E. (2017). "Reprocessing of Low-Density Polyethylene (LDPE) Waste Materials for The Formation of PVC Ceiling Tiles Using Sawdust as A Reinforcement." Journal of Architecture and Civil Engineering, Volume 3 ~ Issue 4 (2017) pp: 09-15.
15. Singh, N., Singh, R., & Ahuja, I. P. S. (2018). "Recycling of polymer waste with SiC/Al2O3reinforcement for rapid tooling applications." *Materials Today Communications*, 15(February), 124–127.
16. Ragaert, K., Hubo, S., Delva, L., Veelaert, L., & Du Bois, E. (2018). "Upcycling of contaminated post-industrial polypropylene waste: A design from recycling case study." *Polymer Engineering and Science*.

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.

Ain Atiqa

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.

Chua Ray Sern , Grayson

Chua Ray Sern, Grayson received his Bachelor's degree in Mechanical Engineering in 2019 from UCSI University, Kuala Lumpur. His field of research is renewable energy.

Investigation of the Mechanical Properties and Applicability of HDPE Recycled Plastic Bags

Ammar A.M. Al- Talib, Nor Fazilah Abdullah, Amar Ridzuan Abd.Hamid, Wan Wai Kit, Tan Jun Hoe.
Department of Mechanical and Mechatronics, Faculty of Engineering,
Technology and Built Environment, UCSI University, 56000 Kuala Lumpur, Malaysia

E-mail: E-mail: ammart@ucsiuniversity.edu.my, norfa@ucsiuniversity.edu.my, amar@ucsiuniversity.edu.my,
ivanwaikit2012@hotmail.com, joonhoe94@gmail.com

Abstract

Plastics have been one of the most used materials in manufacturing products. Disposal of plastic wastes have led to a major problem to surrounding environment. Implementation of plastic recycling is important to the current era where disposed plastic is more than recycled or reused plastics. In this research, optimum heating conditions will investigate for the manufacturing of plastic boards from recycled HDPE plastic bags. Other possible applications are also further researched with comparison to products available in the market. Multiple specimens are produced and undergone microscopic observation, hardness, and compression test. HDPE plastic bag specimens are recycled using a normal home oven and its thermal insulation has been used to prevent heat loss. Various temperatures and duration in oven are tested in this research. The HDPE plastic bags are shredded, placed in a metallic pan, and melted in the oven with controlled heating temperatures and duration. The fully melted material is then poured in a mold and well compressed to produce planks or bars for the required mechanical tests. The experiments have shown that higher temperatures have produced better hardness results. The duration shown influence on the materials in terms of homogeneous bond. A heating temperature of 230˚C and 60 minutes duration in the oven give the better mechanical properties. Solid plastic boards for tables and chairs beside shopping trolley wheels are the suitable applications recommended for the recycled HDPE plastics.

Keywords: HDPE; Plastic Bags Recycling; Plastic Waste.

1. Introduction

In today's sustainable society, plastic is a substance that is frequently utilized and has ingrained itself into our way of life. Global plastic production has gradually increased through the years due to its low cost, versatility, high physical and chemical stability, and simple manufacturing and processing procedures [1]. Common commercial used thermoplastics are polyethylene terephthalate (PETE), high-density polyethylene (HDPE), polyvinyl chloride (PVC), low-density polyethylene (LDPE), polypropylene (PP), and polystyrene (PS), which can be identified with labels found on the products. Due to chemical properties of plastics, it takes decades to degrade. Due to their widespread commercial use, the present amount of trash from them poses a serious environmental risk if improperly disposed of. Common household used plastics are contributing to that factor as most plastic wastes are discarded without being recycled [2]. Plastic recycling process starts with sorting based on shape, size, density, colour, and chemical composition [3].

A astounding 322 million tonnes of plastics were produced globally in2015, an increase of 3.5% from the previous year. Plastic waste generated in Europe annually is 25 million tonnes on average. 29.7% was effectively recycled, 39.5% was used for energy recovery, and remaining of 30.8% was landfill [4]. E-waste has generated an estimated volume of 44.7 million tonnes in year 2016 and with annual growth rate of 3~4%, expectation will be 52.2 million tonnes in year 2021. Despite having a staggering volume in millions, overall global e-waste recycling rate is only 15% [1]. Recycling rates in the United States, China, and Europe are 9%, 25%, and 30%, respectively, and global recycling rate are only at an average of 18% [5].

Plastic wastes have generated about 6300 million metric tons till the year 2015, 79% are discarded, 12% are incinerated, and only 9% are being recycled. Most plastic wastes end up in landfill or incinerated causing more pollution's while recycled plastics can be used to make a variety of things depending on properties of the recycled plastics [2]. Consumption of plastics will gradually increase and several methods of recycling to recover plastics are required which include mechanical recycling, chemical recycling, pyrolysis, and gasification. The most widely practiced recycling method are mechanical and chemical recycling [5].

One of the mechanical recycling methods is primary recycling also known as re-extrusion or closed loop process [6],[7]. Wood plastic composites made from different mixture, poses different strength, also, good combination of recycled plastics will provide better composition properties [8]. Plastic bags are the most popular item to pollute and destroy natural environment according to [9]. Solvents has been also used to dissolve the polymers in plastic recycling [10].

2. Experimental Setup

Simple recycling techniques for plastic bags are suggested to encourage regular people to choose the recycling industry due to its inexpensive startup and processing costs. The proposed procedures are an implementation of mechanical plastic recycling processes to produce plastic boards using compression molding. The procedure is separated into few stages, identification and separation, sorting, cleaning, shredding, heating and molding. Materials are first identified and separated by type according to recycling code, where HDPE plastic bags are chosen.

The following steps are followed for the recycling:

1. Classification and shredding of HDPE plastics.
2. Heating the shredded HDPE plastics in a normal domestic oven.
3. Decide the heating duration inside the oven for better melting.
4. Pouring the melted HDPE plastics into a prepared metal cast.
5. Applying sufficient pressure on the cast to get the homogeneous shapes.
6. Natural cooling at room temperature.
7. Conducting the mechanical testing on the samples.

HDPE plastic bags are chosen for this research as the raw material. The properties of HDPE are stiffness, strength, chemicals and moisture resistant, permeability to gas, ease to process and ease to form. Some product applications of HDPE are bottles, milk jugs, chemical bottles, dish and laundry detergent bottles, trash, and retail bags.

The equipment and apparatus utilized in this research to conduct the experiments are:

i. Plastics shredder
ii. House use thermal insulated oven. 1380W and 19L capacity. Temperature setting of the oven ranges from 0°C to 250°C.
iii. Baking paper and metal trays
iv. Metal cylindrical mold for compression test samples.
v. Press with G-clamps
vi. Hardness testing machine – Shore A Durometer, Teclock Durometer GS-706G with ASTM D2240 standards.
vii. Compression machine – KENCO E-series Analog Compression Machine.
viii. Microscope – Olympus CX31

In this experiment, the temperature ranges from 210°C to 230°C with different duration inside the oven are implemented to produce the different recycled specimens. Each temperature is conducted with variation of 40, 50 and 60 minutes. Duration of 30 minutes or less are shown unsuccessful for complete melting of the HDPE flakes.

Figure 1 and Figure 2 shown sample specimens of recycled HDPE plastic and compression test, respectively.

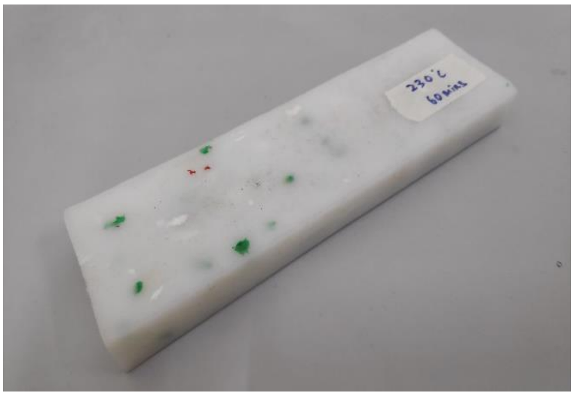


Fig. 1 HDPE plastic specimen

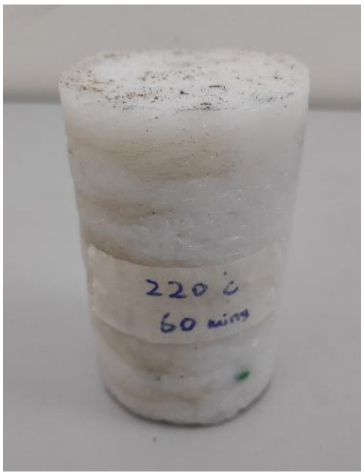


Fig. 2 HDPE plastic compression specimen

3. Result & Analysis

3.1. *Shore Hardness Test*

The hardness test is conducted on specimens according to ASTM D2240 standards, each reading is 12mm from the edges and 6mm in between each point. All data is recorded and tabulated according to time and temperature. Figure 3 is showing the hardness results at the different temperatures and different oven heating duration. It can be observed from the graph that the hardness value is increases with increasing temperature. At every duration, higher temperature results in higher hardness value. Highest hardness value of HDPE specimens at 210℃ is 91.1 while highest hardness value at 230℃ is 95.8. The behaviour of the specimen finally observed that the temperature has an impact on the specimen's hardness, however the highest hardness value was recorded after 60 minutes.

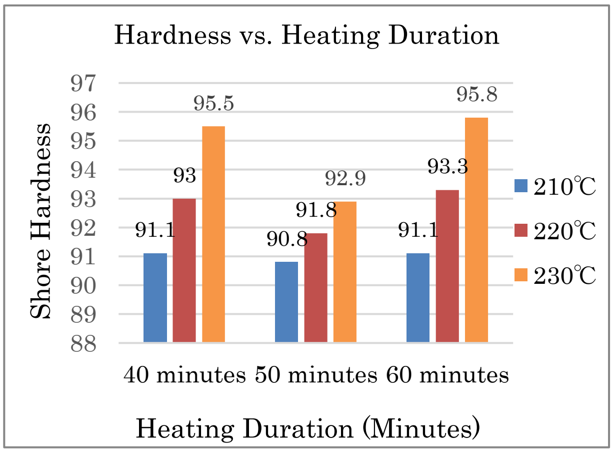


Fig. 3 Hardness comparison at different temperatures and oven heating duration

3.2. *Microscopic Surface Structure*

All specimens from the different recycling conditions have been magnified and observed under a microscope to compare each surface structure. Observation from human naked eye shows smooth surface compared to microscopic magnification which revealed grooves on each surface. Magnification power used to observe for each specimen is 10X. Observation results of surface structure of HDPE plastics when heating at 230^{0}C at different durations are shown from Figure 4 to 6.

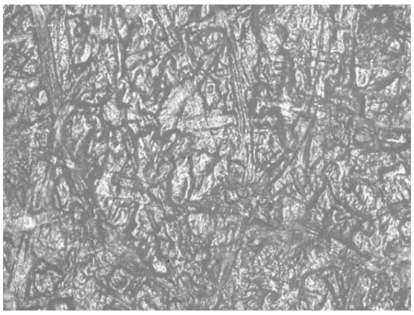


Fig. 4 Specimen at heating duration 40 minutes

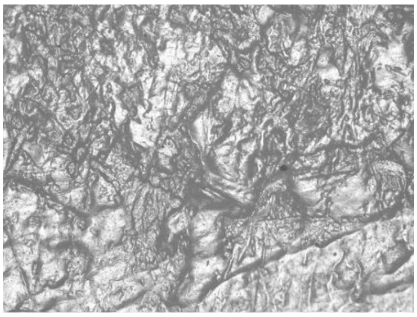


Fig. 5 Specimen at heating duration 50 minutes

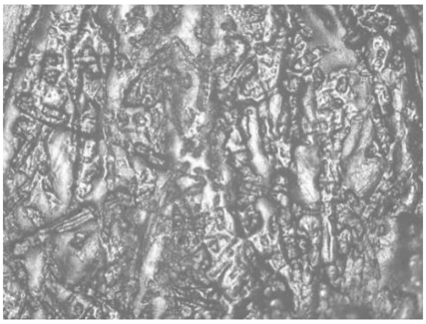


Fig. 6 Specimen at heating duration 60 minutes

3.3. *Compression Test*

Cylindrical specimens have been prepared and tested, using Kenco E-Series Analog compression test machine. After compressing the specimens, HDPE plastics exhibited ductile characteristics as specimens shows barreling from the sides. Cracks occurred on HDPE at certain amount of load and experiment was stopped.

Compressive strength can be calculated with data obtained from compression test. Value obtained from the equipment is Newton, a specific formula will be used to calculate and find the compressive strength value.

$$\text{CS} = \frac{Force\ (N)}{Cross\ sectional\ Area\ of\ Specimen\ (mm)}$$

*CS = Compression Strength

Table 1 Compression Strength of HDPE Plastic

Temperature (°C)	**210**	**220**	**230**
Compressive strength (MPa)	21.04	21.89	20.2

Table 1 shows the values of compressive strength of the recycled HDPE plastics at heating duration 60 minutes. It can be observed that at 220℃, the recycled HDPE is having the strongest compressive strength of 21.89MPa. At higher recycling temperatures, there is no apparent increase in compression strength.

4. Conclusion

In this paper, we have found that the recycling heating temperature and the time duration inside the oven are the dominant factors in the recycling of HDPE plastic bags. A specified temperature should be set for better recycling. Also, the duration inside the oven have shown an effect on the recycled HDPE quality. A heating temperature of 230°C with duration of 60 minutes have shown the best mechanical properties for HDPE recycling. It was noticed that if the recycling temperature was exceeded, the material could burn. Some burnt marks are visible to naked eye, but some can only be revealed through microscopic observation.

When compared to other materials, the recycled HDPE's mechanical qualities show that it can meet all the criteria needed for flat boards to be utilized as table and chair tops. Additionally, it can be utilized to replace the idler rollers and the die pads used for metal forming. Conveyor belts may be supported by idler rolls. Furthermore, recycled HDPE is suggested for the wheels on shopping carts.

Acknowledgment

The authors would like to express the gratitude towards the support from the School of Engineering of UCSI University. Special thanks would be extended to CERVIE office, UCSI University for the endless support to research.

References

[1] Sahajwalla V., and Gaikwad V., 2018. The present and future of e-waste plastics recycling. *Current Opinion in Green and Sustainable Chemistry*, 13, 102-107.

[2] Singh N., Hui D., Singh R., Ajuha I.P.S., Feo L., and Fraternali F., 2017. Recycling of plastic solid waste: A state of art review and future applications. *Composites Part B*, 115, 409-422.

[3] Maris J., Bourdon S., Brossard J.M., Cauret L., Fontaine L., and Montembault V., 2018. Mechanical recycling: Compatibilization of mixed thermoplastic wastes. *Polymer Degradation and Stability*, 147, 245-266.

[4] Zhao Y.B., Lv X.D., and Ni H.G., 2018. Solvent-based separation and recycling of waste plastics: A review. *Chemosphere*, 209, 707-720.

[5] Maris J., Bourdon S., Brossard J.M., Cauret L., Fontaine L., and Montembault V., 2018. Mechanical recycling: Compatibilization of mixed thermoplastic wastes. *Polymer Degradation and Stability*, 147, 245-266.

[6] Ragaert K., Delva L., and Geem K.V., 2017. Mechanical and chemical recycling of solid plastic waste. *Waste Management*, 69, 24-58.

[7] Yuyang S., Umesh G., Takeshi S., Uday K.V., Srikar V., Anthony Y., and Tim O., 2018. CAE method for compression molding of carbon fiber-reinforced thermoplastic composite using bulk materials. *Composites Part A: Applied Science and Manufacturing*, 114, 388-397.

[8] Najafi S.K., 2013. Use of recycled plastics in wood plastic composites – A review. *Waste Management*, 33 (9), 1898-1905.

[9] When The Mermaids Cry: The Great Plastic Tide by Lytle. C (2017, January). Available from <http://plastic-pollution.org/> [Accessed 22 July 2017].

[10] Oblak, P., Gonzalez-Gutierrez, J., Zupancic, B., Emri, I., 2016. Mechanical properties of extensively recycled high density polyethylene (HDPE). Slovenia: *Materials Today: Proceedings*, 3, 1097-1102.

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 PhD 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.

Wan Wai Kit

Wan Wai Kit was graduated Bachelor's Degree in Mechanical Engineering at UCSI University, Kuala Lumpur in 2019. His field of research is recycled plastics material.

Tan Joon Hoe

Tan Joon Hoe was graduated Bachelor's Degree in Mechanical Engineering at UCSI University, Kuala Lumpur in 2019. His field of research is recycled plastics material.

Mr. 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.

Ms. Nor Fazilah Abdullah

She received her Bachelor's degree in aerospace engineering (Hons) from IIUM, Gombak and her Master's degree in Mechanical Engineering from UKM, Bangi. Currently she is pursuing Doctoral of Philosophy programme and also a lecturer from Department of Mechanical & Mechatronics Engineering at UCSI University, Kuala Lumpur. Her research interest in bio-based nanoparticles materials.

Design and Fabrication of Power Generating Treadmill

Ammar A. M. Al-Talib, They Kai Yang, Noor Idayu M.Tahir
Department of Mechanical and Mechatronics, Faculty of Engineering,
Technology and Built Environment, UCSI University, 56000 Kuala Lumpur, Malaysia

Sarah ‘Atifah Saruchi
Faculty of Mechanical and Mechatronics, University Malaysia Pahang, 26600 Pekan,
Pahang, Malaysia
E-mail: ammart@ucsiuniversity.edu.my, 10011748536@ucsiuniversity.edu.my, NoorIdayu@ucsiuniversity.edu.my,
saratifah@yahoo.com

Abstract

This paper aims to take advantage of treadmill’s wasted electrical energy during a person’s workout and utilize the energy for charging electrical appliances. The energy expended on a treadmill during the exercise is all wasted. In order to take advantage of the wasted energy, it could be harnessed by a power generator and stored in a battery bank. The electrical energy generated during the exercise on the treadmill could be utilized to power electronic devices and appliances. The attached power generation machine will not interrupt a person’s workout flow and it can be attached to any treadmill due to its friendly design. Wasted energy is harnessed in this research by a non-traditional manner of using shaft and wheel method. A multi-meter is used to measure the voltage and current and power is then calculated from the readings recorded. Tests have shown that the prototype machine is able to fully charge a 3096 mAh smartphone in 135 minutes and the phone could be fully charged for 2 charging cycles. This power generating machine is showing a good implementation for the Sustainable development Goals (SDG’s).

Keywords: Energy harvesting; Power generator; Treadmill; Using wasted energy

1. Introduction

The Every day, there is a noticeable rise in energy demand. Traditional energy sources, such as fossil fuels, are having long harmful influence on the environment. A theoretical model of an Energy Generating Gymnasiums System (EGGS) that can reuse and transform wasted energy has been presented [1]. The purpose of this power generating treadmill is to utilize an energy harvesting system that moves in response to the motion of a treadmill in order to convert the treadmill's kinetic energy into electrical power [2].

Energy harvesting works by capturing little amounts of energy that would otherwise be lost or squandered as heat, vibration, or light. Renewable energy production is becoming one of the best strategies for energy production. Renewable energy production accounts for an estimated two-thirds of total worldwide power generation and is predicted to more than double by 2040. This increase is

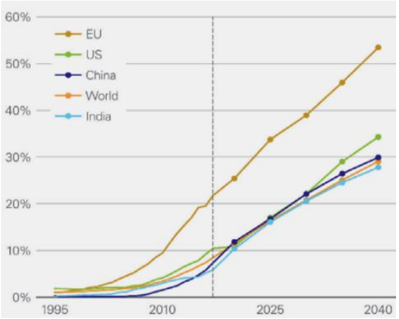


Fig. 1 Glimpse of energy future [3]

regarded seriously as a target to cover at least 30% of total utilized energy [3]. Figure 1 is showing the glimpse of energy future.

Three types of harvesting system are considered for this paper.

Piezo electricity is the charge that builds in certain solid materials (such as crystals, ceramics, and biological matter like as bone, DNA, and different proteins) when mechanical stress is applied [4].

The effect occurs when a ferro-electric material is subjected to a direct application of any kind of mechanical force or when a mechanical force is generated in response to the application of an electric field. Ferro-electric materials' molecular structure displays a local charge separation known as an electric dipole. When an electric potential is applied, the lead zirconate titanate crystal vibrates, producing an indirect effect. PZT is composed of two types of materials: soft and hard [5]. The piezoelectric effect takes place whenever a ferro-electric material is subjected to a direct application of any kind of mechanical force. Piezoelectric materials are a subset of the larger family of ferro-electric materials.

Mechanical energy conversion devices, such as electromagnetic harvesters, are gradually replacing battery-powered electronic devices [6]. The conversion of the mechanical energy that is being harvested into electrical energy is accomplished by a method known as electromagnetic energy harvesting. According to Faraday's law, the amount of electromotive force that is produced by a circuit is directly proportional to the time that passes before there is a change in the magnetic flux linkage that is contained inside the circuit.

Electrostatic converters are capacitive devices that transmit energy when the plates of a variable capacitor separate or the area of the plate changes in response to external mechanical energy [7]. The variable capacitor is the heart of the electrostatic energy harvester; it exploits a change in capacitance to generate a voltage or charge rise in the harvesting device. They are classified into two general categories: continuous charge and constant voltage [8].

Electrostatic harvesting has been found to offer the least amount of complexity, lowest energy density, and smallest integrated current size compared to other energy harvesting techniques [9]. Because of this, it is the most applicable to real-world situations and offers the best potential for long-term power production.

2. Methodology and Experimental Setup

Electromagnetic energy harvesting has been chosen as the energy collecting principle. The treadmill machine generates power by translating the horizontal motion generated by gym exercise to rotational motion [10]. A technical drawing of a redesigned conceptual design for a power generating treadmill machine is proposed in this paper. The final modified design is more comprehensive and feasible in terms of achieving all the objective goals. SOLIDWORKS software has been used to create the detail drawings for the final design in Figure 2.

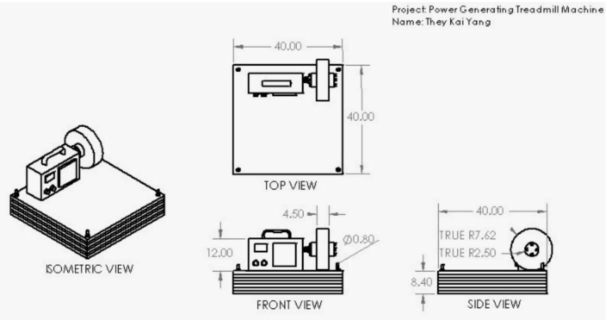


Fig. 2 Isometric and orthographic of the prototype

A functional prototype of the power generating treadmill machine has been fabricated. The generator will be mounted on a base, with a wheel coupled to it through a shaft. The coupling is transferring the rotational motion from the wheel to the generator shaft, which will generate the electrical energy. wooden planks have been chosen to fit the weight requirement and height of the power generator.

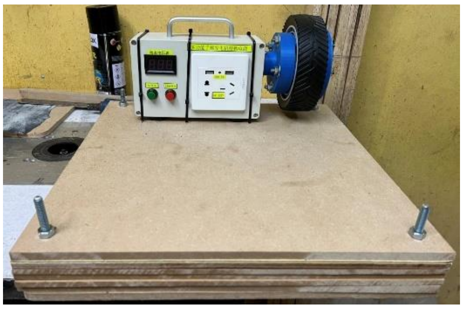


Fig. 3 Completed prototype of power generating treadmill machine

Figure 3 shows the completed prototype of the power generating treadmill machine. The experiments have conducted at three different speeds (5km/h, 7km/h, 9km/h) that an average human is comfortable for walking, jogging and running on.

3. Result and Discussion

Tests are conducted on the designed treadmill generator at three different speeds that an average human is comfortable running on. No load is used in the first set of experiments, and the three comfortable speeds used are shown in Table 1.

SPEED 1	**SPEED 2**	**SPEED 3**
5km/h	7km/h	9km/h
Slow pace	Medium pace	Fast pace

Table 1 Experimental Criteria

The result of the experiments from speed 1 to speed 3 has been recorded and presented in Figure 4.

3.1 Result of Voltage Generated for Three Speeds

Figure 4 is showing the voltage generated. The readings of the multi-meter have been taken after the treadmill was ran for 5 minutes to ensure stability. The voltage readings for 5km/h, 7km/h and 9km/h have been repeated three times and averaged out to reduce inaccuracy in the final value of voltage.

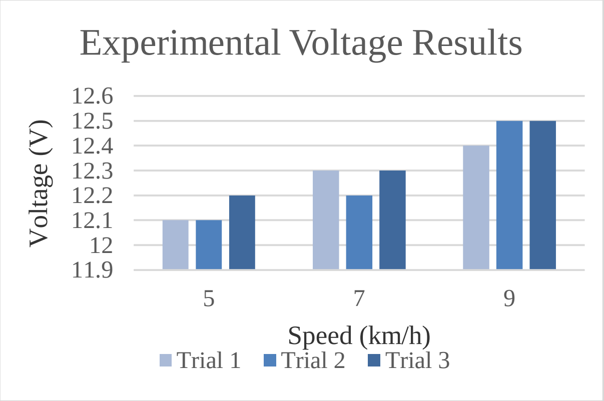


Fig. 4 Graph of experimental voltage results

3.2 Results of Current Generated for There Speeds

Figure 5 shows the experimental current generated by the power generator. The current readings for 5km/h, 7km/h and 9km/h are repeated three times and the results were averaged out to reduce inaccuracy in the final value of current. Figure 5 is showing the overall results of the experimental current generated.

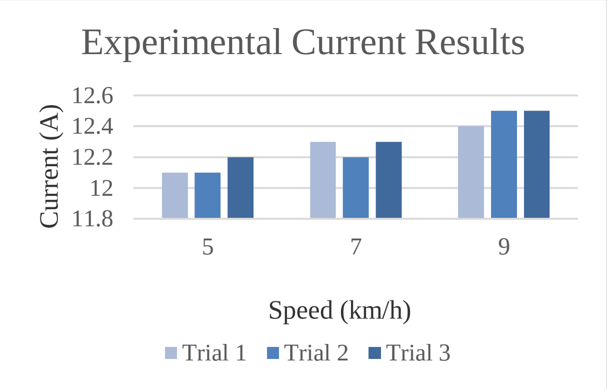


Fig.5 Graph of experimental current results

3.3 Results of Power Generated for Three Speeds

To obtain the value for power generated for the three speeds, the average value from the voltage and current output are used. The value of power in Watts are obtained by using the equation below.

$$P = I \times V$$

Figure 6, shows the results by calculating the average voltage, current and power generated at the three different speeds.

From Figure 6, it can be observed that the power generated from the power generating treadmill machine increases as the speed of the treadmill increases. From this observation, it can be concluded that power generation is directly proportional to the speed of the treadmill. Figure 6 shows the relation of power generation and speed of the treadmill.

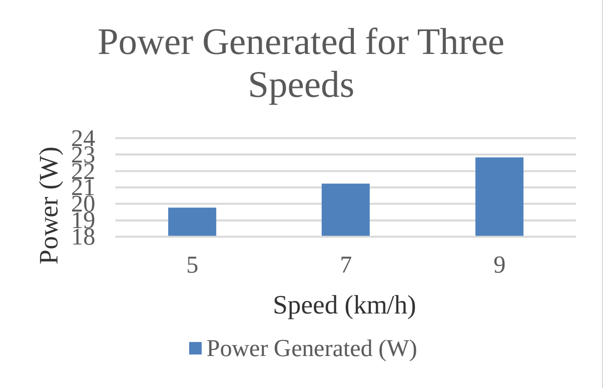


Fig. 6 Graph of power generated for the three speeds

3.4 Application of the power generating treadmill machine

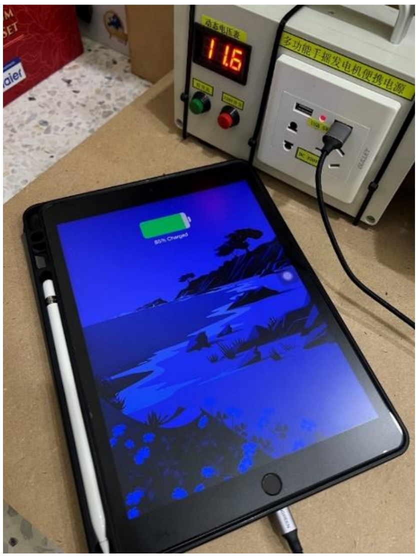


Fig.7 Prototype charging a smart device

For the application of the power generating treadmill machine, the machine is charged fully by the treadmill. A smart device is then connected to the machine via a USB cable using the 5V USB port. Figure 7 shows the machine charging a smart device.

Time (minutes)	Battery Status of Smart Device (%)
0	0
15	15
30	31
45	43
60	55
75	68
90	80
105	90
120	95
135	100

Table 2 Table taken to fully charge a smartphone

According to the results of the conducted experiments, a smart device with a battery capacity of 3095mAh may be fully charged in 135 minutes using a 5V USB port in Table 2. When compared to a standard wall socket power point, the speed is regarded to be slower. The power generated by the prototype is a green energy that harvests wasted mechanical energy from the treadmill.

The lithium batteries utilized in the prototype have a capacity rating of 6600mAh. It is theoretically possible to charge a smart phone with a 3095mAh battery capacity to full battery capacity for 2.13 cycles.

4. Conclusion

A working prototype has been fabricated for the treadmill power generator to demonstrate the concept. Using wheel and shaft mechanism have reduced the intricacy of building the prototype. By incorporating the power generating treadmill machine into the gym, the user may charge their phone or power up any home appliances using waste energy conversion. Experiments at different speeds have shown that the proposed system can charge a smartphone of 3069mAh from 0% to 100% in 135 minutes, which is still within an acceptable time range. It delivers affordable and cost-effective energy to power electric appliances.. Aside from that, it does not affect the environment due to its zero-fuel input and zero greenhouse gas emissions which shows a good implementation of the SDG's.

Acknowledgment

The authors would like to express the gratitude towards the school of engineering at UCSI University for the support throughout the research. Special thanks would be extended to CERVIE office at UCSI University for the endless support to research.

References

[1] Musharraf, M., Iqbal, F., & Saleem, I. (2018). Energy Generating Gymnasiums Machines for Renewable, Sustainable and Green Energy. International Research Journal of Engineering and Technology.

[2] Shingare, P. v, Somvanshi, V. B., Tore, T. P., & Sonawane, V. K. (n.d.). Gym Power Generation Mechanism (Vol. 4).

[3] Yahya, K., Salem, M., Iqteit, N., & Ahmad Khan, S. (2020). A Thermoelectric Energy Harvesting System. In

Renewable Energy - Resources, Challenges and Applications. Intech Open.
[4] Ravi Shekhar, & Prakash Kumar. (2019). Design and Analysis of Piezoelectric Energy Harvesting Systems.
[5] Adhithan, M. A., Vignesh, K., & Manikandan, M. (2015). Proposed Method of Foot Step Power Generation Using Piezo Electric Sensor. International Advanced Research Journal in Science, Engineering and Technology.
[6] Georgiev, N. P., & Raychev, R. P. (2020). Study of four-spring electromagnetic harvesters. IOP Conference Series: Materials Science and Engineering, 878(1).
[7] Aljadiri, R. T., Taha, L. Y., & Ivey, P. (2017). Electrostatic Energy Harvesting Systems: A Better Understanding of Their Sustainability Electrostatic Energy Harvesting Systems: A Better Understanding of Their Sustainability. Journal of Clean Energy Technologies, 5(5), 409–416.
[8] Elliott, A. D. T., Miller, L. M., Halvorsen, E., Wright, P. K., & Mitcheson, P. D. (2015). Which is better, electrostatic or piezoelectric energy harvesting systems? Journal of Physics: Conference Series, 660(1).
[9] Marzencki, M. (2005). Vibration energy scavenging. European Commission research Project VIBES (IST-1-507911) of the 6th STREP Framework Program.
[10] Kumar, M., & Mundada, G. S. (2007). Energy Harvesting from Gym Equipments. International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering ISO, 3297, 2321–5526.

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.

They Kai Yang

They Kai Yang received his Bachelor's Degree in Mechanical Engineering in 2021 from UCSI University, Kuala Lumpur. His field of research is renewable energy.

Sarah ‘Atifah Saruchi

Sarah ‘Atifah Saruchi 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 works as a lecturer under the Department of Mechanical and Mechatronics, Faculty of Technology and Built Environment, UCSI Kuala Lumpur. Her research interests include control, vehicle and artificial intelligence.

Noor Idayu Mohd Tahir

Noor Idayu Mohd Tahir 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.

Japanese Self-Directed Learning System with YouTube Requires Meta-knowledge of Collocation

Hiroki Fxyma
Tainan University of Technology
No.529, Zhongzheng Rd., Yongkang District, Tainan City 710302, Taiwan (R.O.C.)

Abstract

This study proposed a self-directed learning system that extracts subtitle information from YouTube videos and presents it as learning vocabulary. In recent years, Japanese language learners' interests have become more diverse, with easy access to multilingual information through SNS. Thus, it is assumed that there are many potential learners not enrolled in Japanese language institutes or less-than-learners. This study proposes a system for learners to learn Japanese from desired YouTube videos. The basic concept of the system is to extract subtitle information from 10 videos and present high-frequency words. The results of frequency analysis and co-occurrence analysis showed the feasibility of this system. At the same time, it was suggested that the characteristic words of each video should be presented together with their collocations and that meta-knowledge of their usages is required.

Keywords: Japanese Language Education, YouTube-based learning, Less-than-learners, Language Learning

1. Introduction

Motivations for language learning are becoming more diverse as contact with multilingual information through SNS is becoming more common. Therefore, it is expected that the number of potential 'less-than-learners,' who are not explicit learners in the classroom, is increasing.

In this study, I propose a system that allows such potential learners to watch videos of their own interest, focusing on the words they need to know. With this background, several studies have been proposed to incorporate YouTube into language education (e.g., [1][2][3]) In this study, students will create a system that allows them to watch their favorite Japanese-language YouTube videos and learn the vocabulary and grammar used in those videos.

2. Method

2.1. *Self-learning system with YouTube*

In this project, I propose a system that learners (including less-than-learners) to watch their favorite Japanese-language YouTube videos and learn the vocabulary and grammar used in those videos. The configuration of the system is shown in Figure 1.

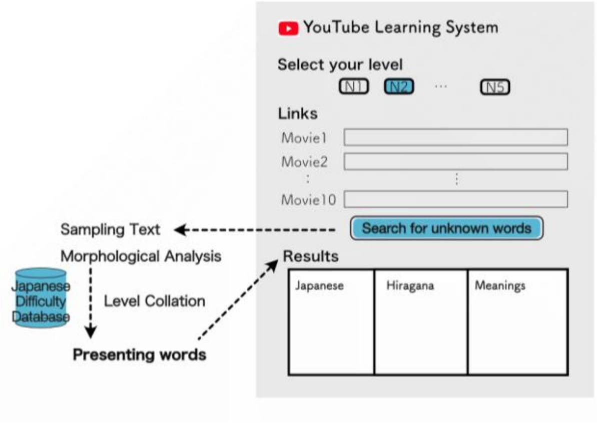


Figure 1 The outline of the self-learning system with YouTube (web application).

Outlines of the system

This system is planned as a web application. Learners will select their Japanese level and input links of videos they want to watch. In order to obtain a sufficient word size for the co-occurrence analysis, it is desirable that the links be entered for approximately 10 videos of 10 minutes in length. If the video length is for several hours, such as a video of a live game, a single video link would be sufficient in size.

How to use the self-study system

1) Visit the system website
2) Select your Japanese language level
3) Enter 5-10 links to the videos you want to watch
4) Press the [Search] button
5) Learn the displayed list of words and grammar list
6) Watch the video you want to watch

2.2. *Video Materials*

In this study, cooking-related YouTube videos will be used as the material for this study, with a view toward educational practice in the food and beverage department where the author belongs. Cooking-related YouTube videos include various genres such as recipe, cooking, mysterious dishes, eating huge volumes, and videos in which the eating sound is played. Of these, this study will use the "Kimagure-Cook" channel, which has the largest number of subscribers among Japanese YouTube channels, as its material.

Kimagure-Cook consists mainly of videos showing one man dismantling and cooking fish and other marine products. In the fish cooking process, terms that are not used in general conversation, such as parts of the fish, cooking utensils, and verbs for cooking, appear. These words are not included in general language instructional materials, or they are considered advanced content. The 10 videos used in this study are listed in the reference.

As language materials, 10 videos were compiled into a corpus, detailed as Table 1. Note that the corpus is analyzed unformatted, assuming that learners use arbitrary videos. (i.e., errors in speech analysis, unknown words, etc., were not removed or corrected).

Table 2

Tokens	18,502
Types	2,800
Sentences	1,952
Paragraphs	1,845
H1 (Movies)	10

3. Analysis and brief results

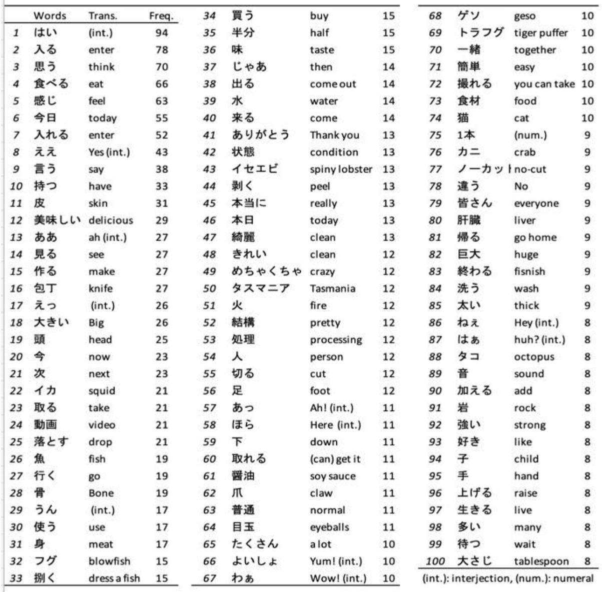
The output of this system is the presentation of learning vocabulary according to the learner's level. For this goal, the system analyzes the words necessary to understand the video in Japanese (the original language).

There are two main types of analysis. One is the frequency analysis of words to extract high-frequency words, and the other is the characteristic words in each of the 10 videos based on co-occurrence relations.

3.1. *High Frequency Words*

Of the words that appeared in the 10 videos, the high-frequency words (set as frequency 8 or higher) were as shown in Table 2. Note that following analysis focuses on verbs and nouns.

Table 1



	Words	Trans.	Freq.
1	はい	(int.)	94
2	入る	enter	78
3	思う	think	70
4	食べる	eat	66
5	感じ	feel	63
6	今日	today	55
7	入れる	enter	52
8	ええ	Yes (int.)	43
9	言う	say	38
10	持つ	have	33
11	皮	skin	31
12	美味しい	delicious	29
13	ああ	ah (int.)	27
14	見る	see	27
15	作る	make	27
16	包丁	knife	27
17	えっ	(int.)	26
18	大きい	Big	26
19	頭	head	25
20	今	now	23
21	次	next	23
22	イカ	squid	21
23	取る	take	21
24	動画	video	21
25	落とす	drop	21
26	魚	fish	19
27	行く	go	19
28	骨	Bone	19
29	うん	(int.)	17
30	使う	use	17
31	身	meat	17
32	フグ	blowfish	15
33	捌く	dress a fish	15
34	買う	buy	15
35	半分	half	15
36	味	taste	15
37	じゃあ	then	14
38	出る	come out	14
39	水	water	14
40	来る	come	14
41	ありがとう	Thank you	13
42	状態	condition	13
43	イセエビ	spiny lobster	13
44	剥く	peel	13
45	本当に	really	13
46	本日	today	13
47	綺麗	clean	13
48	きれい	clean	12
49	めちゃくちゃ	crazy	12
50	タスマニア	Tasmania	12
51	火	fire	12
52	結構	pretty	12
53	処理	processing	12
54	人	person	12
55	切る	cut	12
56	足	foot	12
57	あっ	Ah! (int.)	11
58	ほら	Here (int.)	11
59	下	down	11
60	取れる	(can) get it	11
61	醤油	soy sauce	11
62	爪	claw	11
63	普通	normal	11
64	目玉	eyeballs	11
65	たくさん	a lot	10
66	よいしょ	Yum! (int.)	10
67	わぁ	Wow! (int.)	10
68	ゲソ	geso	10
69	トラフグ	tiger puffer	10
70	一緒	together	10
71	簡単	easy	10
72	撮れる	you can take	10
73	食材	food	10
74	猫	cat	10
75	1本	(num.)	9
76	カニ	crab	9
77	ノーカット	no-cut	9
78	違う	No	9
79	皆さん	everyone	9
80	肝臓	liver	9
81	帰る	go home	9
82	巨大	huge	9
83	終わる	fisnish	9
84	洗う	wash	9
85	太い	thick	9
86	ねぇ	Hey (int.)	8
87	はぁ	huh? (int.)	8
88	タコ	octopus	8
89	音	sound	8
90	加える	add	8
91	岩	rock	8
92	強い	strong	8
93	好き	like	8
94	子	child	8
95	手	hand	8
96	上げる	raise	8
97	生きる	live	8
98	多い	many	8
99	待つ	wait	8
100	大さじ	tablespoon	8

(int.): interjection, (num.): numeral

3.2. *Common words and characteristic words*

These high-frequency words can be categorized into common words that appear in all videos and characteristic words that are unique to each video. To identify the tendency of the words-appearance, the nouns and verbs in the corpus are plotted by correspondence analysis with each video, as shown in figure 2.

Figure 2 shows that the Tasmanian crab and squid videos contain a number of characteristic words. In the Tasmanian crab video, words such as Tasmania, crab,

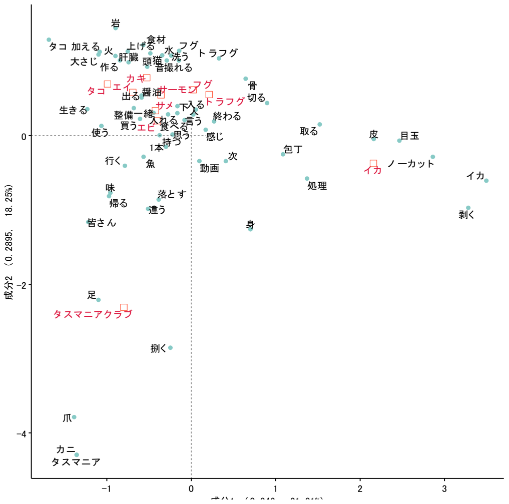


Figure 2. Correspondence analysis of nouns and verbs (frequency over 8) appeared in 10 videos.

claw, and legs seem characteristic words. In the squid videos, words such as squid, uncut, eyeball, skin, and processing are characteristic words. The feature terms, based on co-occurrence relationships (by *Jaccard index*) within each video are shown in Table 3.

Table 3. Characteristic words of each video

エビ		エイ		フグ		イカ	
整備	.069	思う	.034	水	.060	イカ	.054
今日	.054	照り焼き	.028	フグ	.054	皮	.053
思う	.046	大さじ	.027	皮	.042	剥く	.034
入る	.035	肝臓	.027	食べる	.040	取る	.028
大きい	.034	今日	.026	思う	.039	包丁	.027
頭	.028	行く	.025	入る	.038	半分	.025
持つ	.027	作る	.024	悪い	.035	ノーカッ	.020
言う	.026	見る	.024	言う	.028	目玉	.020
固まる	.025	簡単	.020	毒	.028	耳	.017
お願い	.025	猫	.020	多い	.027	身	.017
カキ		サーモン		サメ		タコ	
言う	.033	入れる	.048	魚	.046	タコ	.050
岩	.032	感じ	.047	フライ	.042	入れる	.050
今日	.030	思う	.037	今日	.042	作る	.044
バーベキ	.028	入る	.033	フカヒレ	.034	強い	.042
きれい	.022	今	.027	ソース	.033	使う	.033
野菜	.018	サーモン	.024	持つ	.027	食べる	.031
焼く	.018	上げる	.024	思う	.027	刻む	.029
洗う	.018	動画	.022	マヨネー	.025	頭	.025
買う	.018	作る	.022	買う	.023	美味しい	.024
今	.017	加える	.019	作る	.021	元気	.021
タスマニアクラブ		トラフグ					
タスマニ	.054	入る	.064				
爪	.053	骨	.054				
捌く	.052	感じ	.053				
カニ	.043	食べる	.048				
思う	.041	トラフグ	.045				
足	.037	頭	.042				
落とす	.035	入れる	.041				
美味しい	.034	腎臓	.034				
食べる	.033	思う	.034				
感じ	.029	撮れる	.033				

3.3. *Collocation*

Although presenting only frequent or feature words may have a certain educational effect, presenting collocations may improve the qualitative aspect of word knowledge. In this section, we propose a method of presenting feature words together with collocations, using the vocabulary of the Tasmanian crab and squid videos, a unique set of 10 videos, as an example.

In this section, collocations of characteristic words in squid-video and Tasmanian club video are analyzed by listing words with two or more co-occurrences as major collocations.

Collocation in Tasmanian club video

Take a look at the collocation of feature words in the Tasmanian club video. The first characteristic word, "Tasmania," is a proper noun and does not need to be analyzed by collocation.

The main collocations of "爪 claw" in the second place were '足 legs' and '落とす drop'. The '足 leg' was used in the literal sense of "crab legs". 'Drop(ping)' was used as a transitive verb in the sense of *cutting off crab legs, claws, or fish fins*. In understanding the word "claw," one may be required to have the meta-knowledge that in Japanese, both crab claws and human fingernails are represented by the same word, 爪 */tsume/*. In derivation, the knowledge that both foot and leg can be applied to the Japanese word '足 *ashi*' may also be required.

The main collocations for '捌く/sabaku/ *dress* or *process (a fish)*' were '思う think' and '魚 fish'. The usage of '思う think' was "捌けると思う I think I can dress it" or "捌いていきたいと思う I would like to process it," thus no particularity as a collocation was observed. The co-occurrence with '魚' is the expression '魚を捌く *dress a fish*.' In Japanese, '捌く' is used to describe the chopping up of fish or chicken.

"落とす *drop*" is the transitive form of '落ちる *drop*.' It co-occurs with '頭 *head*,' '関節 *joint*,' '足 *leg*,' and '爪 *claw*' as objects. When processing crabs, '落とす *drop*' is used to describe the cutting off of the legs and other parts of the crab.

Collocation in Squid video

The collocations for "皮 peel [noun]" are '剥く /muku/ peel [verb]', '状態 state', '柚子 yuzu', '引く pull', '感じ feel', '止める stop', '入れる insert', and '剥ける

remove’. The regular expression to *remove the skin* is the phrase “皮を剥ぐ /*kawa o hagu*/ *peel off*” or “皮を剥く /kawa o muku/ *peel off*”. A similar but less common expression, “皮を引く/*kawa o hiku*/ *peel off the skin*,” is used when *removing the thin skin of yellowtail, squid, and other fish*. Derivatively, there is also the usage of “湯引き /*yubiki*/ *hot water-peeling*,” for tomatoes, peaches, and so on.

In Japanese, not only human skin, but also the skin of fish, the peel of fruits such as yuzu, and the bark of trees are described as ‘皮 /*kawa*/ *skin*’. Animal skin is also ‘皮,’ and tanned skins are also pronounced ‘/*kawa*/ (spelled as 革)’. Such metaknowledge is essential to understand the use of the word ‘skin’ as well as ‘*nails / craws* for 爪’ in the Tasmanian Club.

The two highest co-occurring words (with at least two co-occurrences) for “剥く /*muku*/ peel” were ‘皮’ and ‘風’. Of these, ‘風’ was used in all cases to be pronounced /fu/ (means *way*, not as /kaze/ for *wind*), which means “こんな風に剥いていきます. *I will peel in this way*.”

4. Summary & Discussion

In this presentation, we propose a self-directed Japanese language learning system that extracts subtitle information from YouTube videos and presents them as learning vocabulary.

As a method of presenting learning vocabulary, I proposed a strategy to present collocation-aware information in addition to the method of creating a corpus from videos and presenting frequently appearing words.

Word usage in YouTube videos has a special frequency structure depending on the genres of the videos. The analysis suggests that while the frequent words can be simply presented, the characteristic words of each video need to be presented together with collocations and meta-knowledge on the word usage.

Future research issues are the presentation of the vocabulary according to the learner's level, the systematization of meta-knowledge for presenting collocations, and the implementation of the system.

References

1. Albantani, Azkia Muharom, and Ahmad Madkur. 2017. “Musyahadat al Fidyu: YouTube-Based Teaching and Learning of Arabic as Foreign Language (AFL).” *Dinamika Ilmu*, 291–308.
2. Aprianto, Dedi. 2020. “To What Extent Does YouTube Contents-Based Language Learning Promote an English Proficiency?” *Journal of English Language Teaching and Literature (JELTL)* 3 (2): 108–26.
3. Riswandi, Diki. 2016. “Use of YouTube-Based Videos to Improve Students’ Speaking Skill.” In *Proceeding of the International Conference on Teacher Training and Education*, 2:298–306.

YouTube Movies
[All by “Kimagure-cook”, Titles modified]
1. 1匹17万円！世界最大のカニ『タスマニアキングクラブ』をさばいて食べてみた！, 2019. https://www.youtube.com/watch?v=DpK6vo4_sX4.
2. 巨大イカのさばきかた, 2019. https://www.youtube.com/watch?v=qmuxw4iZvjU.
3. 【一人BBQ】市場で見つけた ばけものサイズの『岩牡蠣』炭火と生牡蠣で食いまくる！, 2020. https://www.youtube.com/watch?v=D4Qz0Vscm0o.
4. 元気過ぎて台所をかけまわるマダコをしめて。ぶったぎりにしてワサビで漬け込む料理。, 2020. https://www.youtube.com/watch?v=I_HX9dW49BU.
5. 全身『水』のフグのお腹の中身が･･･。こんなん食べてるの！？, 2019. https://www.youtube.com/watch?v=_nvAfEov9pY.
6. 巨大トラフグさばいてみた!, 2018. https://www.youtube.com/watch?v=zRbV08ykKTM.
7. 水死したエイの腹の中を綺麗に掃除してさばいて料理してみた, 2018. https://www.youtube.com/watch?v=gIYbo3nN72E.
8. 育ちすぎてしまった巨大な危険生物ニシキエビ。すべてが規格外。生きたままさばいて食べた, 2019. https://www.youtube.com/watch?v=Rx-xTpxeOUQ.
9. 200キロのサメをさばいたらお腹の中がすごかった, 2018. https://www.youtube.com/watch?v=W-y-rONjNTc.
10. 【衝撃映像】キングサーモンのお腹の中身がイクラまみれだった, 2018. https://www.youtube.com/watch?v=HQEaFT7Q0nk.

Authors Introduction

Hiroki Fxyma (Ph.D)

H. Fxyma received Doctoral degree from Keio University, Japan in 2018. His research domain includes application of Cognitive Science to Japanese Language Education.

A Study on the Impact of "Flaming" on Content from "A Crocodile Who Will Die in 100 Days"

Taiki Sugimoto
Faculty of Global Management, Chuo University, Higashinakano Hachioji-shi, 192-0393/Tokyo, Japan

Jun Nakamura
Faculty of Global Management, Chuo University, Higashinakano Hachioji-shi, 192-0393/Tokyo, Japan

E-mail: a19.3gem@g.chuo-u.ac.jp, jyulis.77f@g.chuo-u.ac.jp
https://www.chuo-u.ac.jp

Abstract

This study is aimed at capturing the impact of flaming generated through social media on content by Sentiment Analysis and Key Graph Analysis. The analysis and discussion clarify that content that has been under flaming will not stop being criticized for a long period of time, and that being flaming experience itself will be consumed as content.

Keywords: Sentiment Analysis, Key Graph Analysis, ML-Ask, SNS, Flaming.

Introduction

This paper uses the comic strip "A Crocodile Who Will Die in 100 Days," which was posted mainly on Twitter[i], to study the prevalence of flaming content on social networking sites and the associated emotional movement. The comic strip "A Crocodile Who Will Die in 100 Days" was updated daily by comic artist Yuki Kikuchi on Twitter for 100 days, from December 12, 2019, to March 20, 2020.

"A Crocodile Who Will Die in 100 Days" was promoted for commercial deployment shortly after it recorded the highest number of "likes" on Twitter in Japan's history. Therefore, both the content and the product were widely recognized, and it can be said to have succeeded in attracting the attention of the AISAS model[ii] of Internet consumption behavior. However, the trending that briefly attracted favorable attention from an unspecified number of people turned overnight and got under the flame, in which attacks and criticisms flooded in from an unspecified number of people in a short period because of the promoted. In this respect, it is worth noting that it is rare for content that was both trending and flamed not to lead to successful sales.

Previous Research

This thesis focuses on trending and flaming on social networking services (SNS). This chapter describes the positioning of this thesis regarding previous research on flaming on the Internet in an SNS.

---

[i] A social networking service launched by Obvious (Twitter, Inc.) in July 2006 and used by 217 million users per day as of April 2021 [1].

[ii] Proposed by Dentsu in 2004. The acronym stands for Attention, Interest, Search, Action, Share, and is used to understand consumers' attitudes and behavior in order to develop marketing strategies appropriate for each stage.

Concerning flaming, defined flaming as "a requirement that a large amount of criticism is written on CGMs such as social media in a short period of time and that the criticism spread to multiple online services rather than only to a single online service" [2], "A Crocodile Who Will Die in 100 Days" which meets the above definition because the criticism spread across various media, including Twitter, in a short period after the promotion for commercial deployment.

The previous research also addressed actual flame cases [2], this paper, similarly, addresses a real flame case. Approaches to flames have been attempted by categorizing Tweets and aggregating them in time series. This paper differs from the previous study in that it focuses on a single case of "A Crocodile Who Will Die in 100 Days" and uses Sentiment Analysis and Key Graph Analysis to study the flames. The research on the presence of others on SNS has indicated the desire to examine the factors that cause flaming by content and to elucidate the processes that lead to flaming [3]. Since this paper quantitatively analyzes and examines the flaming of content, this paper is academically and socially significant research that also contributes to the clarification of the process.

Purpose

The purpose of this paper is to clarify the impact of flaming on content. This paper shows how flaming have affected peoples' emotions.

Methods

This chapter describes the data, sentiment, and Key Graph Analysis used in this thesis.

This paper collected a total of 172,562 Tweets which including either or both the official name of the subject, "*100 日後に死ぬワニ*," the abbreviation of the subject, "*100 ワニ*," and the author of the subject, "*きくちゆうき*." Tweet data were collected from two periods, based on the events and Google trends. before the film's release (April 1, 2020) to (June 30, 2021) and From around the start date of the film release (July 1, 2021) to around the end date of the film release (August 8, 2021). Sentiment Analysis and Key Graph Analysis are conducted on the Tweets.

Sentiment analysis

Sentiment Analysis analyzes the Tweet data for each time series, and quantitatively quantifies the emotional fluctuations seen in the user's transmissions.

Sentiment Analysis using ML-Ask is a Sentiment Analysis method that focuses on a wider range of emotions. Emotions are classified into 10 categories using the Emotional Expression Dictionary: “sadness”, “shame”, “anger”, “dislike”, “fear”, “surprise”, “like”, “excitement”, “peace”, and “joy” [4].

Key Graph analysis

Key Graph Analysis analyzes the same Tweet data for the same period as the Sentiment Analysis, to objectively visualize the information and impressions that users were transmitting.

For the analysis, the Dice coefficients based on equation (1).

$$DSC(A,B) = \frac{2|A \cap B|}{|A|+|B|} \qquad (1)$$

Results

Sentiment analysis results

Using the collected data, Sentiment Analysis was conducted for each period, and the percentage of each sentiment was calculated for each period. This was done to make it easier to compare the transition, since the number of Tweets differed from period to period. Fig.1 below show a graphical summary of the percentage transition of each estimated emotion.

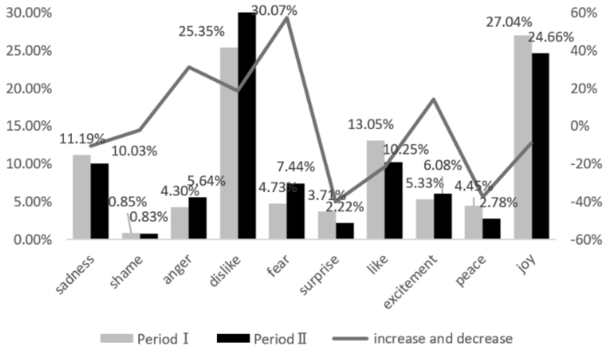


Fig. 1: Sentiment Analysis (April 1, 2020 - June 30, 2021, and July 1, 2021 - August 8, 2021)

Key Graph analysis results

In this chapter, the clusters that can be read from the Key Graph Analysis results are indicated by numbers in each figure to clarify what kind of context was present in the Tweets.

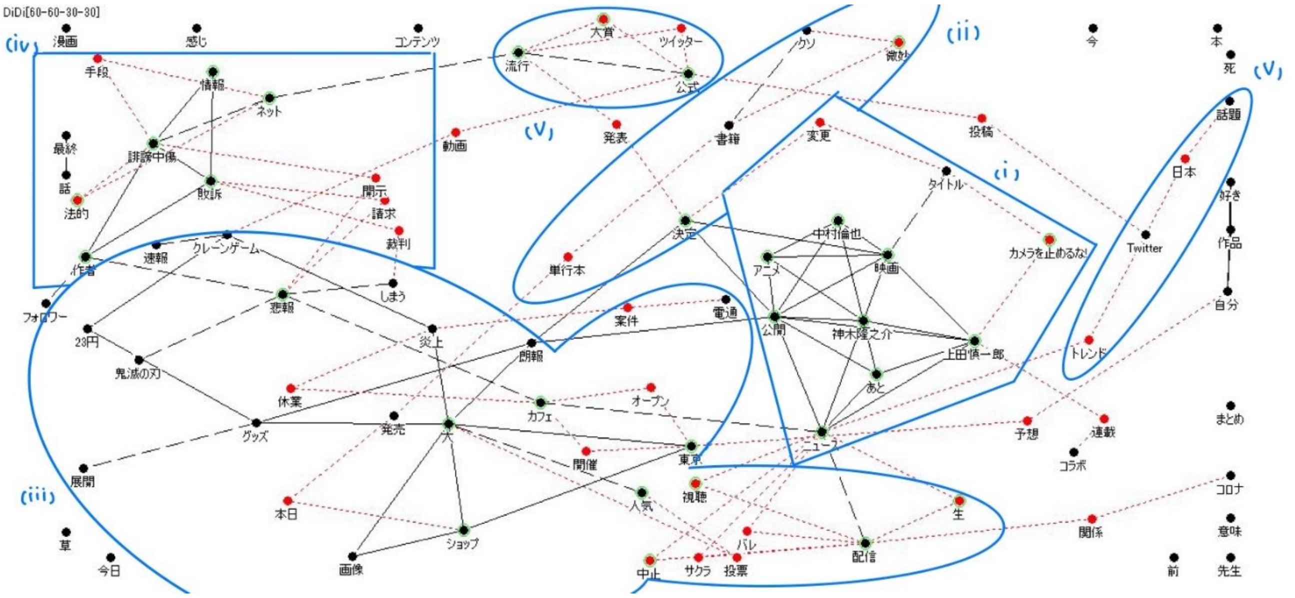

DiDi[60-60-30-30]
(iv)
漫画
感じ
コンテンツ
流行
大賞
ツイッター
公式
クソ
微妙
(ii)
今
本
死
(V)
手段
情報
ネット
最終
話
法的
誹謗中傷
敗訴
開示
請求
裁判
動画
(V)
発表
書籍
変更
投稿
タイトル
(i)
話題
日本
好き
作品
Twitter
自分
作者
速報
クレーンゲーム
しまう
悲報
フォロワー
決定
中村倫也
アニメ
映画
カメラを止めるな!
単行本
案件
電通
公開
神木隆之介
上田慎一郎
トレンド
23円
炎上
朗報
鬼滅の刃
休業
カフェ
オープン
あと
グッズ
発売
大
東京
ニュース
予想
連載
まとめ
コラボ
展開
開催
人気
視聴
生
コロナ
(iii)
本日
ショップ
バレ
配信
関係
意味
草
今日
画像
中止
サクラ
投票
前
先生

Fig. 2: Tweet data Period I (April 1, 2020 - June 30, 2021)

(i) Information and reaction about the film of this content.
(ii) Negative tweets about this content book.
(iii) Ridicule this content. E.g., compared film of this content with other popular work box-office (“Demon Slayer” (Koyoharu Gotouge)). The café of this content that closed within three days of its opening. The merchandise of this content was sold for ¥23 in a crane game. The turned-out shill are existing in the distribution of this content.
(iv) Reactions and ridicule a lawsuit filed by the author of this content in response to a defamation statement.
(v) Information and reaction to this content being hot topic about.

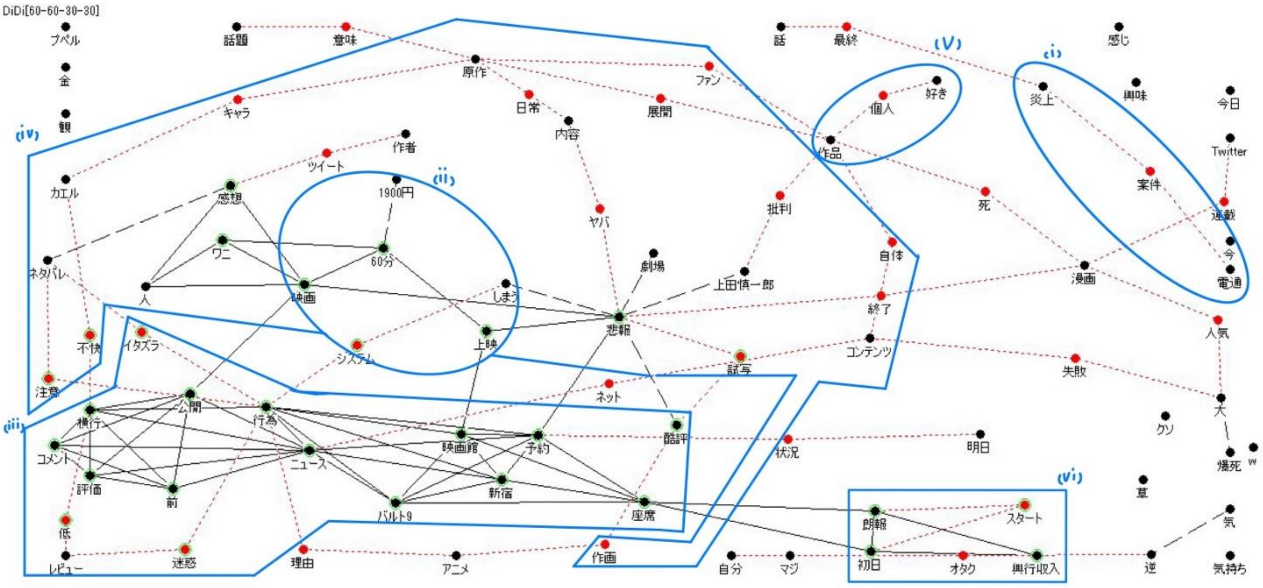

DiDi[60-60-30-30]
ブペル
話題
意味
原作
ファン
話
最終
(v)
(i)
感じ
金
キャラ
日常
展開
個人
好き
炎上
興味
今日
観
(iv)
内容
作品
Twitter
作者
ツイート
カエル
感想
1900円
(ii)
ヤバ
批判
死
案件
連載
ワニ
60分
劇場
自体
今
ネタバレ
人
映画
しまう
上田慎一郎
漫画
電通
悲報
終了
不快
イタズラ
上映
システム
人気
試写
コンテンツ
失敗
注意
ネット
(iii)
横行
公開
行為
大
酷評
クソ
映画館
予約
状況
明日
コメント
ニュース
爆死
w
評価
前
新宿
(vi)
草
バルト9
座席
朗報
スタート
気
低
レビュー
迷惑
理由
アニメ
作画
自分
マジ
初日
オタク
興行収入
逆
気持ち

Fig. 3: Tweet data Period II (July 1, 2021 - August 8, 2021)

(i) Reaction to the flaming by the alleged relationship between this content and Dentsu.
(ii) Information and reaction to the fact that the film for this content has a running time of 60 minutes and the price of the film ticket is ¥1,900.
(iii) Tweets about trolling behavior, such as rampant low-rated reviews before the film's release, or reservations without intent to see the film after its release.
(iv) Information and reactions about this content film.
(v) Positive comments on the film of this content.
(vi) Information and ridicule of the box-office revenue for the film in this content.

Discussion

Combining the data obtained from the results of the Sentiment Analysis and the results of the Key Graph Analysis, this section discusses the emotional swings in each period.

Emotional movement from Period Ⅰ (Before the film's release) to Period Ⅱ (Film release)

In the Tweet data, " dislike ", "fear" is increased and the "like" "joy" is decreased can be interpreted from the clusters of vandalism and criticism generated for this content film (Fig.2, Fig.3). As for the increase in "fear", it is likely that it is caused by the uncertainty of the film adaptation.

The results of the Sentiment Analysis and Key Graph Analyses suggest that the reason why the high ratio of joy is occurring in the tweet data may be due to the content of the flaming. The flaming began on March 20, 2020, and even though the film's release date is July 9, 2021, more than a year later, the effects of the flaming remain in a different form. The flaming period was unremarkable because it was seen as an issue to be discussed, but since then, it has become content as “flaming content,” and there is a segment of the population that gleefully watches court cases brought against harassment and defamation by the author (Fig.3). This includes cynicism and sarcasm.

Conclusions

The purpose of this paper is to visualize the “A Crocodile Who Will Die in 100 Days,” flaming, which is a flaming period, and to understand how the flaming have affected peoples' emotions.

The study found that the analysis revealed the following about the impact of flames on content

Once content goes up in flaming, it has become clear that content that has been under flaming will not stop being criticized for a long period of time, and that being flaming experience itself will be consumed as content.

This paper does not fully investigate the bias of the Tweeting users. In the case of a specific flaming, it is possible that the noisy minority phenomenon is being caused [5]. This phenomenon is the diffusion of information only by users belonging to a specific community.

Therefore, it is necessary to quantitatively evaluate whether the information spread at the time of the flaming incident is biased toward users in a certain community by investigating user bias using KL Divergence, as in the previous study.

References

1. Twitter IR. (2022). "Q4' 2021 Shareholder Letter." retrieved on December 25, 2022.
2. H. Yoshino, (2018). "A Study of Media Environments that Generate Internet Flaming and Characteristics of Flaming Participants." Graduate School Office, Chuo University, 809. (in Japanese)
3. A. Matsui, (2020). "Influence of the presence of Others in Social Networking." Japan marketing journal, 40 (3). (in Japanese)
4. M. Ptaszynski, P. Dybala, R. Rzepka, K. Araki, F. Masui, (2017). "ML-Ask: Open Source Affect Analysis Software for Textual Input in Japanese." Journal of Open Research Software, 5 (1).
5. F. Toriumi, T. Sakaki, M. Yoshida, (2020). "Social Emotions Under the Spread of COVID-19 Using Social Media." Transactions of the Japanese Society for Artificial Intelligence, 35(4).

Authors Introduction

Mr. Taiki Sugimoto

He is currently a 4th year student, a member of the Jun Nakamura Laboratory of Chuo University, Japan. The areas of his research interest are cognitive science and Artificial Intelligence. In 2023, he will join to the master's degree course of Chuo University.

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.

Protocol analysis for constructing Verbalizing Support System

Yuki Hayashi
Grad. School of Science and Engineering, Chiba University, 1-33 Yayoi-cho, Inage-ku, Chiba-shi, Chiba, Japan
National Institute for Japanese Language and Linguistics, 10-2 Midori-cho, Tachikawa-shi, Tokyo, Japan
E-mail: chiba-u.jp

Abstract

In this paper, I will show a result of an experiment. The experiment was conducted using voice recorder and the participants was required to report what they think orally. I have been conducting experiments to know how people express their feelings and thinking in their daily lives. Previously, I suggested the basic concept of Verbalizing Support System (in Japanese). However, in this experiment, I tried to show the process of verbalizing people's feelings and thinking. In addition, I'll show the revised concept of Verbalizing Support System compared with the previous paper.

Keywords: Protocol analysis, IADL, Dementia, Verbalizing Support System.

1. Introduction

In Japan, many people are interested in health. As people get older, however, physical activities decrease because of retirement, bereavement, and so on. In addition, there are many people who cannot go outside because of backache, arthralgia, and so on. The Care-nursing Support Website (Kaigo Ouen Net) says that people should look after aged people so that they can become independent. It is important that society as a whole support aged people, but social welfare spending is increasing year by year. Therefore, it is also important that people are not too dependent on society but stand on their own feet.

Generally, aging will be proceed if a person does not communicate with many people or does not often go out. Otake said that “The shortage of social communication is one of the main factor of dementia.[1]” However, when you want to do that, you should take psychological stress into account. For instance, the Music Puzzle is proposed as the way of the brain activation[2]. In this puzzle, people rearrange the sound-cell and complete a song. However, this method may be difficult to use constantly in daily life.

Therefore I will suggest to support people in the situation of shopping activities. Shopping activities are one of the daily activities, and people (both young and aged) should do that usually by themselves. There are many people who visit a real grocery shop and there are also people who use online shop or delivery service. Moreover, convenient stores are visited not only by young and also by aged recently in Japan, so shopping activities are one of the essential daily activities for many people.

In the following chapters, I will examine the intelligent aspect of shopping activities and propose the Verbalizing Support System

2. Collection protocol

I conducted the following observations to examine that shopping are intelligent activities. In these observations, I asked participants to report orally what they thought. Participants were 3 house-workers. The observation were conducted in the grocery store which each participant often visit. In the store, I traced participants and collected their utterance on shopping. I asked them in advance to report orally what they thought, for instance, what they would purchase, what they would make as dinner or how they would change their plan.

In previous researches on consumers' behavior, they usually observe consumer in a real grocery store. However, this way of observation may be difficult for observers and for participants. Abe pointed out that "The way of observation in my research requires lots of time and labor, so it is difficult to collect lots of shopping activities data.[3]"

Accordingly, I developed the Shopping Simulator. In this simulator, user can move a character and walk around virtual store. Upper part means a basket which the character has, and lower part means a floor in the virtual store. If you want a particular item, you can drag the picture of item from floor to basket. On the other hand, if you stop purchasing an item, you can drag the picture of item from basket to floor. There are many areas in the virtual store, for instance, vegetables, bakery or snacks. When a character approaches a certain area, user can see information on each item, which information fundamentally consists of name, amount and price.

Participants were 34. 12 of them were house-workers, 11 of them were boarders and 11 of them were home-students. I collected 37 shopping activities data and constructed a corpus based on participants' utterances.

I extracted knowledge about price, judgement about freshness or any other thinking from corpus, and we worked up these thinking into the list.

3. Verbalizing Support System

I aim to extract more active utterances and conversations from participants. In this chapter, I will propose the Verbalizing Support System in which people can perform intelligent activities with shopping. This system consists of the following steps.

First, user can plan their shopping activities. User can input an item name or a menu name by manuscript or voice. In this step, you can communicate with another user. For instance, you can ask "Why will you purchase that?" You can also give useful information or idea to the user, "Now is the best season for Pacific saury", "Today is hot, so how about eating fine noodle?", and so on.

Secondly, user can perform shopping. I recommend that they use Shopping Simulator so that they can chat or converse freely.

Thirdly, user can check the result of their shopping activities. User can use this result in the next step and compare it with plan made in firsr step.

Finally, user can experience more intelligent activities than the other steps. They can recollect the process of their shopping. They will experience intelligent activities. In addition, they can talk freely, for instance, what they find in shop, what they want to claim about shop, and so on.

4. Discussion

As I mentioned, I want to encourage people to perform intelligent activities. Now, how about the criteria of "intelligent"? I prepare the following definition: when a participant can show the "novel" information about shopping activities to the other participants (or the experimenters), this participant's activities are regarded as "intelligent". In the collection of shopping activities data, most participants perform shopping alone. However, I think their shopping activities should be evaluated by the other people. Accordingly, "novel" information will be discovered by them and chance for utterance or conversation activation will be also discovered.

5. Conclusion

In this paper, I showed a result of an experiment. The experiment was conducted using voice recorder and the participants was required to report what they think orally. In this experiment, I tried to show the process of verbalizing people's feelings and thinking. In addition, I showed the concept of Verbalizing Support System.

References

1. Otake Mihoko. "Conversation Support Robot for Promoting Cognitive Activity of Older Adults". Journal of the Japanese Society for Artificial Intelligence. vol.29, No.6, pp.591-598. 2014.
2. Noguchi Takafumi, Chida Kazunori, Sano Yoshihiko, Inamori Sakae. "Development of the educational materials using open-ended musical sound parts to foster endurance and concentration for the disabled person". Japanese Society for Information and Systems in Education. Vol.24, No.1,7. 2009. in Japanese.

3. Abe Shuzo. “Shopping Activities and Information Processing-a trial of protocol analysis-”. Yokohama management study. vol.4, No.2, pp.33-48. 1983.

Authors Introduction

Yuki Hayashi

He received the degree of Master of Arts from Graduate School of Humanities and Social Sciences, Chiba University.

Story generation during appreciating an artwork based on an actual tale (Ugetsu-monogatari)

Akinori Abe
Faculty of Letters, Chiba University, 1-33 Yayoi-cho, Inage-ku,
Chiba 263-8522, Japan

Abstract

In this paper, I will show a result of an experiment. The experiment was conducted using an artwork based on an actual tale (Ugetsu-monogatari). Actually the tale is rather old one. I have been conducting experiments to know how people can generate stories from the artworks. Previously I did not use artworks based on actual tales. Only the exception was artworks based on ``Alice in Wonderland." These artworks were used to see how Visual Thinking Strategies (VTS) functions during art appreciation. Of course in the experiment, we asked participants to generate stories. However, in this experiment, I tried to see the process of the story generation during art appreciation.

Keywords: Art appreciation, Ugetsu Monogatari, Story generation

1. Introduction

We have been studying how to help art viewers during art appreciation ([9], [10], [11], [12], [1]). Since sometimes it is rather difficult especially for the novice viewers to understand and enjoy abstract artworks, with a curation, we tried to help viewers' art experiences.

As Leder et al. suggested [7], [8], information may help the viewers in an art appreciation. How can we measure the level of understanding of viewers? One strategy is to make viewers to generate stories in the artworks. Previously we conducted such an experiment [1], [12], [13], [4], [5]. However, in the experiments, we did not use artworks based on actual tales. Only the exception was artworks based on “Alice in Wonderland” [13]. These artworks were used to see how Visual Thinking Strategies (VTS) functions during art appreciation. In the experiment, we asked participants to generate stories. However, our main purpose was to see how the others' opinion would influence one's art appreciation. In this paper, I will show the experiment that was conducted by using an artwork based on an actual tale (Ugetsu-monogatari).

In this experiment, I tried to see the process of the story generation during art appreciation. Accordingly I offered information in several levels. My hypothesis is that the level of understanding will be improved according to the quality or quantity of given information. In the following sections, I will show the experiment, its results, and my analysis.

2. Experiment

My hypothesis is that the level of understanding artworks will be improved according to the quality or quantity of given information. I think that the level of understanding can be measured by generated story of the artworks. Therefore, I conducted an experiment in order to determine which factor (information) will change the viewers' viewpoint in an art appreciation. In the previous papers [2], [3], we analysed the results only from the score of the value and preference of artworks. However, in this paper, I will show the change of participants' viewpoint of the artwork. For that, we asked the participants to generate a story of the artwork.

2.1. *Participants*

Participants were 20 adults including university students and aged person and their age were from 19 to 58 years old. Number of females was 8 and that of males was 12. Their naked visions or corrected visions were normal. First, all the participants were asked to answer about their art education background.

Fig. 1. Experiment

Then, main questions followed.

2.2. *Stimuli*

The experiment was conducted in a lecture room in Chuo University on June 3, 2022. I used an artwork created in 1970 as a stimuli. The artist is Donge Kobayashi (小林ドンゲ) (1926–). The title of the artwork is “Dear my lover, now, let’s go together (いとしいお前よさあ一緒に行こう).” The artwork was placed in front of the participants like Figure 1. The participants could look at the artwork closely.

2.3. *Method*

All participants answered the questions on the worksheet during appreciating the artwork. The questions were as follows;

・Question 1: About this artwork, please write freely what you think, what you feel, and a story in the artwork. And please make a title of the artwork.

・1.1 Do you like such an artwork?

All evaluations are conducted by using the SD (Semantic Differential) method with the score of 1 to 5. For instance, the score 1 for dislike, the score 5 for like and the score 3 for neutral. The score 1 for value is low and the score 5 for value is high.

・ 1.2 Why do you evaluate so?

・ 2.1 How much do you understand the story of the artwork?

・ 2.2 Why do you evaluate so?

・ 3.1 How much are you interested in the artwork?

・ 3.2 Why do you evaluate so?

I showed the information such as a name and a history of the artist. In addition, I showed the title and techniques of the artwork. Then I asked the questions as follows;

・ Question 2: Please make stories in the artwork with referring to the information. Does the above information influence your mind?

In addition, the same questions such as the above (1.1–3.2) followed.

I showed information like “this artwork was created as an attached of Ugetsu-monogatari (雨月物語) by Shusei Ueda (上田秋成).” Then I showed the story in original language and the story translated into the current language. In addition I showed a very short abstract of the story. Then I asked the questions as follows;

・ Question 3: Which part of the story the artwork draws? Please write impressions of the artwork and changes in mind after reading the information as many as you can.

In addition, the same questions such as the above (1.1–3.2) followed.

・ the last question: If you change your mind after obtaining the information, please write down. For instance, a content or a story in the artwork. Do you like or dislike the artwork?

At the last, please write the opinion to the experiment. For instance, you needed more time, and the selection of the artwork was bad, etc. If you conduct this type of experiment, you will ask the other questions...

The duration of the experiment was about 30 minutes, for one artwork, it will be longer than the previous experiments.

3. Result and Discussions

From the results in [9], we observed that frequent museum visitors tended to answer that information about artist’s life, historical background and theme of the artwork were helpful. We also observed that many non-frequent museum visitors felt information about techniques used in the artwork was helpful. Findlay [6] pointed out five attributes (Provenance, Condition, Authenticity, Exposure, Quality) must be known for determination the value of artworks. However, since most of artworks were obtained in the primary market, factor as provenance will not be considered. Therefore, in the experiment, we prepared information about techniques and artist’s history. In this paper, I will analyse the result without considering participants’ artistic education. The following answers were originally written in Japanese,

3.1. *Answers to the Question 1*

egard a boy as a girl. Actually he seems a girl. Accordingly they made stories from the context of love of male and female. For instance, some participants described as follows:

・A

I feel a dark image close to a death. And I feel his/her facial expression is dark. Not from his/her facial expression, but from the existence of a skeleton I image a death. In addition, since I see the fact that he/she is surrounded by flowers, and a skeleton exists, I image a death.

・B

A person whom she has been seeing (she has loved) goes away (perhaps the person is described as a skeleton, so the person is dead?), so she has been sad. The scene her hair is wrapped up implies that she wants to be embraced by the person. The person reminds that they walked around a park with a flower garden with being blown by the wind... their hair flowing in the breeze.... Then the person wake up in the morning after dreaming the person wants to meet her again.

Most of participants referred to the skeleton and flowers. Actually they thought the boy as a girl. So they generated stories from the context of love. They felt the situation of the death but generated romantic and sad stories.

3.2. *Answers to the Question 2*

・A

The title “Now let’s go together” was what the skeleton said to the female. Then I strengthen my interpretation that the artwork implicates a death. From the fact that a calling such as “Now” is spoken and from the facial expression of the female, I feel the sad impression of a death.

・B

I changed my mind. -> In the question 1, the main person is a girl in the left. However, after reading the question 2, the main person changes to a “skeleton” in the right. I cannot see the facial expression of the skeleton, accordingly various imaginations emerged. That is, I changed point of view from that the skeleton is a dead man to that it is a symbolic living male.

I offered a profile of the artist and the title of the artwork. Some participants created the new stories such that a person bring his lover to the world of death. Anyway stories are still a love between male and female. In addition, still they feel a death situation.

3.3. *Answers to the Question 3*

・A

I changed my interpretation that the person I thought a female was a boy, and the person I thought a skeleton was a monk in the Daichu-ji (大中寺) temple. Because I thought that the artist expressed the skeleton as after the death of the boy, monk in the Daichu-ji temple losing a power of live transformed into a demon. The boy died and the monk is still living. However, the artist drew in an opposite way. I felt that this artwork contains a theme that the problem is that in our life, a soul is more important than a body.

・B

I changed my imagination; The female in the left is expressed as “a female whose gender is twisted.” -> The female transformed into a demon and finally was dead like a skeleton.

Finally the participants were aware that a girl they thought is a male (monk). Of course they change their stories. In this phase, they read a whole story. Some of them understand the context behind the artwork, then accept the artwork. But some cannot accept the artwork. The artist expresses this story as a passion in the end of the century. Accordingly the artwork becomes somewhat uncomfortable, unusual and extraordinary. I think her artworks seems like style of Odilon Redon.

3.4. *Answers to the last question*

・A

From the beginning, it is difficult to read the story of the artwork. For me, the artwork is very interesting. Because after obtaining the true story, I changed my interpretation. After feeling the artist’s intention, I remember to go to the 21st century museum of contemporary art, Kanazawa (金沢21 世紀美術館) in this spring. In the museum, some artworks were not provided artists’ intention nor background of the artworks. It is my personal opinion.... it will be good to compare the background story with my own interpretation during an art appreciation.

・B

The story is unrealistic. (It is rare for me to read this type of novel.) This artwork symbolically expresses the extraordinary situation. The artwork is drawn one scene of the story. In a certain sense, I understand the situation. -> The fact that the artwork expresses the unrealistic matters means that it contains a considerable stories. If this were an artwork drawn by, for instance, Picasso, it would not contain a story or be regarded as a completely different level of artwork. As for an artwork such as this, items (skeleton, flower, hair, female...) constructing an artwork have a certain meaning. I feel the depth and the texture of the context again.

Some participants said that the story was very difficult. Actually, this story was written in from 1768 to 1776. That is, it can be regarded as classic. I offered a story translated into modern Japanese, in addition. But the story itself is rather different from sense of people in younger generation. An opinion that is interesting is; since this artwork is composed of certain shapes (a human, a skeleton, and flowers can be recognized), if the story is added, no effect of story is shown. If this artwork were more abstract, they could capture the content differently according to the given story. This opinion seems interesting. We may consider the opinion. Previously we found the similar phenomenon in the appreciation of artwork by Jean Fautrier [9]. The artwork is an abstract artwork. A certain participant after reading the title (Foˆerˆet (forest)), the participant could create a story “I found a fairy in a forest.” in the context of a forest.

3.5. *Discussion*

My hypothesis is that the level of understanding will be improved according to the given information. The above results are, in a sense, very interesting. Most of participants pay attention to the skeleton and the flowers. In addition, first, they thought the boy as a girl. Then the first impression was love. However, according to the existence of the skeleton, they also felt a death story. After obtaining the title of the artwork, a certain participant noticed the calling such as “Now” is spoken, he/she thought the artwork implicated a death situation. By this phase, the participants read the story from the aspect of death and love. After obtaining the whole story, most of the participants changed their mind. Because the person whom they thought was a girl was actually a monk, and the person whom they thought was a skeleton was actually a boy. They read the artwork in the context of creepy and scared. The effect of information is very clear in this experiment. However, the problem is that the story is very difficult to read and extraordinary.

4. Conclusions

In this paper, I conducted the experiment to see whether the level of understanding would be improved according to the given information. My hypothesis was that the level of understanding will be improved according to the quality or quantity of the given information.

I used an artwork created in 1970 as stimuli. The artist is Donge Kobayashi (小林ドンゲ) (1926--). The artwork is not so abstract. The participants can recognize what are drawn in it. However, it may rather symbolic. Because it contains a lot of items such as flowers and a skeleton. They may not be an attribution of European artworks. However they may contain certain meanings. The result is as follows;

After obtaining the title of the artwork, a certain participant noticed the calling such as ``Now" is spoken, he/she thought the artwork implicated a death situation. By this phase, the participants read the story from the aspect of death and love. Finally, after obtaining the whole story, most of the articipants changed their mind. Because the person whom they thought was a girl was actually a monk, and the person whom they thought was a skeleton was actually a boy. They read the artwork in the context of creepy and scared. Thus according to the given information, participants might change their thinking to the artwork. The result is very interesting from a view point of a narrative generation. Because I could observe the change of a thinking procedure and a story generation in the art appreciation.

A problem is that I used a rather classic novel and it is rather extraordinary. Accordingly for some participants, it is difficult to read the novel. In the future experiment, I have to select a more readable and normal story.

References

1. Akinori Abe and Kotone Tadaki: Captions with Several Levels of Explanation, In *Proc. of KES2019*, pp. 2335–2344 (2019)
2. Akinori Abe, Kotaro Fukushima and Reina Kawada: How will sense of values and preference change during art appreciation?, *information 2020, 11(6), 328 (Special Issue on CDEC: Cross-disciplinary Data Exchange and Collaboration)*, (2020)
3. Akinori Abe: Expression of the taste of Akinori Abe, Kotaro Fukushima: How will the information change the sense of values during art appreciation?, *Proc. of KES2020, Procedia Computer Science, Volume 176*, pp. 3083–3092 (2020)
4. Akinori Abe: How will art appreciations change according to information change?, *Proc. of ICAROB2022, OS31-9*, pp.1054--1057 (2022)
5. Akinori Abe: Abduction in Art Appreciation, Philosophies 2022, 7, 132 (2022)
6. Michael Findlay:*The Value of Art*, Prestel (2012)
7. Helmut Leder, Belke, B., Oeberst, A., & Augustin, D.: A model of aesthetic appreciation and aesthetic judgments, In *British Journal of Psychology, 95(4)*, pp. 489–508 (2004)
8. Helmut Leder, Claus-Christian Carbon, and Ai-Leen Ripsas: Entitling art: Influence of title information on understanding and appreciation of paintings,” *Acta Psychologica, 121*, pp. 176–198 (2006)
9. Kotone Tadaki and Akinori Abe: Museum Visitors’ Behavioral Change Caused by Captions, In *Proc. of the 2nd. Int’l Workshop on Language Sense on Computer in IJCAI2017*, pp. 53–58 (2017)
10. Kotone Tadaki and Akinori Abe: The Influence of Story Writing Worksheets on Art Appreciation, In *Proc. Of ICAROB2018*,, pp. 687–690 (2018)
11. Kotone Tadaki and Akinori Abe: The Potentiality of First Person :Analysing Narratives During Art Appreciation, In *Post-Narratology Through Computational and Cognitive Approaches (Ogata T., Akimoto T. eds.), Information Science Reference*,, pp. 375–396 (2019)
12. Kotone Tadaki and Akinori Abe: The Influence of Story Creating Activities While Appreciating Abstract Artworks, In *Advances in Artificial Intelligence. JSAI 2019 (Ohsawa Y. et al. eds.), Advances in Intelligent Systems and Computing, vol 1128, Springer Verlag*, pp. 126–135 (2020)
13. Shunsei Tokida and Akinori Abe: What Kind of Influence Can Be Obtained during Groupwork in Art Appreciation?, *Proc. of KES2022, Procedia Computer Science*, pp. 3998–4007 (2022)

Authors Introduction

Dr. Akinori Abe

He received an M.E. and Doctor of Engineering (PhD) from the University of Tokyo in 1988 and 1991 respectively. His main research interests are abduction, analogical reasoning, data mining, Chance Discovery, curation and language sense processing (rather emotional aspects of language processing). He worked in NTT Communication Science Laboratories from 1991 to 2000, After that, he worked several places, and he moved to Faculty of Letters, Chiba University in April 2012. Since then he is a full professor of Chiba University.

Designing a Narrative Generation Game Based on the Russian Invasion of Ukraine

Jumpei Ono
Faculty of Software and Information Technology, Aomori University,
10-1 Seishincho, Edogawa, Tokyo, 134-0087, Japan

Takashi Ogata
Faculty of Software and Information Science, Iwate Prefectural University,
152-52 Sugo, Takizawa, Iwate, 020-0693, Japan
E-mail: j.ono@aomori-u.ac.jp, g031q057@s.iwate-pu.ac.jp
t-ogata@iwate-pu.ac.jp

Abstract

In a previous paper, we proposed an automatic narrative generation game. In this paper, we propose a game concept based on the framework of the automatic narrative generation game. The proposed game is based on the ongoing Russian invasion of Ukraine (beginning in February 2022). The game is based on an analog game in which participants advance a shared storyline mainly through dialog. The proposed game facilitates story generation. Accordingly, as an application, we intend to use the game to support the creation of stories.

Keywords: Story Generation, Automatic Narrative Generation Game, Ukraine, TRPG

1. Introduction

Caillois [1] defined play as "an activity that is essentially isolated and carefully insulated from the rest of life, which is usually completed within the definite limits of time and place." In general, play is an isolated and fictional activity. Play is generally executed through a shared narrative among participants under the control of the rules of play. In the space of play, participants experience events in a simulated manner through their own independent actions during the period specified by the rules of the particular play. As play is a fictional activity, it has no direct impact on the real world. Therefore, failures are allowed in play.

The above qualities in play are also present in games (the "play" referred to by Caillois includes a wide range of activities such as sports, theater, raffles, and merry-go-rounds). The "play" we refer to as "games" herein narrows the scope to games as defined by Juul [2]. Notably, there are examples of using games as a medium for proactively simulating a narrative. For example, games such as table-top role-playing games (TRPGs) and live-action role-playing games (LARPs) are used to simulate fictional and real-life stories. In particular, the player plays the role of a character in a fictional story or real historical event. LARPs allow players to simulate a subjective perspective in a story by playing the role of a character; correspondingly, their nature has attracted attention as a method for learning cultural heritage [3]. However, according to Ensslin [4], the voices and perspectives of narrators in novels and films have been replaced by the experiential qualities in games. In this context, whether a game itself can be considered a "narrative" has been debated.

In particular, there are games concerning events in the real world (such as those for learning about cultural heritage) and intended for educational purposes; these are known as "serious" games. They have been defined as "(digital) games that are developed and used to solve problems in various areas of society" (Toru Fujimoto, CEDEC 2022 lecture,*). One recent example is *Ukraine War Stories* [5], an adventure game based on the Russian invasion of Ukraine beginning on February 24, 2022, and still ongoing as of this writing. The game was released in October 2022 and a demo version is currently available. There are three storylines set in Hostomel, Bucha, and Mariupol, respectively. All of the game's images are based on actual events and eyewitness testimonies. In addition, the images used in the game are based on photographs taken on the battlefield.

---

* https://cedec.cesa.or.jp/2022/session/detail/115

In a previous paper, we proposed a narrative auto-generating game modeled on a TRPG [6]. This paper proposes a game concept based on the framework of the auto-generated narrative game. This paper is also an extended version based on [7]. As with *Ukraine War Stories*, our game is also based on the Russian invasion of Ukraine beginning in February 2022 and still ongoing (September 2022). In this paper, we refer to the invasion as the "Ukrainian War."

2. What's TRPG

TRPG is an analog game in which participants advance a shared storyline, mainly through dialog. The shared storyline works as a constraint. There are two roles in TRPGs. One is the game master (GM), and the other is the player (PL). In playing a TRPG, the GM and PLs share a single storyline called a scenario.

Figure 1 is an image of a scenario in TRPG. The GM has a fair overview of the scenario and, depending on how the PL's actions turn out, modifies the scenario's scenes to steer the story in a certain direction. The resulting story can be a tragedy or comedy. However, as TRPGs are intended to be entertaining, the ending should usually satisfy both the GM and PLs. PLs insert new scenes according to the scenarios provided by the game master. PLs do not know the entire scenario; PLs insert scenes based on the scripts of the various stories memorized by the PLs.

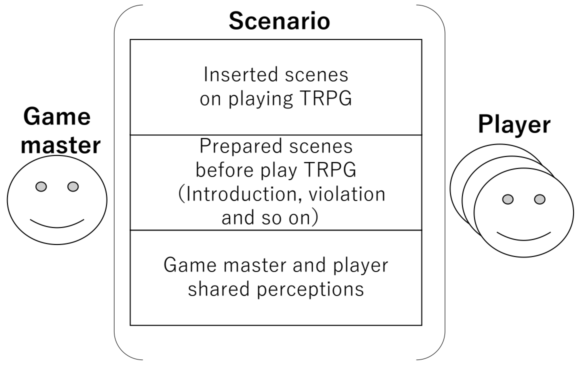

Scenario
Inserted scenes on playing TRPG
Prepared scenes before play TRPG (Introduction, violation and so on)
Game master and player shared perceptions
Game master
Player

Fig. 1. Image of a scenario in TRPG

Because TRPG is a game and virtual experience, it provides an environment in which it is easy (and acceptable) to make mistakes. There have been attempts to use TRPG to support leisure time activities for children with developmental disabilities [8]. In addition, there may be large or small differences in the continuation of the shared story as recalled between the participants. The main plot of the story may change according to such differences. In addition to "actually playing," there are other ways to enjoy TRPGs, such as by recording and publishing scenes of the game in writing or video, or live-streaming the game on a video distribution site. In other words, the game itself is a work of art that can be watched and enjoyed. The game proposed herein facilitates story generation; as an application of this game, we intend to support the creation of stories through the game.

3. Proposed Game Format

The proposed game is a text adventure game in which the player must maintain morale (which is being lowered by public opinion that the nation should surrender) and keep the nation alive. In terms of existing works, the proposed game is similar to a game that simulates the manipulation of public opinion through actual or fake news, as discussed below.

Headliner [9] and *Headliner: Novi News* [10] are adventure games in which the player becomes the editorial director of a news bureau or newspaper in a fictional nation and manipulates public opinion. The game is not about fake news *per se*, but the selection of the articles submitted by writers can drastically change the state of affairs in the country where the game takes place. *Plague Inc.* [11] is a simulation game in which the player becomes an epidemic, aiming to evolve and destroy the human race. In 2019, under the supervision of the American fact-checking organization Politifact and others, a mode for spreading fake news was added to the game. *Bad News* [12] is a game developed by researchers at the University of Cambridge for simulating the spread of fake news.

In terms of story creation support, the game proposed herein is similar to a game called "Solo Journaling RPG" (https://itch.io/jam/solo-journaling-rpg-jam). Solo Journaling RPG concerns creating a story and writing a record from an introspective game experience. The story is created by the player by matching relationships between events and the storyline from fragments and sequences of events selected using random number generators such as dice and playing cards. The main focus is on how the players themselves (or the characters they have created) feel about the events, as recorded in their diaries or monologs. As another example, *Skyworthy* (developed by A Couple of Drakes, https://acoupleofdrakes.itch.io/skyworthyzine) is a game in which a player acts as a flying boat with many sailors on board and writes a diary of his observations of the lives of the crew members. Other narratives include a ronin who saves a village from bandits who owe him a night's lodging and meal as inspired by Akira Kurosawa's "Seven Samurai," an adventure story about the inhabitants of a fairy tale world, and a life memoir for a vampire with one hour left to live.

4. Model for Story Generation

We propose a model for story generation based on the automatic story generation game proposed by Ono and Ogata [7]. Figure 2 shows the model. A scenario is a structure that determines the scenes that start the story and scenes that can occur during the story. The story is a collection of scenes told by the GM. The story state changes with the scenes. Each PL monitors the state of the story and interfere with the story when the state of the story meets certain conditions. Here, "interference" refers to the process of rewriting the scenes that the GM plans to tell. The user creates a single PL.

Figure 3 shows the process of story generation in the proposed model. The GM inserts a scene consisting of an event and corresponding states into the story. The GM inserts a scene according to the story framework prepared in advance (the framework is prepared by a user or another generation system). States has information about the characters, objects, and places in a scene.

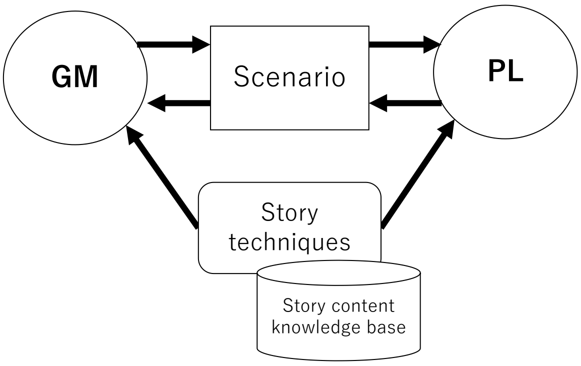


Fig. 2. Model of story generation

Based on the scene inserted by the GM, the PL inserts a scene following the scene or overwrites the scene with a different scene. PL's insertion gives small change to a story. The PL inserts a scene that is related to the scene inserted by the GM. Therefore, the scenes inserted by the PL often follow the flow of scenes that the framework has. In this model, scene relationship means a causal or temporal relationship. On the other hand, overwriting gives big change to a story. By rewriting a scene that the framework has, it cuts off the flow of scenes that the framework has in the future. Overwrite expects to generate different stories.

The GM changes the scenes to be inserted based on the scenes indicated by the PL. The GM will insert a scene that deviates from the prepared story framework, especially if the PL's overwrite creates a significant change in the story. The following paragraphs describe the methods of generation more specifically.

The GM and PL control the characters in the scenario. The characters have story techniques that change the storyline in the scenario. As mentioned above, the story is based on the Ukrainian War. We prepared the storyline for the Ukrainian War based on the article "The invasion of Ukraine in chronological order" published by Sankei Shimbun (https://www.sankei.com/article/20220301-KYQCEGX7GRFBPJBC2BFQPXQ754/). The article concerns the Ukrainian War and is arranged in chronological order, starting with the publication of the article at 12:18 on February 24, 2022. Notably, the article is still being updated. The characters are the people and government in Ukraine, people and government in Russia, and "World Commander Hashimoto-kun" [13]. The PL corresponds to one of the characters and the GM corresponds to the other characters.

The generation algorithm is as follows. In this description, a story is a variable for storing the stories to be generated. The list of processes is a variable storing the reserved story generation processes.

(i) The user enters a scenario.
(ii) The user determines PL1.

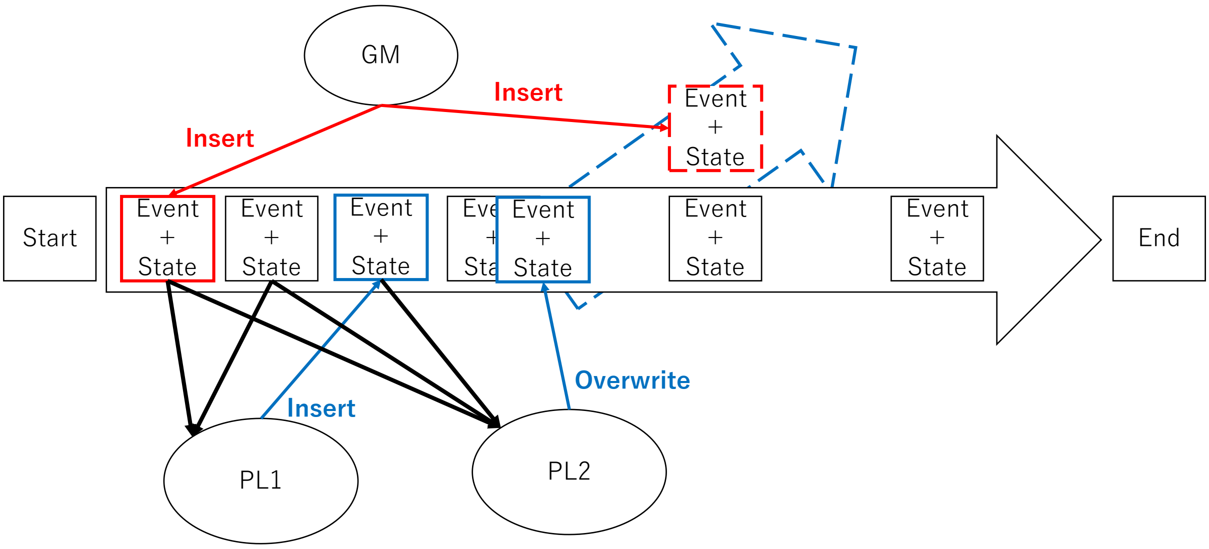


Fig. 3. Image of story generation

(iii) The process from the GM is inserted into the list of processes.
(iv) One process is selected from the list of processes.
(v) Scenes are inserted into the story based on the retrieved process.
(vi) A process is inserted into the list of processes (or not) based on the last scene added to the story.
(vii) If the list of procedures is empty, generation is terminated. Otherwise, the algorithm returns to Step (iv).

Morale is also set as a game mechanism. Morale is expressed as an integer with an initial value of 100. There are no lower or upper limits. In addition to increasing or decreasing depending on the scenes inserted by the GM, the morale also changes depending on the scenes inserted by the PL. A scene increases or decreases morale by a certain amount. The GM transitions the war as the story progresses. The PL inserts a scene in which anti-war statements decrease morale. The GM inserts a scene in which the PL fully accepts anti-war views and surrenders if morale falls below 0. The PL's objective is either to accept the surrender and end the story as an anti-war faction or to continue the story until victory as an anti-war faction.

5. Example of Story Generation

As mentioned above, the system uses Sankei Shimbun's timeline of the war in Ukraine. In particular, the timeline is used as transcribed through April 31, 2022 and classified by item. The "date" and "time" refer to the date and time the article was published. The content is the title of the article and indicates the event that took place. "Subject" indicates the classification of the event. "Military" refers to military actions and damage. "Negotiations" refers to meetings, conferences, and speeches. "Economy" refers to finances, imports, and exports. "Anti-war" refers to anti-war speeches and demonstrations. "Sports" refers to sports played by athletes.

In the story generation, the GM first presents a scene in which Vladimir Putin decides on a military operation in eastern Ukraine. In response, the PL, acting as "World Commander Hashimoto-kun" (hereafter referred to as "Hashimoto-kun"), is presented with the following instructional statement: "The mission of a country with no military power (such as Ukraine) should be to protect its citizens, and should refrain (in Hashimoto Toru's opinion) from raising grandiose ideals that are beyond its own power." He also mentions that "Japan, which is constrained by Article 9 of the Constitution, should do the same." In the context of Ukraine (with low military strength), this directive statement encourages flight and decreases its morale as a nation. The anti-war PL (user) responds "yes" or "no" to the statement. Morale is decreased when the directive statement is affirmed and increased when it is denied. The directive does not necessarily only decrease morale; in alternative cases, the increase or decrease in morale owing to an affirmation or denial is reversed.

6. Conclusion

We proposed a game concept based on the Ukrainian War using the framework of a narrative auto-generating game. Based on the timeline of the Ukrainian War as published by Sankei Shimbun, we prepared a story of the Ukrainian War up to April 30. We then summarized the characters and progression of the story based on the Ukrainian War story within the framework of the story auto-generating game. We then presented an example of story generation using the prepared story. In this example, the story progressed when the user (i.e., the player) accepted or rejected the statements of instructions given by "World Commander Hashimoto-kun."

Acknowledgements

This research was supported by JSPS Grant-in-Aid for Scientific Research JP21K17870.

References

1. Caillois, R., TadaM., Tsukazaki, M. (Trans.), Asobi To Ningen. Kodansha, 1990.
2. Juul, J., Matsunaga, S. (trans.), Harf Real: Kyojitsu no aida no video game. New Games Order, 2016.
3. Mochocki, M., *Role-play as a Heritage Practice*. Taylor & Francis, 2021.
4. Ensslin, A., Video Games as Complex Narratives and Embodied Metalepsis. *Routledge Companion to Narrative Theory*. pp. 411-422. Routledge, 2022.
5. Starni Games (2022). "Ukraine War Stories." Starni Games.
6. Ono, J., Ogata, T., Monogatari Jido Sesei Game Ni Okeru Odoroki To Monogatari: Odoroki Ni Motoduku Story Seisei Notameno Gap Giho. In, Ogata, T. (ed.), Post-narratology No Shoso: Jinko Chino No Jidai No Narratology Ni Mukete 1. Shin'yosya, 2021.
7. Ono, J., Ogata, T., Narrative Generation Game Based on the Ukraine War. The Japanese Society for Artificial Intelligence SIG-LSE0C31, 69, 143-148, 2022.
8. Horiguchi T., Table-talk Role Playing Game (TRPG) no Jissen. In, Kato, K. (ed.), Tsuge, M. (supervision), Hattasu Shogai No Aru Kodomo / Wakamono No Yoka Katsudo Shien Handy series Hattatsu Shogai Shien / Tokubetsu Kyoiku Navi. Kaneko Shobo, 2021.
9. Unbound Creations, The Headliner. Unbound Creations, 2017.
10. Unbound Creations, Headliner: Novi News. Unbound Creations, 2018.
11. Ndemic Creations, Plague Inc.. Ndemic Creations, 2021.
12. Cambridge Social Decision-Making Lab., Bad News Game. https://www.getbadnews.com/en/play, 2018.
13. Ogata, T, 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.

Authors Introduction

Dr. Jumpei Ono

Jumepi Ono 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

Takashi Ogata 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.

Weather Forecast System for Mobile Devices

Hu Jiahao, Neesha Jothi, Thong Chee Ling
Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia

E-mail: neesha@ucsiuniversity.edu.my
www.ucsiuniversity.edu.my

Abstract

In the 21st century, the Internet has developed rapidly, and its technology has reached an unprecedented level. In the 21st century, many personal computers appeared in daily life, and people began to be familiar with obtaining the required information through the Internet. Subsequently, as mobile development technology became increasingly popular and rapidly occupied the Internet market, the development of mobile terminals also made a breakthrough. At present, Android and iOS platforms dominate mobile platforms. Most applications are developed based on a single platform, and the development efficiency is not ideal. Therefore, cross platform technology has gradually entered the vision of developers, easing the problem of cross platform development. It provides opportunities to meet people's needs for real-time, accurate and diverse weather information. The development of such applications should focus on the design of interfaces and the realization of key functions.

Keywords: healthcare, mobile app, mobile application

1. Introduction

We are now in an era where technology is changing our lives. How to better meet the needs of users has always been an important topic of mobile platform development, and the eternal theme of weather has naturally become a center. After all, weather information plays an important role in people's daily travel, agriculture, industry and other important fields [1]. The project selected is based on the functions realized, taking weather as the theme and weather query as the purpose to find, mine and process appropriate data sources, and finally obtain data to provide real-time weather query [2].

2. Literature Review

In these papers, we have made a lot of analysis on the selection of the system arrangement framework. First of all, native development is a topic that cannot be bypassed across platforms. Native development comes first and then cross platforms. Native development refers to the development of applications in a specific operating system, mainly using the SDK provided by the corresponding platform. For example, the Android platform uses Java or Kotlin to call the API (application programming interface) provided by the SDK to develop Android applications. Similarly, the iOS platform uses Swift or Objective-C to call the API to develop Apple applications. The developed finished applications can only be used on specific operating systems they can only be put on the market on specific platforms.

2.1. *Cross Platform Development*

Cross platform technology means that the software development process does not depend on a certain hardware environment or operating system, and it can be applied to different operating systems or platforms through one-time development without modification or minor modification [3].

At present, many high-level program languages can be cross platform, such as C++, Python and other development languages. Take Python for example, in Linux or Windows environments, you only need to configure the script provided by Python and then install it to compile and run the same Python file (but the libraries in the code may support different platforms). On the basis of language cross platform, developers will classify and deploy the resources of different platform systems to encapsulate them into a new cross platform library, providing an interface for more upper layer work, step by step from bottom to top, and finally achieving the realization of cross platform framework. Developers use cross platform frameworks to develop cross platform applications, and then compile them from top to bottom to generate executable applications on different platforms. This is a simple cross platform application generation process.

2.2. *Vue+SpringBoot platform*

Before development, we need to set up an environment, that is, Vue's environment and SpringBoot's environment. Vue environment setting: the development environment of Vue3 is selected. There is an image interface available in Vue3, which is convenient. Spring Boot development preparation: The IDEA development environment is used, so there is already a Spring Boot shortcut development method in the IDEA. The busy classes do not need to be deployed.

2.3. *Status Management*

State management is an eternal topic in responsive programming. The idea of discussing and solving state management problems in a framework that supports responsive programming is consistent [4]. This section mainly introduces vuex, which is specially designed for vue.js application development state management mode. It uses centralized storage to manage the state of all components of the application and uses corresponding rules to ensure that the state changes in a predictable way [4]. Vuex is also integrated into the official debugging tool devtools extension, which is also the state management mode used in this project.

2.4. *Asynchronous Requests*

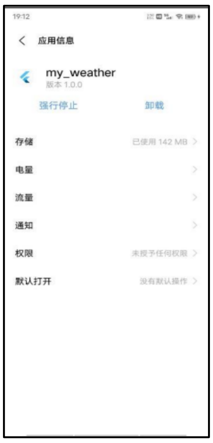
The core operation of Spring Boot+Vue's front and rear end separation is to complete data synchronization through asynchronous requests, which can be divided into many different situations, such as GET requests or POST requests, whether the parameters are ordinary variables or JSON, and how to operate based on RESTful architecture. Asynchronous requests in Vue are completed using the axis component. Axis is an HTTP client based on Promise for browsers and nodejs, which can be used in browsers and node.js.

2.5. *Database*

The background data of this project is realized through Request and connection to MySQL. MySQL is mainly used as the database, which is a mainstream choice. It is excellent in terms of performance and volume. Of course, there are also some disadvantages. However, the database is not the focus of our discussion here, and it can be used temporarily. Most of the data of this project is obtained through Request request using the API provided by related websites. The data obtained through the Request API are all in the json string format. JavaScript provides the function of converting json to dictionary type. After converting json to dictionary type, we created a variety of model classes to facilitate the use of the required data, and extracted and used data in the model class format. The database used on the mobile platform is mainly SQLite, which is a lightweight database embedded in our mobile client. However, due to the functional requirements, a large amount of data should be stored in the remote database. SQLite is not suitable for remote connection, so MySQL was finally selected. Similarly, when reading the database data, we will eventually convert it to the Model class format for the supplier to use.

3. Results

Opening the software for the first time requires permission as shown in Fig.1. Entering the software page without location permission is shown in Fig.2, but you can still search for the specified city. There is an animation effect when loading, which enhances the smoothness of software use, as shown in Fig.3.



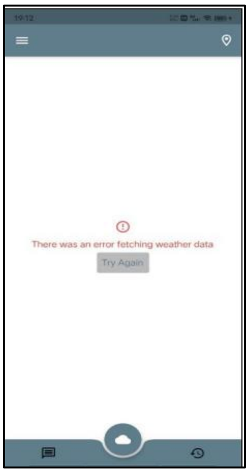
19:12

应用信息

my_weather
版本 1.0.0

存储 已使用 142 MB
电量
流量
通知
权限 未授予任何权限
默认打开 没有默认操作

Fig.1. Authority application



There was an error fetching weather data

Fig.2. Permission failed

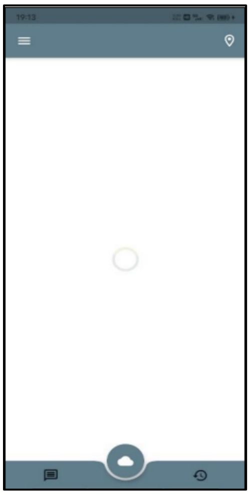
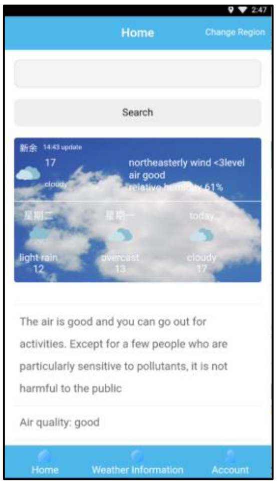


Fig.3. Loading page

After obtaining permission to use the location, enter the software, and the home page will display the city's name, weather conditions, real-time temperature, maximum and minimum temperatures of the day, hourly weather in the next 24 hours, body temperature and other indexes. As shown in Fig.4 Sliding the upper middle area can convert the module, so that the upper middle area can be converted into a weather line chart for the next 7 days. Similarly, sliding the 24-hour weather module can be viewed horizontally, as shown in Fig.5.

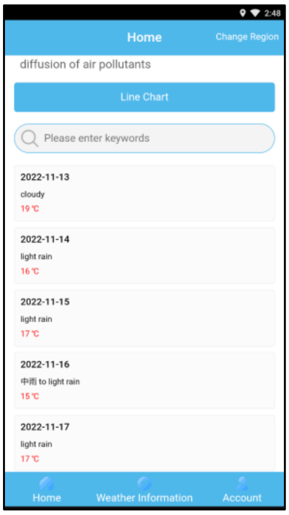


2:47
Home
Change Region
Search
新余 14:43 update
17
cloudy
northeasterly wind <3level
air good
relative humidity 61%
星期二
星期一
today
light rain
12
overcast
13
cloudy
17
Home
Weather Information
Account

The air is good and you can go out for activities. Except for a few people who are particularly sensitive to pollutants, it is not harmful to the public

Air quality: good

Fig.4. Main page 1



2:48
Home
Change Region
Line Chart
Please enter keywords
Home
Weather Information
Account

diffusion of air pollutants

2022-11-13
cloudy
19 ℃

2022-11-14
light rain
16 ℃

2022-11-15
light rain
17 ℃

2022-11-16
中雨 to light rain
15 ℃

2022-11-17
light rain
17 ℃

Fig.5. Main page 2

Click the button on the top right to open the city modification interface. This interface includes city modification and city positioning. The city modification will conduct a fuzzy search according to the entered city, and automatically select the city that best matches the search city name. The interface is shown in Fig.6, and the selection results are shown in Fig.7. The positioning effect has been tested when entering the software.

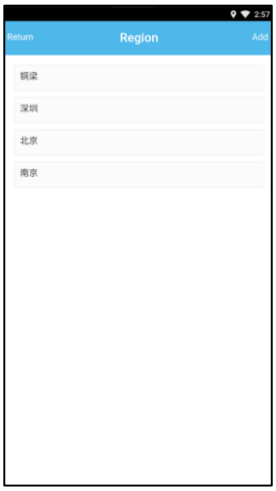
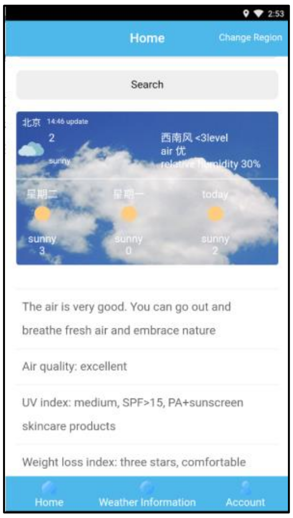

2:57
Return
Region
Add
铜梁
深圳
北京
南京

Fig.6. City change


2:53
Home
Change Region
Search
北京 14:46 update
2
sunny
西南风 <3level
air 优
relative humidity 30%
星期二
星期一
today
sunny
3
sunny
0
sunny
2
Home
Weather Information
Account

The air is very good. You can go out and breathe fresh air and embrace nature

Air quality: excellent

UV index: medium, SPF>15, PA+sunscreen skincare products

Weight loss index: three stars, comfortable

Fig.7. Result page

Click the button on the top right to open the city management interface. This interface contains a list of visited cities. We can complete the city switch by clicking the specified city, or long press the specified city to remove the target city information. After visiting several cities, open the city management interface as shown in Fig.8. After selecting Beijing City, the interface will jump as shown in Fig.9. We can remove some cities that are no longer in common use according to the needs. After long pressing the three cities, the removal results are shown in Fig.10.

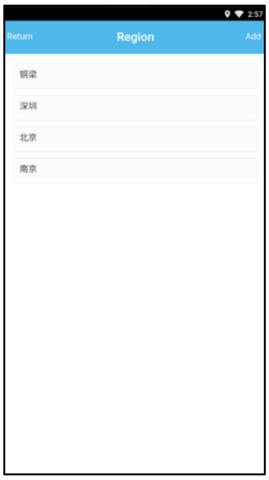
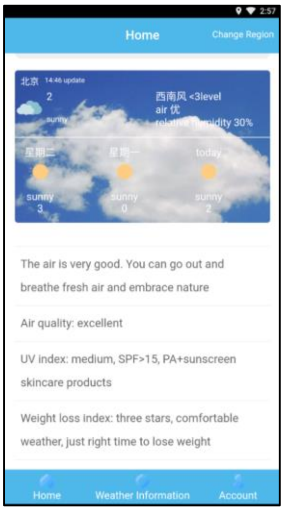

2:57
Return
Region
Add
铜梁
深圳
北京
南京

Fig.8. City management


2:57
Home
Change Region
北京 14:46 update
2
sunny
西南风 <3level
air 优
relative humidity 30%
星期二
星期一
today
sunny
3
sunny
0
sunny
2
Home
Weather Information
Account

The air is very good. You can go out and breathe fresh air and embrace nature

Air quality: excellent

UV index: medium, SPF>15, PA+sunscreen skincare products

Weight loss index: three stars, comfortable weather, just right time to lose weight

Fig.9. Result of picking up

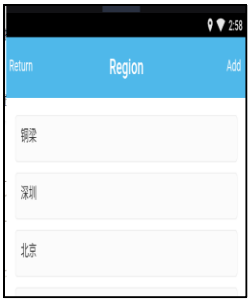

2:58
Return
Region
Add
铜梁
深圳
北京

Fig.10. Result of removing

- Click the button at the bottom right of the software to jump to the historical weather page, which displays the historical weather information of the currently selected city in the form of line chart and table, as shown in Fig.11. Due to the limited size of the mobile terminal equipment, some areas are not enough for display. The software allows users to read the data completely through the sliding screen, as shown in Fig.12, or visually read the data in the horizontal axis direction through the horizontal screen.

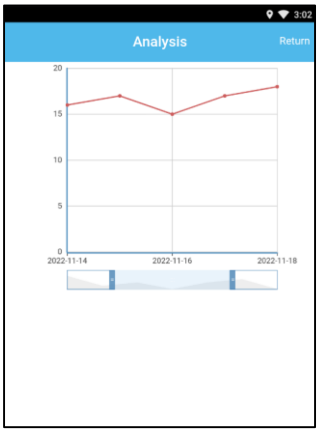
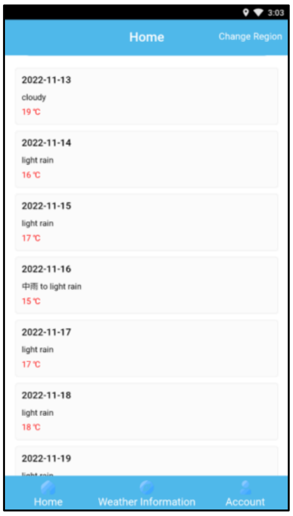


Fig.11. History weather page 1



2022-11-13
cloudy
19 ℃

2022-11-14
light rain
16 ℃

2022-11-15
light rain
17 ℃

2022-11-16
中雨 to light rain
15 ℃

2022-11-17
light rain
17 ℃

2022-11-18
light rain
18 ℃

2022-11-19

Fig.12. History weather page 2

4. Conclusion

Based on cross platform technology, this paper uses Vue+spring boot front and rear end separation framework to finally design a weather forecast application that can truly face the mobile intelligent platform. Through the introduction and comparison of various cross platform technologies, the advantages of Vue+spring boot framework are clearly pointed out. At present, the Vue+spring boot framework is still on the rise, and developers are actively trying to adapt to this front and back separation framework. Due to the limited technology, this article only gives a brief opinion on Vue+spring boot framework and makes a simple practice. In the development process, Vue+spring boot framework fully reflects its efficiency, whether it is convenient to write code or fast to debug programs. It is believed that as technology continues to flourish, developers will build a more excellent and huge ecological community. However, when we are exposed to new technologies, we need to spend twice as much time to learn. Developers who have experienced these technologies understand the extra burden of using new technologies [5].

References

1. Palmieri M , Singh I , Cicchetti A . Comparison of cross-platform mobile development tools[C]. International Conference on Intelligence in Next Generation Networks. IEEE, 2012.
2. Ximing Hu Design and implementation of a high performance remote procedure call framework based on responsiveness [d] Hangzhou University of Electronic Science and technology, 2021
3. Guohong Gao, Xiaohui Guo, Kaining Gao Draw weather forecast app based on Mui
4. Haohand Dong Research on the development prospect of cross platform mobile app based on fluent [j] Information and computer, 2019000 (015): 197-199
5. Okazaki, N. *et al.* (2015) “Security Evaluation System for Android applications using user's reviews and permissions,” *Journal of Robotics, Networking and Artificial Life*, 2(3), p. 190. Available at: https://doi.org/10.2991/jrnal.2015.2.3.12.

Authors Introduction

Mr. Hu Jiahao

He received his bachelor’s from the Institute of Computer Science and Digital Innovation (ICSDI), UCSI University, Malaysia.

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. Thong Chee Ling

She is an Associate Professor at Institute of Computer Science and Digital Innovation, UCSI University. Her research interests include IT Innovation in solving real-life problem. She has received three research grants from UCSI University Malaysia as Principal Investigator for projects entitled ‘Android-based Time Tracker for Shuttle Bus ‘, ‘iOS Time and Location Tracker for Shuttle Bus’ and ‘Mobile Car App for Travel Agency’. These projects are fully copyrighted by Intellectual Property Corporation of Malaysia (MyIPO). In 2021, Thong was awarded the "Product Innovation and Commercialization Award" by the university.

Development of a Novel E-Learning System for Improved Usability

Soo Yang Yew
Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia

Shabana Anjum Shaik, Shayla Islam
Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia

Abstract

E-learning is a learning approach that combines organized instruction with the use of technological resources such as laptops, computers, and tablets. It allows students to study at any time and from any location with the help of Information Communication Technologies (ICT) that can connect instructors and pupils who are separated by thousands of miles. The criticality of this study is to fulfill students' and instructors' expectations of using virtual learning systems while having thrilling and interesting learning and teaching experience. This proposed system will help them to know more about the technology tools functioning and teach them a better understanding of using the system's features. Most lecturers and students faced challenges in utilizing e-learning systems with such factors being reviewed by some researchers.

Keywords: e-learning, ethics, technologies, online learning, pandemic

1. Introduction

E-learning is a learning approach that combines organized instruction with the use of technological resources such as laptops, computers, and tablets. It allows students to study at any time and from any location with the help of Information Communication Technologies (ICT) that can connect instructors and pupils who are separated by thousands of miles. E-learning has provided on improving education's efficacy and efficiency by assisting students' needs such as offering access to updated content, having quick delivery of lessons, and the ease of having unlimited access to the content. Most institutions use tools such as Microsoft Teams, Google meet, Zoom , and so on to impart the e-learning facility to the learners [1].

The critical of this study is to fulfill students' and instructors' expectations of using virtual learning systems while having thrilling and interesting learning and teaching experience [3]. This proposed system will help them to know more about the technology tools functioning and teach them a better understanding of using the system's features. Most lecturers and students faced their challenges in utilizing e-learning systems with such factors being reviewed by some researchers [9],[10],[11].

According to (Wood, 2022), the usage of e-learning has been increasing around the world since the covid-19 pandemic occur [4]. During that period, it was an important instrument for effectively conducting the teaching-learning process [2]. Therefore, institutes of higher learning have been forced to swiftly migrate to distance and online education as the result. Students rarely go back to the usual physical classes to avoid getting infected. Many universities have adapted to investing in online learning as they are interested in finding the most effective ways to distribute course content online, engage students, and perform evaluations [5],[6].

2. Background Study

There are some of the current e-learning systems that have been discussed by past researchers related to novel e-learning systems and the latest technologies.

According to [13], e-learning benefits the education industry in both urban and institutional settings by offering a variety of instruments for students' managerial and instructional needs at the institutional level. the administrative elements, like student enrolment, and the instructional elements, such as making course materials available. An alternative to conventional servers that can better utilize resources is cloud technology as it has been offered as a cutting-edge approach to enabling IT solutions because of its elasticity, speed, scalability, and adaptability capabilities.

E-learning is more than just learning online. It offers solutions that incorporate management and system standardization technologies according to [11]. The researcher provides that different institutions and people are interested in using cloud computing because of its recent emergence. It suggested there are many technologies on the cloud platform that should be fully utilized in the development of cloud-based e-learning. However, cloud computing cannot completely be reliable as some challenges in cloud platforms have been identified.

Based on [12], researchers suggest ensuring that students have access to the same assessment and learning materials, most existing e-learning platforms include traditional e-learning systems. The proposed adaptive e-learning model's integration of Big Data technologies allowed for the consideration of novel techniques and approaches to deliver the most relevant learning materials to each learner while providing an adaptive e-learning approach. It can give additional learning time to students to solve issues for slow learners. However, there are some discussions on related issues in big data on PLP.

3. Problem Statement

There are three problems have been identified as below:

- Usability issues in using e-learning systems among students and instructors

The students may find it difficult to adapt to e-learning immediately due to a lack of IT background and hard to communicate with a lecturer or their friends. They will also feel less interested in learning online without their help. Some studies reviewed the challenges which students faced without technological knowledge [4]. For Instructors who are accustomed to lecturing courses in physical classrooms, they will face similar issues like inexperience in using technological tools such as teaching platform training, inability to get focus on students' learning attention, keep the classroom in order, coordinate classroom discussion, and provide feedback and discussion via the internet [5]. Due to that, there must have some guidance and instruction for them to know how to use the e-learning system and get familiar with it.

- Ethical and trustworthy in using an e-learning system

Most of the time, users do not realize that ethical issues could cause problems while using an e-learning system [7]. For example, cheating during exams would be easily caught when students are in a typical classroom context. however, in e-Learning, the detection methods have not been clearly spelt out. In this case, they are likely to accept academic dishonesty as it is common for them to do so during online learning periods. In addition, they can also copy from other people or unauthorized sources and put it as their work. While they also have no confidence and trust in the e-learning system [13]. Therefore, they must understand the ethics and avoid being caught by exam protocol because of doing this unethical work. Thus, the development of a novel e-learning system enables them to have more confidence for a brighter online learning experience.

- Lack of completeness resources and material in online learning the problem and the solution are depicted in Table 1.

Not all disciplines and subjects are appropriate for digital learning especially STEM subjects such as science, psychology, and engineering where the use of practical knowledge is required. It is impossible to completely adopt online learning in the long term as a student's practical work must be evaluated in person by their lecturers physically. Without the practical, they will feel difficult in understanding their representative course and only rely on learning its theories. As the result, the interest in learning online will be also gradually reduced. Therefore, a digital board must be included, and providing more practical tasks can attract more attention from students during the online learning environment [8],[9],[10],[11].

Table 1 Problem and Solution

Problem	Solution
Usability issues in utilizing an e-learning system among students and instructors	Develop a simplified and user-friendly system that is easy to understand and use.
Ethical and trustworthy in using e-learning system	Educate the students regarding academic dishonesty Implement additional validation features within the system
Lack of completeness resources and material in	Provide some interesting learning resources to have a more enjoyable learning

online learning	experience.

4. Aim and Objective

The project’s purpose is to satisfy the students by providing them with a virtual learning environment through an improvised e-learning system. This will offer them a similar learning experience as face-to-face classes during the hybrid period. Some related issues will be identified and provide suitable solutions for solving them.

There are three project objectives:
- To develop an advanced e-learning system that is convenient and efficient for students and lecturers to use.
- To study existing modern technologies that enable the implementation of a practical environment simulation within the e-learning system.
- To test and evaluate, investigate, and identify the security of the current e-learning system.

5. Methodology

The methodology used in this project to develop an e-learning system is the Prototyping model.

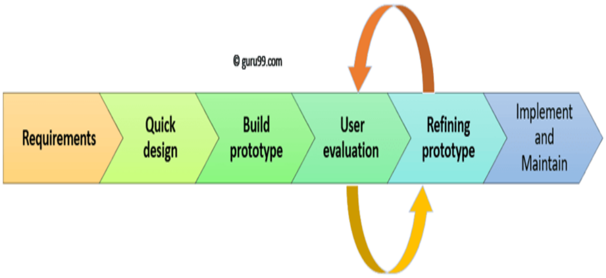


Figure 1: Step of prototyping methodology (Martin, 2022)

Figure 1 show that there are five phases of prototyping and each of them will be described below. Table 2 demonstrate the phases of the approach and deliverables.

1- Requirements gathering and analysis
The requirements of the system are defined in detail. During the process, the users of the system are surveyed and formally interviewed to know what their expectation from the system is.

2- Quick design
In this stage, a simple design of the system is created. However, it is not a complete design. It gives a brief idea of the system to the user. The quick design helps in developing the prototype.

3- Build a Prototype
The information acquired during rapid design is used to create a real prototype in this step. It's a scaled-down version of the needed system.

4- Initial user evaluation
In this stage, the proposed system is presented to the client for an initial evaluation. It helps to find out the strength and weaknesses of the working model.

5- Refining prototype
If the user is dissatisfied with the existing prototype, it must be refined in response to the user's input and ideas.

6. Implementation and Product Maintenance

The final system is tested and deployed to the client.
Any feedback will be considered for further improving the system.

Table 2 Prototyping phases

Prototyping	
Phases	Deliverables
Gather requirements	• Survey results based on what they expected
Design	• A simple idea
Development a Prototype	• Build an e-learning
Initial user evaluation	• Testing the prototype
Refining prototype	• Evaluation of the prototype
Implementation	• Deploy the system and maintain

7. Result Analysis and Synthesis

The analysis is needed to gather the data and to be analyzed to complete the system. There are various ways to gather data such as interviews, questionnaires, and research based on requirement data. The survey is a good way to know the target needs and to collect opinions from users who have used e-learning conveniently. The target audience for this survey is people who are students, and employees who work in various jobs. The survey is released on 2nd October 2022 until 25th October 2022 and a total of 100 students responded. The platform used to conduct the survey is google forms posting the link on social media such as WhatsApp Group and Instagram.
After the survey has been conducted, most of the users benefit by using an e-learning system for attending

classes. Most of the users get the new norm of using it during the technological era. Although some responders will still prefer face-to-face learning, users are very convenient and easy to use the e-learning system even before the pandemic when some of the users have been using it for their own learning purposes. There are some issues that users currently face such as incomplete resources, and a lack of security features. Users believe that using e-learning systems can help them verify their assignment work without having to meet with the lecturer. They will be honest with their work without plagiarizing from another source.

7.1. UML Diagram

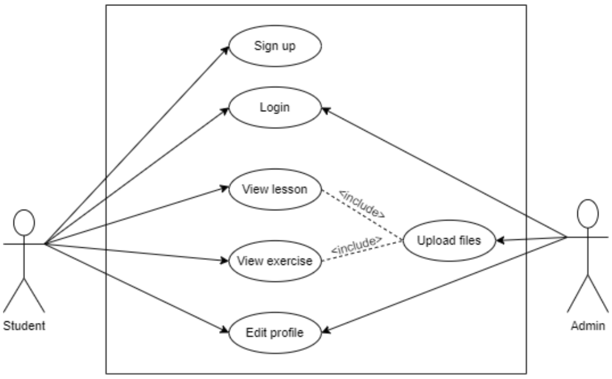


Figure 2 e-learning system use case diagram

A use case diagram is a diagram that describes the interaction between actors and the system. The activities can be carried out by students and the admin when they are using it. Showed in Figure 2, there are two actors who interact with the e-learning system, namely the student and the admin. Students need to sign up by filling in their personal information which includes username and password. Then the registered user can log in to view lessons, and view exercises uploaded by the admin. Admin needs to log in to upload files into view lesson, and view exercises for students to view them. In addition, both admin and students can edit their respective profiles. Only the admin can see their student's full information that has been registered. Students would need to request permission from the admin to see it after making changes.

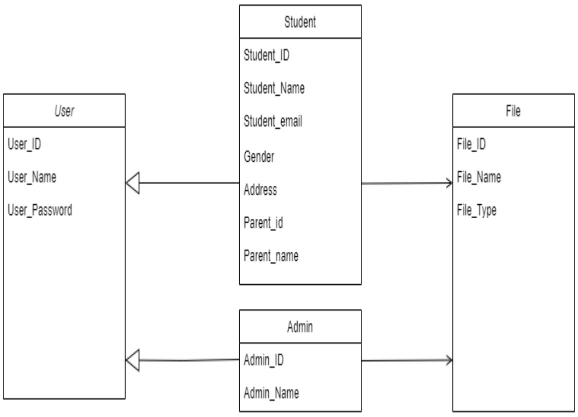


Figure 3 e-learning system class diagram

The class diagram of the e-learning system is illustrated in Figure 3. There are five classes namely student, admin, user, and file. For student's class diagram, it contains the student id, student name, student email, gender, address, parent id, and parent name. Admin's class diagram includes admin id and admin name. Both relate to a user class diagram which includes the user ID, username, and user password. Both also relate to file class diagrams which contain file id file name and file type.

8. Evaluation

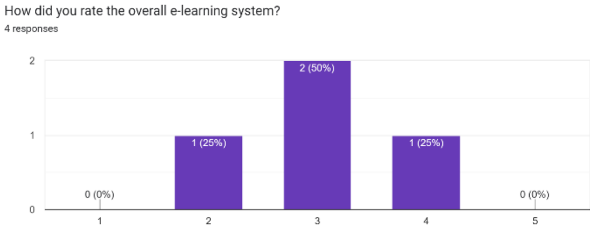


Figure 4 poll of the e-learning system

On a scale of 5, as shown in Freigu 4, two users feel moderate about using the e-learning system while one user feels a bit dissatisfied, and one feels satisfied with using it.

Overall, based on the poll, respondents are satisfied with the usefulness of the e-learning system. However, there are some suggestions given by respondents to develop the system further, for example, by giving view file options. The ideas are noted down and developers will make additional improvements to the system before publishing the system in the future and hope to meet user-expectation

9. Conclusion and Further Work

The project is to give students and lecturers convenient and efficient use of the system as it has features such as adding lessons and exercises for students to download it. Users can make changes to their profiles in the future.

9.1 Main Contribution

The major contribution made to this project is that it has provided file sharing for the lesson and exercise to students. The e-learning system can help benefit students learning in hybrid mode. It can enhance usability that includes adding, editing, viewing, and deleting lessons and exercise files. Students' personal information will be monitored by the admin when the user makes changes to it.

System Limitation
Most of the system development has some disadvantages due to a lack of knowledge during the build prototype phase. This system consists of several disadvantages such as:

Lack of messaging tool
The system lacks a way for users to communicate via messaging.

Unable to upload videos
The system doesn't allow the admin to upload lesson videos for students to view.

System interface
The system design might be simple and least attractive for users due to not having much design implemented.

Future Research
The system is believed to be further enhanced by introducing a video upload function. Next, the system can also be improved by adding a video viewing function. Finally, the system can improve a view function that allows users to view the files before downloading them.

References

1. Kannadhasan, S., Shanmuganantham, M., Nagarajan, R., & Deepa, S. (2020). The Role of Future E-Learning System and Higher Education. International Journal of Advanced Research in Science, Communication and Technology, 12(2), 261-266.
2. Wood, J. (2022, January 27). These 3 charts show the global growth in online learning. World Economic Forum. Retrieved June 13, 2022,
3. Cooke, S. (2021, May 12). How to create a full student experience through Virtual Learning Systems. eLearning Industry. Retrieved June 16, 2022,
4. Alqahtani, A. Y., & Rajkhan, A. A. (2020). E-learning critical success factors during the covid-19 pandemic: A comprehensive analysis of e-learning managerial perspectives. Education sciences, 10(9), 216.
5. Ahmadon, F., Ghazalli, H. I. M., & Rusli, H. M. (2020, December). Studying during pandemic: A review of issues from online learning in the middle of COVID-19. In 2020 6th International Conference on Interactive Digital Media (ICIDM) (pp. 1-4). IEEE.
6. Chang, C. L., & Fang, M. (2020, June). E-Learning and online instructions of higher education during the 2019 novel coronavirus diseases (COVID-19) epidemic. In Journal of Physics: Conference Series (Vol. 1574, No. 1, p. 012166). IOP Publishing.
7. Almseidein, T., & Mahasneh, O. (2020). Awareness of ethical issues when using an e-learning system. International Journal of Advanced Computer Science and Applications, 11(1), 128-131.
8. Burke, L. (2020, October 27). Inside higher ed. Long-term online learning in pandemic may impact students' well-being. Retrieved June 5, 2022.
9. Almaiah, M. A., Al-Khasawneh, A., & Althunibat, A. (2020). Exploring the critical challenges and factors influencing the E-learning system usage during COVID-19 pandemic. Education and information technologies, 25(6), 5261-5280.
10. Aboagye, E., Yawson, J. A., & Appiah, K. N. (2021). COVID-19 and E-learning: The challenges of students in tertiary institutions. Social Education Research, 1-8.
11. Aini, Q., Budiarto, M., Putra, P. O. H., & Rahardja, U. (2020). Exploring e-learning challenges during the global COVID-19 pandemic: A review. Jurnal Sistem Informasi, 16(2), 57-65.
12. Khoury, A., Eddeen, L. N., Saadah, D., & Harfoushi, O. (2011). E-learning: Justifications and obstacles. International Journal of Emerging Technologies in Learning (iJET), 6(3), 53-56.
13. Al-Shehri, M. (2017). Code of Ethics of Teaching-Learning for an e-Learning System. International Journal of Computer Applications, 166(5), 16-20.

Authors Introduction

Mr. Soo Yang Yew

He is currently student at Institute of Computer Science and Digital Innovation (ICSDI), UCSI University, Kuala Lumpur, Malaysia. His research interest includes Mobile and web Application development.

Dr. Shaik Shabana Anjum

She received a B.Eng. degree in Computer Science and engineering and M.Eng. degree (Hons.) in Software Engineering from Anna University, Chennai, India, in 2010 and 2012, respectively, and a Ph.D. degree in Computer Science from the University of Malaya, Malaysia, in 2018. She is an IEEE member. She has also served as a Research Assistant and Post-doctoral research fellow for projects involving traffic congestion and the Internet of Things (IoT) with the Centre for Mobile Cloud Computing Research, Faculty of Computer Science and Information Technology, University of Malaya. She is currently a Lecturer at UCSI University, Malaysia. Her research interests include the IoT, wireless sensor networks, radio frequency identification, ad hoc networks, cognitive radio, and energy harvesting. She has been accoladed with many awards at international competitions and appreciation at the faculty level for her research projects.

Dr. Shayla Islam

She is Currently, Associate Professor in UCSI University, Malaysia. She has completed PhD degree inEngineering from Electrical and Computer Engineering (ECE) departmentat International Islamic University Malaysia (IIUM) in 2016 under Malaysian International Scholarship (MIS). She completed her MSc.at the department of Electrical and Computer Engineering in InternationalIslamic University Malaysia in 2012. She completed her BSc. in ComputerScience and Engineering from International Islamic UniversityChittagong, Bangladesh.

A Development of a Prototype based Mobile Pet Care Application

Gan Ai Leen
Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia

Shabana Anjum Shaik
Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia

Chloe Thong Chee Ling
Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia

Abstract

The keeping of domesticated pets has become a norm in today's society, and taking care of them has become a part of the humane culture. With the rapid advancement of technology in the current era, more people are beginning to utilize it in their everyday life. This project aims to create prototype based mobile pet care application that is able to make the users life easier. The pet care application is based on the android application. Its objective is to provide the users a way to book the necessary appointment through the use of mobile phone without having to do it manually. Additionally, the application will include a chat function which can connect with the users and the admin. The application will enable the pet owners to get timely updates about their pets when they are away and under pet care. It will also provide options for grooming and healthcare facilities for the pet. This project aims at developing an initial prototype of the pet care android application.

Keywords: Mobile Application, Prototype, Pet, Pet Care Application, Android Application

1. Introduction

The 21st century is truly a marvelous era with the introduction of new technologies, especially wireless communication, these technologies have shaped the very way we live and has brought forth many conveniences and benefits to many different sectors of society, economy, countries and many more. The Internet too can be seen as a large-scale information distribution which centers around the many users. In today's world, people of all ages will definitely have at least a mobile phone in hand, and it has played a part in their everyday life. It has become an important factor to the people as they are unable to be separated from it. This change has brought forth an impact for the people in their daily habits due to the advancement in technology. Just like mobile phones, the pet care industry is also playing a part in the everyday life for people who own pets. The pet industry is ever so expanding at a fast rate, this is because the demand in pet population and pet related services is evolving at a very rapid rate and is also emerging in the global market. Pets are a part of our everyday life, since the start of the old century until now, a majority of households will surely own a pet or two. This is because there are benefits to keeping pets in household, studies have shown that owning a pet can increase productivity and lessen stress due to the bond between the people and their pets. What is pet? In Oxford dictionary terms, it is referred to as domesticated animals which are kept in home for the sole purpose of companionship, either for a family or a respective person.[1] Psychologists has proven that people owning pets have a positive impact on their happiness. For example, people who are lonely or just want a companion are told to keep a dog or a cat to quell their loneliness. The general public have viewed the very idea that pets are not only for companionship, but they can also serve as substitutes for psychologist or physicians, as the idea of living with pets have shown to improve mental health, and longevity.[2] In comes the pet care system. Why do we need a pet care system? The pet care application system helps in dealing with problems which involves pet services. It also brings together both the pet owners and the people who provide these services together. The use of the pet care systems allows pet owners to have the ability to choose what kind of services they need, while keeping things like the pricing, proximity, or experience of the pet care service into consideration.[3]. Google is one of the top searching website engines used in the whole world with an estimate of 5.6 billion searches each day. 30% of the people will use Google to book appointments at various places such as restaurants, clinic, movies, etc. According to statistics, 88% of people will research product information before they make their purchase via online.[4]

Pet ownership is increasing every year, and along with it the demand of quality vet service and pet grooming is also increasing. This in turn has encourage the growth of the sector called the Internet of things (IoT) technology. With using IoT, pet owners are able to monitor pets' health, book for services for their pet, or even remotely interact with their pets.[5] IoT has made it possible for any monitoring of pets easy with just a click any mobile application through internet connection.[6]

Technology has advanced till this date and will surely improve in the future. There are multiple pet care developments that offer a variety of pet care apps especially on mobile phones. The apps include various features for the users' daily needs, such as walking and playing meter, health checkups, and video chats. These apps are made to help manage a pet owners time and worries as it can help them to monitor their pet's health and fitness.

The reason why mobile application is preferred than websites is because the users are able to access the app faster due to it being on the mobile phone, whereas for websites, the users need to turn on their laptop/computer. Mobile application also has a lower maintenance cost and developing time compared to websites. The use of the mobile application helps to motivate pet owners to use them to help keep their pets in tip top shape anytime.

2. Literature Review

For one person, to understand the meaning of companionship with animals and its lives plays an important role in the way professional pet care services are organized and well-delivered.[1] Research have shown that pet owner take precautionary step to unsure their pets get well deserved treatment, and as such, they expect veterinarians to do the same.[2] As we know, there are an abundance of websites dedicated for our pets and to get to know their symptoms and treatment, but the information provided and quality of these sites differ greatly among each other. Experts have expressed their concern on whether the consumers are able to discern which is authentic and which is not.[3] Studies have shown that some sites found online gave unreliable or inaccurate source of information regarding the pet symptoms, thus leading to some online consumers having second thought about online pet care systems.[4] According to Yifan Ge, the pet industry is an ever-growing industry which seeks to develop in this line of market should not only be good at discovering and understand what customers need, but also try to understand from the aspect of product quality, sales channel, price and promotion. If one wants to succeed in the pet line industry, product quality would be the utmost importance. Not only that, but they must also grasp the importance of improving it too.[5] It has been pointed out of the advantages of the new era of communication infrastructure, where the Internet has provided a form of threshold to communicate with nearly everyone worldwide. Whereas in the real world, it is limited to the distance and time.[6] The pet care service is one of those web-of-object (WOA) architecture service which provides a real-time and adaptive service control based on the environment change. It adopts a service overlay network concept for providing new services.[7]

The is where the application of the Internet of Things (IoT) comes in. IoT is vast, and yet, also versatile to use in almost everything. The capability it possesses is the connection of various gadgets in one network, especially of the use of the mobile phone, via the Internet, through sensors, are making remarkable formula to achieve the goal of smart recognition, position tracking, monitoring, and many more. [8],[9],[10] With the use of these technology, difficult task such as the caring of animals can be automated through the help of IoT.[11] The pet care application is a type of mobile app which specializes in pet related service. Its sole purpose is to bring together pet owners and the people who work in the service together. The app gives the owners the ability to choose which services they need and the time they want into consideration. [12],[13]

During the Covid-19 pandemic, it was a difficult situation for pet owners and the staff to come face-to-face to address the health issue of the pet. In the year 2019, a digital health survey which was conducted by

Accenture suggested that nearly half of the people aged between 22 to 38 years old only used routine healthcare services. They desire a medical service which is far more convenient and easier to access than the usual healthcare which was used by the previous generation. For example, they hope to have easy access to medical services through the means of online or through the use of mobile applications (apps).[14] This can also be applied to pet related healthcare.

It is preferable under the development of the pet care application to include a chat function. The application is based on real time and the objective of the chat function is to create an online platform which can connect the pet owners and the staff for them to chat in regards to checking the status of their pet or so on.[15]

Table 1: Comparison Table

Mobile Applications	JomPaw	Rover	Meowtel	Prototype mobile application (Pets App)
Pet Boarding/Pet Hotel	✓	✓	✗	✓
Pet Salon	✗	✗	✗	✓
Pet Taxi	✓	✗	✗	✓
Vet Service	✗	✗	✗	✓
Pet Tips	✗	✗	✗	✓
Chat & message function	✓	✓	✓	✓

2.1. *Similarities between these mobile applications*

Based on the comparison table above which is Table 1, the similarities that all these app have are the chat & message function. In order to have a conversation between the pet owner and the staff, this function is a definite must in all pet care system.

2.2. *Differences between these mobile applications*

Most of the applications above do not provide the services such as veterinarian care, and also pet tips. The pet tips should be included in all pet care system in the case where an emergency happen, the pet owners will know what to do before seeking professional help. As for the vet services, this service is also important in all pet care as the pet owners do not have to find other places to treat their pets.

2.3. *Pets App Mobile Application*

The proposed prototype mobile application shown in the comparison table above which is Table 1 will feature services and added feature which the system does not have. Pet Paradise mobile application differs from most pet applications is that to be able to use it, the user will have to be a member. It is free to sign up and they would be able to use most of the features on the app.

3. Problem statement

3.1. *Problem 1: Limited options for existing pet care system*

The flaws for the existing mobile application include less options for pet owners to choose. Pet owners would like have an app which has many services for them to choose. Enhancement for the proposed mobile application will include more options to showcase an all-in-one app.

3.2. *Problem 2: Complex User Interface (UI)*

Mobile application may be one of the best options for quick services, but some of the existing mobile application's UI design is complex, and thus lead to the users' dissatisfaction. Mobile applications UI design has to be made simple and easy to understand for people of all ages, especially towards old people as their learning process is slow due to old age. And the input controls for the UI have to be easy to search since users are usually pressed for time.

3.3. *Problem 3: No chat function*

Most pet application do not come with a chat function. Having the difficulty of being at the place physically, pet owners would most definitely prefer to have a messaging function that can connect them to the staff.

This is to ensure that they can chat or be kept updated on their pets if they were to worry about them.

4. Methodology

The objective is to study the existing pet care mobile applications that is available on the app store, by doing this, it can help to identify both the strength and weakness of the current pet care system. Once the objective is found, the prototype pet care mobile application will be designed and developed which will include extensive options to all pet owners. Once the development of the prototype system is done, evaluation on the performance of the proposed prototype mobile application will be done by conducting testing to get the users perception.

The major research problem that developers aim to solve is by collecting data from users to identify potential users who are more likely to use which features in the system, and to seek comments and suggestions, which can be crucial for the development of the system. Questionnaire was chosen because it allows faster collection of data from larger data sets. Hence, this can be used to help developers study the user's needs and wants. The survey questionnaire form will be sent to random UCSI students from various background in order to collect an objective data for the prototype application development.

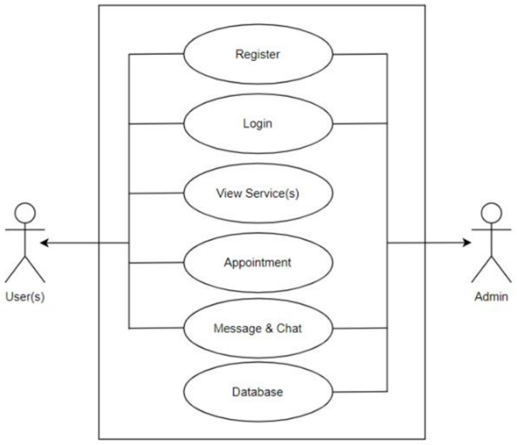
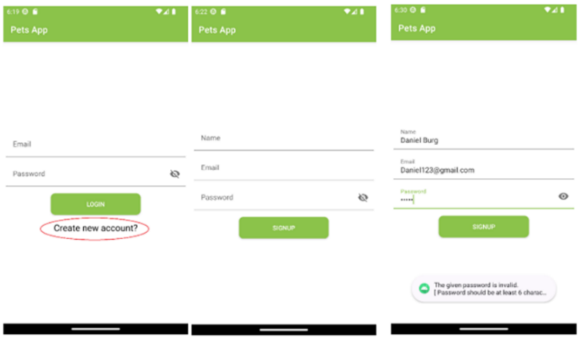


Figure 1: Use Case Diagram

The use case diagram shown in Figure 1 shows two actors, the users and the admin. The users are able to access the mobile application to register as a member. Once the registration is successful, they will be able to log into the application, which they will then process to view the services provide on the mobile app. The users are also able to book an appointment and chat with the staff, which are the admin. The same goes for the admin, they are able to registers and login into the system but only on the admin side. The admin can also chat with the users, the only difference is that admin has database access to view all information that is kept in regards to the mobile app.

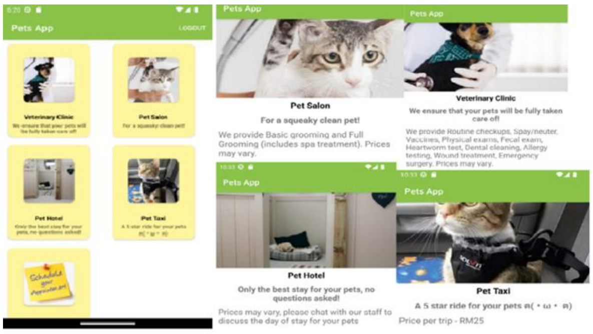


Name
Daniel Burg
Email
Daniel123@gmail.com
Password
•••••

The given password is invalid.
[Password should be at least 6 charac...

Figure 2: Registration of pet owners to use the pet care mobile application

In Figure 2, it shows where the users can click on the create a new account, once the user has clicked on it, it will lead them to a new page for them to input their name, email, and their preferred password. Once they have filled up the necessary form, the user can click on the signup button and they will be a registered member of the mobile app. It also shows what will happen when a new user includes a password of less than 6 characters. A notification will pop up at the bottom of the screen indicating that it is an invalid password.



Pet Salon

For a squeaky clean pet!

We provide Basic grooming and Full Grooming (includes spa treatment). Prices may vary.

Veterinary Clinic

We ensure that your pets will be fully taken care of!

We provide Routine checkups, Spay/neuter, Vaccines, Physical exams, Fecal exam, Heartworm test, Dental cleaning, Allergy testing, Wound treatment, Emergency surgery. Prices may vary.

Pet Hotel

Only the best stay for your pets, no questions asked!

Prices may vary, please chat with our staff to discuss the day of stay for your pets

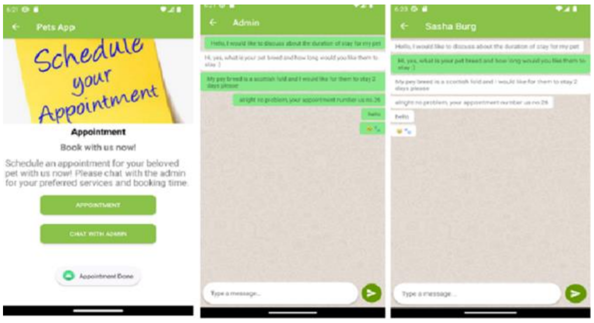
Pet Taxi

A 5 star ride for your pets ฅ(• ω • ฅ)

Price per trip - RM25

Figure 3: Homepage and pet services option after a successful registration

Once the users have successfully logged in, it will lead them to the homepage which is shown in Figure 3, and shows the users the various services that are available for them.



Appointment

Book with us now!

Schedule an appointment for your beloved pet with us now! Please chat with the admin for your preferred services and booking time.

Hello, I would like to discuss about the duration of stay for my pet

Hi, yes, what is your pet breed and how long would you like them to stay :)

My pey breed is a scottish fold and I would like for them to stay 2 days please

alright no problem, your appointment number us no.26

hello

Hello, I would like to discuss about the duration of stay for my pet

Hi, yes, what is your pet breed and how long would you like them to stay :)

My pey breed is a scottish fold and I would like for them to stay 2 days please

alright no problem, your appointment number us no.26

hello

Figure 4: The active chat application to converse with the admin

In Figure 4, it shows the button which is used to chat with admin in regards to discussing which and what services, and also the booking time that the user wishes to book. It also shows the perspective of the chat function for both the user and the admin, on the left is where the user chats with the admin while the right is vice versa.

5. Result

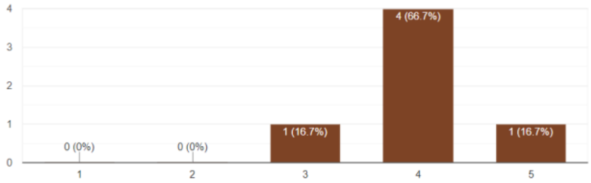


Figure 5: Graph of the prototype testing survey

Based on the Figure 5, it shows the responses from the prototype testing survey, out of all 6 users, most of the participants have a good experience when using the prototype system. While the majority of the respondents have shown positivity on the prototype app, we can conclude that because it is a prototype app, improvements can still be made to it to ensure it can become a full-fledged, working app before it can be released for public use.

6. Conclusion

The development of modern technology is rapidly changing with each passing day, Nowadays, mobile phones are no longer restricted to just making phone calls or sending messages. In the 5G era that we are now living in, smart phones have now been integrated into everyone's life, no matter the age group.

Mobile phones have brought forth many conveniences throughout our life. It has allowed us to communicate at an efficient rate, which includes online chat and video call, streaming videos, or the commonly used mobile applications that has made the life of people easier. It is believed that in the near future, website usage will be phased out slowly, as mobile technology is growing stronger as the years go by.

As of now, the current prototype system will definitely need to undergo changes in the future to be a full-fledged, usable mobile application. Appointment will need to undergo changes and add a form with calendar or clock feature for ease of booking. Improvement to the user interface will also be a part of the future development.

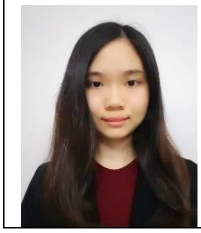
References

1. Pet. (2006). In Oxford Word power Dictionary (3rd ed). Oxford: Oxford University Press
2. Herzog, H. (2011). The impact of pets on human health and psychological well-being: fact, fiction, or hypothesis? Current directions in psychological science, 20(4), 236-239.
3. Mahamuni, S. R. (2021). Pet Care Web Application (Doctoral dissertation, CALIFORNIA STATE UNIVERSITY, NORTHRIDGE).
4. https://www.pymnts.com/news/retail/2018/omichannel-ecommerce-consumer-habits/.
5. Chen, Y., & Elshakankiri, M. (2020). Implementation of an IoT based pet care system. In 2020 Fifth International Conference on Fog and Mobile Edge Computing (FMEC) (pp. 256-262). IEEE.
6. Aguirre, E., Lopez-Iturri, P., Azpilicueta, L., Astrain, J.J., Villadangos, J., Santesteban, D. and Falcone, F. (2016). Implementation and analysis of a wireless sensor networkbased pet location monitoring system for domestic scenarios. Sensors, 16(9), p.1384.
7. Christine Morley & Jan Fook Prof (2005). The importance of pet loss and some implications for services, Mortality, 10:2, 127-143,
8. Jason B. Coe, DVM; Cindy L. Adams, MSW, PhD; Brenda N. Bonnett, DVM, (2007). PhD, Journal of the American Veterinary Medical Association, Vol. 231, No. 10, Pages 1510-1518
9. Eysenbach G, Kohler C. (2002). How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. BMJ.
10. Murray E, Lo B, Pollack L, et al. (2003). The impact of health information on the internet on the physician-patient relationship: patient perceptions. Arch Intern Med.
11. Ge, Y., Thongrawd, C., & Duangthongsuk, W. (2022). ORBITZ HOUSE" PET CARE ON-LINE STORE.
12. Schroeder, Carolin. (2015). Through space and time: Mobile apps in Urban Development.
13. Lee, N., Lee, H., & Lee, H. (2016). Things-aware smart pet-caring system with internet of things on web of object architecture. In 2016 International Conference on Information and Communication Technology Convergence (ICTC) (pp. 1247-1252). IEEE.

14. Patel, K.K. and Patel, S.M. (2016). Internet of things-IOT: definition, characteristics, architecture, enabling technologies, application & future challenges. International Journal of Engineering Science and Computing, 6(5).
15. Liyanage, W. L. S. V., & Wedasinghe, N. (2021). Implementation of Smart Pet Care Applications in an IoT Based Environment.
16. Cheng, Y. H. (2019). A development architecture for the intelligent animal care and management system based on the internet of things and artificial intelligence. In 2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIC) (pp. 078-081). IEEE.
17. Mahamuni, S. R. (2021). Pet Care Web Application (Doctoral dissertation, CALIFORNIA STATE UNIVERSITY, NORTHRIDGE).
18. Vanshri, S., Veena, P., Priyanka, D., & Priyanka, G. (2018). Pet Care System based on android Application. International Journal for Research in Applied Science and Engineering Technology, 6(3), 1915-1919.
19. Belton, P. (2019). Would you be happy to see your doctor online? BBC News.
20. Mohammed, M. A., Bright, A. S., Apostolic, C., Ashigbe, F. D., & Somuah, C. (2017). Mobile-Based Medical Health Application-Medi-Chat App. Int. J. Sci. Technol. Res, 6(5), 70-76.

Authors Introduction

Ms. Gan Ai Leen



She is currently student at Institute of Computer Science and Digital Innovation (ICSDI) UCSI University, Kuala Lumpur, Malaysia. Her research interest includes Mobile and web Application development.

Dr. Shaik Shabana Anjum

She received a B.Eng. degree in Computer Science and engineering and M.Eng. degree (Hons.) in Software Engineering from Anna University, Chennai, India, in 2010 and 2012, respectively, and a Ph.D. degree in Computer Science from the University of Malaya, Malaysia, in 2018. She is currently a Lecturer at UCSI University, Malaysia. Her research interests include the IoT, wireless sensor networks, radio frequency identification, ad hoc networks, cognitive radio, and energy harvesting. She has been accoladed with many awards at international competitions and appreciation at the faculty level for her research projects.

Dr. Chloe Thong Chee Ling

She received her PhD in Information Systems, from Universiti Putra Malaysia in 2015. She is currently an Associate Professor in UCSI University, Malaysia. Her research interest include Applied Coputing, Information Management and Digital Inoovation.

Modeling of an Environmentally Independent and Contactless Speed Sensor for Measuring the Speed of Ships, Submarines, and Aircraft in Relation to the Ground Development of Image

Jakaria Mahdi Imam, Mohammad Aminul Islam, Norrima Mokhtar, S. F. W. Muhammad Hatta
Department of Electrical Engineering, Faculty of Engineering, 50603 University of Malaya, Malaysia

Heshalini Rajagopal
Institute of Computer Science and Digital Innovation, UCSI University, 56000 Kuala Lumpur, Malaysia

E-mail: s2133586@um.edu.my, aminul.islam@um.edu.my, norrimamokhtar@um.edu.my ,sh_fatmadiana@um.edu.my

Abstract

We are presenting a theoretical and mathematical model for an environment nondependent contactless speed sensor which can measure directly the horizontal speed of ships, submarines, and aircraft with respect to the ground. Currently, available standalone onboard speed sensors used in ships, submarines, and aircraft measure the speed of the vehicle with respect to the water or air; not ground. Thus, they are unable to measure the ground speed of these vehicles directly. In this paper, we have shown that a novel speed sensor can be designed by using small size dropped inside a vacuum chamber. The theoretical and mathematical model of the proposed sensor was further validated by simulation. The simulation results showed that unlike the conventional methods our proposed method can measure the speed of ships, submarines, and aircraft with respect to the ground directly. Another issue is that current speed sensors are environment-dependent, meaning they or their probes require some type of touch with the operational environment. This environment dependency further affects the accuracy of these sensors, as well as lays these sensors or their probes at risk of collision with external objects. However, the requirement to mount the sensor or its probes outside the vehicle is removed by our suggested sensor. The complete sensor assembly can be placed inside the vehicle.

Keywords: 3D Pose Estimation, Accelerometer, Computer Vision, Free-Fall Ball, Gyro, Inertial Measurement Unit (IMU), Inertial Navigation System (INS), Image Processing, Simultaneous Localization And Mapping (SLAM), Speedometer

1. Introduction

At present ships and submarines use Electromagnetic Log (EM Log), Pitometer Log, and other types of speed sensors for their speed measurement[1][2][3]. But all these sensors measure the speed of the vessel with respect to water, not ground[4]. As a result due to tide or current, vessels can’t measure their speed with respect to the ground directly[4]. For example, if a ship is anchored at sea under a strong current, its speed sensor shall show that it is moving at high speed[5], though, in reality, it is standing still with respect to the ground[5]. Similarly, aircraft use the pilot tube for their speed measurement with respect to air, not ground[6], [7]. As a result due to air, aircraft can’t measure their speed with respect to the ground directly[8][9][10][11]. For instance, if a helicopter is hovering at zero speed under the heavy wind, its speed sensor shall show that it is moving at high speed[11]. Though in reality, it is standing still with respect to the ground[11].

Fig.1 shows the current method used by ships for ground speed measurement. At present ships measure their ground speed by calculating the EM Log speed information, propeller speed information, wind sensor information, and the Tide Table information[5].

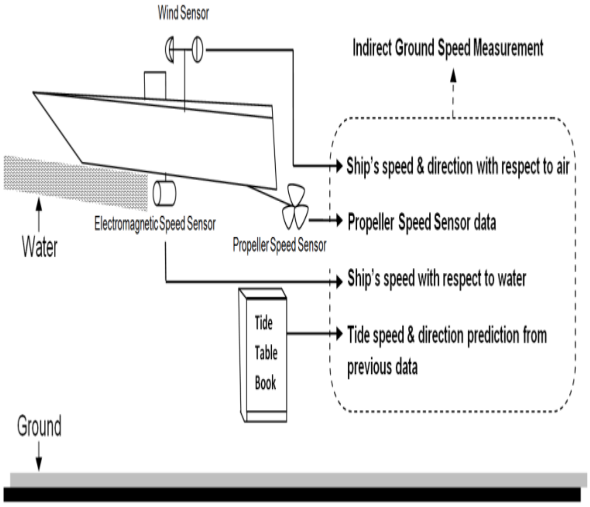

Wind Sensor
Indirect Ground Speed Measurement
Ship's speed & direction with respect to air
Propeller Speed Sensor data
Electromagnetic Speed Sensor
Propeller Speed Sensor
Water
Ship's speed with respect to water
Tide
Table
Book
Tide speed & direction prediction from previous data
Ground

Fig.1.Current method used by ships for ground speed measurement

However, because of faults in the EM Log and Tide Table, this technique yields significant errors in the ground speed measurement[1]. Same way, in the absence of GPS and external sources an aircraft has to measure its ground speed by using air speed, altitude, pressure, density, and other information collected from various onboard sensors. This will also introduce large errors in ground speed measurement due to inaccuracies of input parameters. Though measurement of speed with respect to ground is an important aspect of navigation for water and air vehicles, due to the above reasons ships, submarines, and aircraft cannot measure their ground speed directly. Moreover, currently available speed sensors used in water and air vehicles are environment-dependent; which means these sensors or their probes need some sort of contact with the operating environment. As a result operating environment parameters such as temperature, density, salinity, chemical properties, etc. further affect the accuracy of these sensors. Besides as these sensors or their probes need to be placed outside the vehicle, it initiates further risk of collision with external objects.

GPS is used for speed measurement with respect to the ground, but GPS is not a standalone solution for ships and aircraft as it is susceptible to jamming and spoofing[12][13]. On the other hand, GPS doesn't work when submarines operate underwater[14]. In rare cases, accelerometers are used by vehicles for ground speed measurement[15][16]. In such cases, the inaccuracy in ground speed measurement is increased over time because of numerous sorts of bias, noise, and errors in accelerometer sensors[17]. As the error increases at a large amount with time, they are not used alone; rather they are used with other sensors for measuring ground speed[15][16][17]. To date, no environment non-dependent contactless standalone onboard speed sensor is known to exist, which can measure the speed of water and air vehicles directly with respect to the ground. Hence we proposed a new method of an environment non-dependent contactless standalone onboard speed sensor (just like a standalone Gyro Sensor or Accelerometer Sensor) which can measure the speed of water and air vehicles with respect to the ground directly; without knowing other information such as speed of current, RPM of the propeller, etc.

In this paper, we proposed a new approach for the theoretical and mathematical model of a contactless speed sensor which can measure the speed of the ships, submarines, and aircraft with respect to the ground directly. Another important focus of our approach was to completely eliminate the environmental dependency. The theoretical and mathematical model of the proposed sensor was further validated by simulation.

2. Materials and Methods

2.1. *Theoretical Background*

In 1638, Galileo established the principle of compound motion[18]. According to this principle, in projectile motion the horizontal motion (velocity in x-axis and y-axis) and the vertical motion (velocity in z-axis) are independent of each other; that is, neither motion affects the other. For better understanding let us assume at earth an object is at free fall in an airless environment. According to Galileo the three-axis velocity profile of the object is shown in Table 1[18].

Table 1. Properties of the free-falling object in the airless environment

Time (Sec)	Velocity in x-Axis (m/s)	Velocity in y-Axis (m/s)	Velocity in z-Axis (m/s)
0.1	5	10	0.98
0.2	5	10	1.96
0.3	5	10	2.94
0.4	5	10	3.92
0.5	5	10	4.9
0.6	5	10	5.88
0.7	5	10	6.86
0.8	5	10	7.84
0.9	5	10	8.82
1.0	5	10	9.8

Table 1 shows that if there is no other external resistance on a free-fall object, gravity does not affect the horizontal velocity (x-axis and y-axis velocity) and they remain unchanged over time. Only the vertical velocity (z-axis velocity) is changed over time due to gravity. It means that if we can measure the x-axis and y-axis velocity of the object only once, we know that these two axis velocities shall remain constant until the object touches the ground. This understanding shows the possibility of developing a new sensor by arranging small size free falling balls inside a vacuum chamber for direct measurement of the horizontal velocity (along the x and y-axis) of a vehicle with respect to ground.

2.2. *Description of Device/ Sensor*

Previously it is discussed and shown that currently available speed sensors used in water and air vehicles cannot measure ground speed directly and they need some sort of contact with the operating environment. They need propeller speed information, tide table information, and other information for indirect ground speed measurement Fig.1. Unlike these sensors our proposed speed sensor can measure ground speed directly without any other extra information like propeller speed information, tide table information, etc. and it does not require any sort of contact with the operating environment. It can be placed at any suitable place inside the vehicle Fig.2.

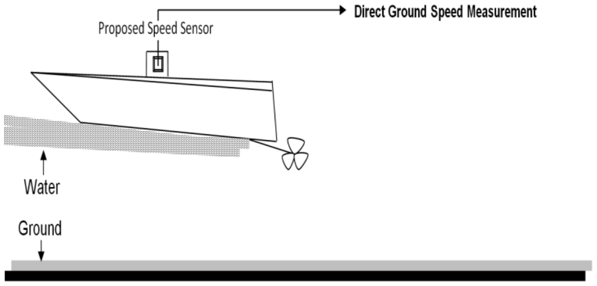


Fig.2.Proposed contactless speed sensor placed inside a ship for direct measurement of ground speed

Fig.3 shows the basic configuration of the proposed speed sensor. The proposed speed sensor consists of an enclosed vacuum chamber. A continuous free-fall ball dropping mechanism is arranged inside this vacuum chamber. Inside the vacuum chamber the balls fall from the top in a single line. Once they reach the bottom of the vacuum chamber, they are again sent to the top. The arrangement is done in such a way that at least 2 balls shall always be at free fall. A 3D Range Finder is placed at the perimeter of the vacuum chamber. The 3D Range Finder is capable of determining the 3D distance of at least 2 free-falling balls at the same time.

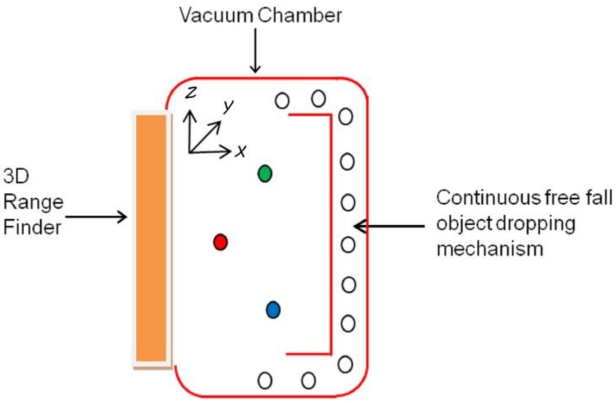


Fig.3.Basic configuration of the sensor

2.3. *Operating Principle*

For better understanding firstly a single free-falling ball is discussed using Fig.4. Subsequently, multiple free-falling balls are discussed using Fig.5.

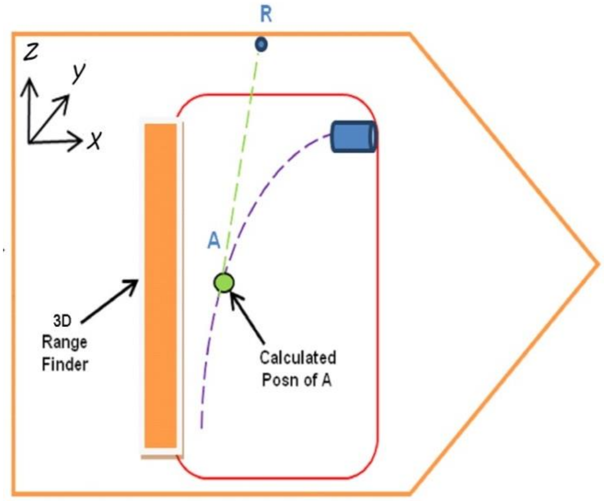


Fig.4.Sensor assembly placed inside a vehicle

During initial set up the sensor assembly is fixed rigidly inside the vehicle. Then it is aligned with the vehicle in all 3 axes. To get the system to work we need to know the 3 axis orientation of the vehicle with respect to earth, Gravity model and the relative distance between the falling ball and the vehicle at all times.

To measure the 3-axis global orientation of the vehicle a 3-axis gyro may be used. To measure the gravity model a 3-axis accelerometer may be used. To measure the relative distance between the falling ball and the vehicle (point R) a LIDAR, RADAR, SONAR, Stereo Camera, Time of Flight Camera, Kinect, or any other suitable device referred to as 3D Range Finder may be used. When the sensor starts operating, the first ball A starts falling. During free-fall ball A moves at constant velocities along x and y-axes as described in Table 1 with respect to ground until it touches the vacuum chamber. At the beginning of the operation at time t0, we need to know the x and y-axis velocities of falling ball A with respect to ground from external sources (like GPS) only for once. For subsequent operations, the external source shall not be required anymore.

From the understanding of Table 1 once we get the 2 axes' horizontal velocity of the falling ball A at any instance t0, afterward at any moment we can predict its 2 axes' horizontal velocity with respect to ground by simple mathematics without the necessity of any external source/sensor. As a result, the falling ball A can now be used as a reference for speed with respect to ground.

For better understanding let us assume, that just after t0 the vehicle started to move at various speeds and directions. Now if we can measure the relative distance between the falling ball A and the vehicle at 2 different times, then using the falling ball as a reference for speed, we can find out the 2 axes' average horizontal velocity of the vehicle in between these times.

At the end of the free fall when ball A touches the vacuum chamber, it can no longer be used as a reference for speed. But the referencing system must go on for measuring the velocity of the vehicle continuously. To solve this problem a continuous loop of multiple dropping balls (similar to ball A) in a single line is introduced which is shown in Fig.5.

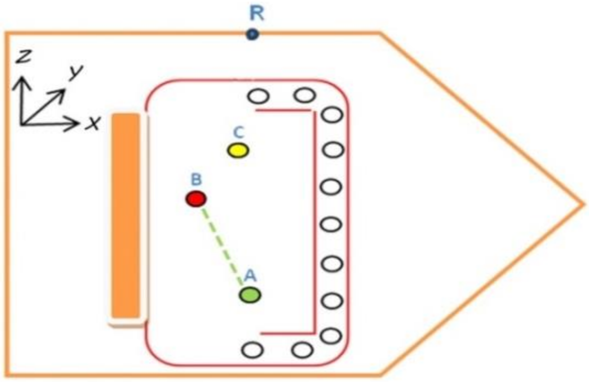


Fig.5.Operation of multiple free-falling balls inside the sensor

The loop shall work in a way so that at least 2 balls shall always be at free fall. As a result after ball A, the next ball B shall start falling. Then the reference shall be shifted from ball A to ball B for determining the 2-axis horizontal velocity of the vehicle. This reference shifting must be done before ball A touches the bottom. Ball B qualifies to be a reference once we find out its 2-axis horizontal velocity. The method of finding out the 2-axis horizontal velocity of ball B is almost similar to the method we found out the 2-axis horizontal velocity of the vehicle. To do this we need to measure the 2 axes (horizontal) relative distance of the falling ball B from falling ball A at 2 different times using the same 3D range finder. Then using the falling ball A as a reference, we can find out the 2-axis horizontal velocity of falling ball B.

Once we know the 2-axis horizontal velocity of ball B, it can be used as a reference instead of ball A. Now ball B can be used in the same manner as ball A for determining the 2 axes' average horizontal velocity of the vehicle. In the same way, the reference shall be subsequently shifted to balls C, D, E, F, G, H, and so on. Fig.6 shows the basic operating principle of the speed sensor in the flowchart.

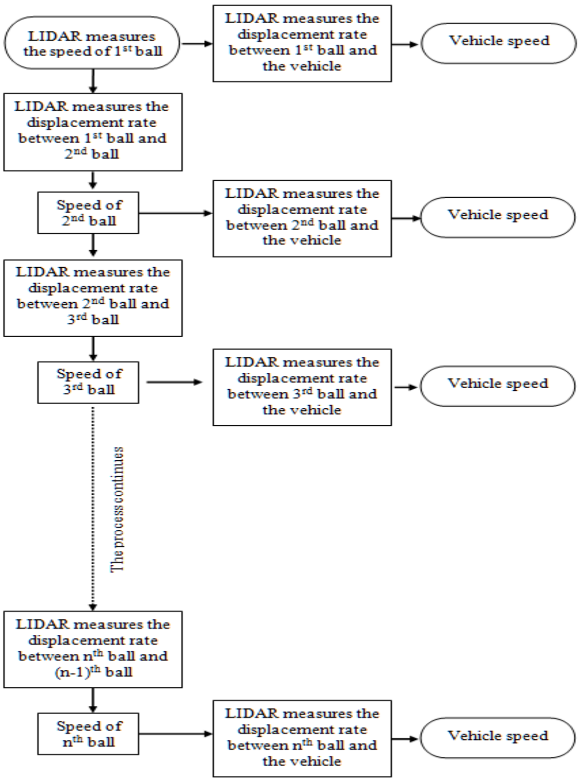

LIDAR measures the speed of 1st ball
LIDAR measures the displacement rate between 1st ball and the vehicle
Vehicle speed
LIDAR measures the displacement rate between 1st ball and 2nd ball
Speed of 2nd ball
LIDAR measures the displacement rate between 2nd ball and the vehicle
Vehicle speed
LIDAR measures the displacement rate between 2nd ball and 3rd ball
Speed of 3rd ball
LIDAR measures the displacement rate between 3rd ball and the vehicle
Vehicle speed
The process continues
LIDAR measures the displacement rate between nth ball and (n-1)th ball
Speed of nth ball
LIDAR measures the displacement rate between nth ball and the vehicle
Vehicle speed

Fig.6.Operating principle of the proposed speed sensor in flowchart

2.4. *Mathematical Modeling*

Table 2 shows the notation used in this work.

Table 2. Notations used in this work

Symbol	Meaning
T	Time (second)
v_{M_x}	Velocity of the current ref ball on x-axis (m/s)
v_{M_y}	Velocity of the current ref ball on y-axis (m/s)
d_{RM_x}	Distance between current ref ball and vehicle ref point R on x-axis (m)
d_{RM_y}	Distance between current ref ball and vehicle ref point R on y-axis (m)
v_{V_x}	Velocity of the vehicle on x-axis (m/s)
v_{V_y}	Velocity of the vehicle on y-axis (m/s)
d_{NM_x}	Distance between current ref ball and next ball on x-axis (m)
d_{NM_y}	Distance between current ref ball and next ball on y-axis (m)
v_{N_x}	Velocity of the next ref ball on x-axis (m/s)
v_{N_y}	Velocity of the next ref ball on y-axis (m/s)

Initially, we know the 2-axis horizontal velocity of the falling ball A. Let us assume we measured its velocity along the x and y-axis as $v_{A_{xt0}}$ and $v_{A_{yt0}}$ respectively at time t0. As long as A doesn't touch the bottom, afterward at any time t1 its horizontal velocity can be predicted along the x and y-axis as:

$$v_{A_{xt1}} = v_{A_{xt0}} \tag{1}$$

$$v_{A_{yt1}} = v_{A_{yt0}} \tag{2}$$

In the same way at time t2 , 2 axis horizontal velocity of the falling ball A can be predicted as:

$$v_{A_{xt2}} = v_{A_{xt1}} = v_{A_{xt0}} \tag{3}$$

$$v_{A_{yt2}} = v_{A_{yt1}} = v_{A_{yt0}} \tag{4}$$

Now we need to measure the relative distance of the falling ball A from the vehicle reference point R on the horizontal 2 axes at 2 different times t1 and t2.

Let us assume, we could measure the horizontal 2 axis relative distance of the falling ball A from the vehicle reference point R in the horizontal 2 axes at time t1 and t2 as $d_{RA_{xt1}}$, $d_{RA_{yt1}}$ and $d_{RA_{xt2}}$, $d_{RA_{yt2}}$ respectively. So, in-between time t1 and t2 the 2 axes' average horizontal velocity of the vehicle can be calculated as:

$$v_{V_{x(t1\,\&\,t2)}} = v_{A_{xt2}} - \{d_{RA_{xt2}} - d_{RA_{xt1}}) / (t_2 - t_1)\} \tag{5}$$

$$v_{V_{y(t1\,\&\,t2)}} = v_{A_{yt2}} - \{d_{RA_{yt2}} - d_{RA_{yt1}}) / (t_2 - t_1)\} \tag{6}$$

In the same way we can continue to find out the 2 axis average horizontal velocity of the vehicle in between the times t2 & t3 , t3 & t4 , t4 & t5 , ... , tn-1 & tn and the above equations can be simplified as:

$$v_{A_{xtn}} = v_{A_{xt(n-1)}} \tag{7}$$

$$v_{A_{ytn}} = v_{A_{yt(n-1)}} \tag{8}$$

$$v_{V_{x(tn-1\,\&\,tn)}} = v_{A_{xtn}} - \{(d_{RA_{xtn}} - d_{RA_{xtn-1}})/(t_n - t_{n-1})\} \tag{9}$$

$$v_{V_{y(tn-1\,\&\,tn)}} = v_{A_{xtn}} - \{(d_{RA_{ytn}} - d_{RA_{ytn-1}})/(t_n - t_{n-1})\} \tag{10}$$

Let us assume later at time t6 the next ball B appeared. We measured the 2 axes (horizontal) relative distance of the falling ball B from falling ball A at time t6 as $d_{BA_{xt6}}$, $d_{BA_{yt6}}$ and at time t7 as $d_{BA_{xt7}}$, $d_{BA_{yt7}}$respectively. So, at time t7 the 2 axis horizontal velocity of falling ball B can be calculated as:

$$v_{B_{xt7}} = v_{A_{xt6}} + \{(d_{BA_{xt7}} - d_{BA_{xt6}})/(t_7 - t_6)\} \tag{11}$$

$$v_{B_{yt7}} = v_{A_{yt6}} + \{(d_{BA_{yt7}} - d_{BA_{yt6}})/(t_7 - t_6)\} \tag{12}$$

The above equations can be simplified as:

$$v_{B_{xtn}} = v_{A_{xtn-1}} + \{ (d_{BA_{xtn}} - d_{BA_{xtn-1}})/(t_n - t_{n-1})\} \tag{13}$$

$$v_{B_{ytn}} = v_{A_{ytn-1}} + \{ (d_{BA_{ytn}} - d_{BA_{ytn-1}})/(t_n - t_{n-1})\} \tag{14}$$

Once we know the 2-axis horizontal velocity of ball B, it can be used as a reference for speed instead of ball A. In the same way, the reference shall be subsequently shifted to balls C, D, E, F, G, H, and so on. Now, if we consider the current reference ball as M and the next reference ball as N, then for a continuous falling ball system, we need to find out the 2-axis horizontal velocity of the next ball from the current ball. In that case, the reference shifting equations can be expressed as:

$$v_{N_{xtn}} = v_{M_{xtn-1}} + \{(d_{NM_{xtn}} - d_{NM_{xtn-1}})/(t_n - t_{n-1})\} \tag{15}$$

$$v_{N_{ytn}} = v_{M_{ytn-1}} + \{(d_{NM_{ytn}} - d_{NM_{ytn-1}})/(t_n - t_{n-1})\} \tag{16}$$

Once we find out the 2-axis horizontal velocity of the next ball, it should be used as the current reference/marker.

At any time the 2 axes horizontal velocity of the current reference of a continuous falling ball system can be predicted as:

$$v_{M_{xtn}} = v_{M_{xtn-1}} \quad (17)$$
$$v_{M_{ytn}} = v_{M_{ytn-1}} \quad (18)$$

For a continuous falling ball system the 2 axis average horizontal velocity of the vehicle can be expressed as:

$$v_{V_{x(tn-1\,\&\,tn)}} = v_{M_{xtn}} - \{(d_{RM_{xtn}} - d_{RM_{xtn-1}})/(t_n - t_{n-1})\} \quad (19)$$

$$v_{V_{y(tn-1\,\&\,tn)}} = v_{M_{ytn}} - \{(d_{RM_{ytn}} - d_{RM_{ytn-1}})/(t_n - t_{n-1})\} \quad (20)$$

3. Results and discussion

A computer simulation was designed by using the software MIT Scratch to validate the theoretical and mathematical model of the proposed sensor. To validate the mathematical equations described in the mathematical modeling section, simulation was done both for static and moving vehicle. All distances were calculated in meter and all velocities were calculated in meter per second (m/s).

In the static test, the simulation was done by placing the sensor assembly inside a vehicle whose velocity was zero in all three axes. But this velocity information was kept hidden from the sensor. Simulation results of the Static Test are shown in Table 3.

In the dynamic test, the simulation was done by placing the sensor assembly inside a vehicle that had a constant velocity 5 of m/s on the x-axis and 10 m/s on the y-axis. But this velocity information was kept hidden from the sensor. Simulation results of the Dynamic Test are shown in Table 4.

During the static test, the sensor could successfully find out the velocity of the vehicle as 0 m/s on the x-axis and 0 m/s on the y-axis at all times by using eqn. 19 and 20 respectively. These results are shown in v_{V_x} and v_{V_y} columns of Table 3 respectively.

During the dynamic test, the sensor could successfully find out the velocity of the vehicle as 5 m/s on the x-axis and 10 m/s on the y-axis at all times by using eqn. 19 and 20 respectively. These results are shown in v_{V_x} and v_{V_y} columns of Table 4 respectively. The dynamic test was done multiple times for a vehicle moving at different constant velocities. During each test, vehicle velocities were kept hidden from the sensor. In all cases, the sensor could successfully find out the velocity of the velocity correctly.

The simulation results showed that theoretically, the proposed sensor is capable of measuring the water and air vehicle speed with respect to the ground directly. Unlike conventional sensors, the proposed sensor completely eliminates the requirement for putting the sensor or its probes keeping outside the vehicle. The complete sensor assembly can be placed inside the vehicle. As a result, in comparison with currently available speed sensors, it may provide higher safety in case of collision with external objects. The proposed sensor is contactless and environment nondependent. As a result, unlike conventional sensors, its accuracy shall not be affected by environmental parameters like temperature, density, salinity, chemical properties, etc. The computer simulation was designed based on ideal physics and math equations. But in the real world, these results may vary. To overcome this limitation the proposed sensor needs to be investigated with practical experimentation.

4. Conclusion

Presently available speed sensors used in ships, submarines, and aircraft cannot measure the vehicle speed with respect to the ground directly. We proposed a theoretical model for a new speed sensor that can measure the horizontal speed of water and air vehicles with respect to the ground directly. The model has also been validated by designing a computer simulation. The simulation results showed that the proposed sensor is capable of measuring the water and air vehicle speed with respect to the ground directly. Unlike the conventional sensors, our proposed sensor is environmentally independent. It completely eliminates the requirement of putting the sensor or its probes outside the vehicle. The complete sensor assembly can be placed inside the vehicle. The proposed sensor is restricted in some aspects because we assume an ideal behavior of the falling balls in an airless environment. This limitation has to be investigated experimentally. If further experiments show good results a completely new type of speed sensor may be introduced which shall be able to measure the speed of water and air vehicles with respect to the ground directly. As well as it has the potential to provide better accuracy, safety, and applicability than currently available speed sensors used in water and air vehicles.

Table 3. Simulation results of the static test

Sample No.	***T* (s)**	**v_{M_x} (m/s)**	**v_{M_y} (m/s)**	**d_{RM_x} (m) (15)**	**d_{RM_y} (m)**	**v_{V_x} (m/s)**	**v_{V_y} (m/s)**	**d_{NM_x} (m)**	**d_{NM_y} (m)**	**v_{N_x} (m/s)**	**v_{N_y} (m/s)**	***Observation***
1st Ball is used as a reference												
1	0.1	10	10	1	1	0	0					
2	0.2	10	10	2	2	0	0					
3	0.3	10	10	3	3	0	0					
4	0.4	10	10	4	4	0	0					
5	0.5	10	10	5	5	0	0	4	4			2nd Ball appeared
6	0.6	10	10	6	6	0	0	4	4	10	10	
7	0.7	10	10	7	7	0	0	4	4	10	10	
8	0.8	10	10	8	8	0	0	4	4	10	10	
9	0.9	10	10	9	9	0	0	4	4	10	10	
2nd Ball is used as a reference as 1st Ball is lost												
10	1.0	10	10	6	6	0	0					
11	1.1	10	10	7	7	0	0	6	6			3rd Ball appeared
12	1.2	10	10	8	8	0	0	6	6	10	10	
13	1.3	10	10	9	9	0	0	6	6	10	10	
3rd Ball is used as a reference as 2nd Ball is lost												
14	1.4	10	10	4	4	0	0					
15	1.5	10	10	5	5	0	0	4	4			4th Ball appeared
16	1.6	10	10	6	6	0	0	4	4	10	10	
17	1.7	10	10	7	7	0	0	4	4	10	10	
18	1.8	10	10	8	8	0	0	4	4	10	10	
19	1.9	10	10	9	9	0	0	4	4	10	10	
4th Ball is used as a reference as 3rd Ball is lost												
20	2.0	10	10	6	6	0	0					
21	2.1	10	10	7	7	0	0	6	6			5th Ball appeared
22	2.2	10	10	8	8	0	0	6	6	10	10	
23	2.3	10	10	9	9	0	0	6	6	10	10	
5th Ball is used as a reference as 4th Ball is lost												
24	2.4	10	10	4	4	0	0					
25	2.5	10	10	5	5	0	0	4	4			6th Ball appeared
26	2.6	10	10	6	6	0	0	4	4	10	10	
27	2.7	10	10	7	7	0	0	4	4	10	10	
28	2.8	10	10	8	8	0	0	4	4	10	10	
29	2.9	10	10	9	9	0	0	4	4	10	10	
6th Ball is used as a reference as 5th Ball is lost												
30	3.0	10	10	6	6	0	0					
31	3.1	10	10	7	7	0	0	6	6			7th Ball appeared
32	3.2	10	10	8	8	0	0	6	6	10	10	
33	3.3	10	10	9	9	0	0	6	6	10	10	
7th Ball is used as a reference as the 6th Ball is lost												
. . .												
nth Ball is used as reference when (n-1)th Ball is lost												

Table 4. Simulation results of the dynamic test

Sample No.	***T* (s)**	v_{M_x} **(m/s)**	v_{M_y} **(m/s)**	d_{RM_x} **(m)** **(15)**	d_{RM_y} **(m)**	v_{V_x} **(m/s)**	v_{V_y} **(m/s)**	d_{NM_x} **(m)**	d_{NM_y} **(m)**	v_{N_x} **(m/s)**	v_{N_y} **(m/s)**	***Observation***
1st Ball is used as a reference												
1	0.1	15	20	1	1	**5**	**10**					
2	0.2	15	20	2	2	**5**	**10**					
3	0.3	15	20	3	3	**5**	**10**					
4	0.4	15	20	4	4	**5**	**10**					
5	0.5	15	20	5	5	**5**	**10**	4	4			2nd Ball appeared
6	0.6	15	20	6	6	**5**	**10**	4	4	15	20	
7	0.7	15	20	7	7	**5**	**10**	4	4	15	20	
8	0.8	15	20	8	8	**5**	**10**	4	4	15	20	
9	0.9	15	20	9	9	**5**	**10**	4	4	15	20	
2nd Ball is used as a reference as 1st Ball is lost												
10	1.0	15	20	6	6	**5**	**10**					
11	1.1	15	20	7	7	**5**	**10**	6	6			3rd Ball appeared
12	1.2	15	20	8	8	**5**	**10**	6	6	15	20	
13	1.3	15	20	9	9	**5**	**10**	6	6	15	20	
3rd Ball is used as a reference as 2nd Ball is lost												
14	1.4	15	20	4	4	**5**	**10**					
15	1.5	15	20	5	5	**5**	**10**	4	4			4th Ball appeared
16	1.6	15	20	6	6	**5**	**10**	4	4	15	20	
17	1.7	15	20	7	7	**5**	**10**	4	4	15	20	
18	1.8	15	20	8	8	**5**	**10**	4	4	15	20	
19	1.9	15	20	9	9	**5**	**10**	4	4	15	20	
4th Ball is used as a reference as 3rd Ball is lost												
20	2.0	15	20	6	6	**5**	**10**					
21	2.1	15	20	7	7	**5**	**10**	6	6			5th Ball appeared
22	2.2	15	20	8	8	**5**	**10**	6	6	15	20	
23	2.3	15	20	9	9	**5**	**10**	6	6	15	20	
5th Ball is used as a reference as 4th Ball is lost												
24	2.4	15	20	4	4	**5**	**10**					
25	2.5	15	20	5	5	**5**	**10**	4	4			6th Ball appeared
26	2.6	15	20	6	6	**5**	**10**	4	4	15	20	
27	2.7	15	20	7	7	**5**	**10**	4	4	15	20	
28	2.8	15	20	8	8	**5**	**10**	4	4	15	20	
29	2.9	15	20	9	9	**5**	**10**	4	4	15	20	
6th Ball is used as a reference as 5th Ball is lost												
30	3.0	15	20	6	6	**5**	**10**					
31	3.1	15	20	7	7	**5**	**10**	6	6			7th Ball appeared
32	3.2	15	20	8	8	**5**	**10**	6	6	15	20	
33	3.3	15	20	9	9	**5**	**10**	6	6	15	20	
7th Ball is used as a reference as the 6th Ball is lost												
.												
.nth Ball is used as reference when (n-1)th Ball is lost												

References

1. Griswold, W. Lyman, “Underwater logs,” Navigation: Journal of the Institute of Navigation, vol. 15, no. 2, pp. 127-135, Summer 1968.
2. C. G. Smith, and j. Slepian, “Electromagnetic ship’s log,” U.S. Patent 1249530, Dec. 11, 1917.
3. R. J. Paredes, Quintuña, M. T. M. Arias-Hidalgo and R. Datla, “Numerical flow characterization around a type 209 submarine using OpenFOAM” Fluids, vol. 6, no. 2, pp. 1-23, Feb. 2021, Art. No. 66.
4. The American Practical Navigator: Bowditch, Paradise Cay Publications, Bethesda, MD, USA, 2010.
5. P. Gloaguen, M. Woillez, S. Mahévas, Y. Vermard, and E. Rivot, “Is speed through water a better proxy for fishing activities than speed over ground,” Aquatic Living Resources, vol. 29, no. 2, pp. 1-8, Oct. 2016, Art. No. 210.
6. J. Mcorlly, “Pilot Tube,” U.S. Patent 2399370, Apr. 30, 1946.
7. Mechanics of flight, Pearson Education Limited, Harlow, ESX, UK, 2006.
8. Principles Of Flight: Aircraft General Knowledge, Flight Performance and Planning (Private Pilot's Licence Course), Airplan Flight Equipment Ltd, Wythenshawe, MCR, UK, 2005.
9. Tietjens, O.K.G. & Prandtl, L. 1957 Applied Hydro and Aeromechanics: Based on Lectures of L. Prandtl, vol. 2. Courier Corporation.
10. Fluid Flow Handbook, McGraw-Hill Professional, Blacklick, OH, USA, 2002.
11. W. Gracey, “Measurement of aircraft speed and altitude,” NASA Langley Research Center, Hampton, VA, USA, Rep No. NASA-RP-1046, May 1, 1980.
12. A. Pinker and C. Smith, “Vulnerability of the GPS Signal to Jamming,” GPS Solutions, vol. 29, no. 2, pp. 1-8, Oct. 2016, Art. No. 210.
13. A. Grant, P. Williams, N. Ward and S. Basker, “GPS jamming and the impact on maritime navigation,”. The Journal of Navigation, vol. 62, no. 2, pp. 173-187, Apr. 2009.
14. G. Taraldsen, T. A. Reinen and T. Berg, “The underwater gps problem,” presented at the Conf. Oceans 2011 IEEE, Spain, Jun. 6-9, 2011.
15. G. Xu, Y. Zhao, K. Xu, H. Xu and Z. Cheng, “A speed measurement method for underwater vehicle based on pulse speedometer and accelerometer,” presented at the 23rd Int Conf. Offshore and Polar Engineering, Anchorage, Alaska, Jun. 30 - Jul. 5, 2013.
16. Hwang, J. K., Uchanski, M. and Song, C. K. (2005). Vehicle speed estimation based on Kalman filtering of accelerometer and wheel speed measurements. Int. J. Automotive Technology 6, 5, 475–481.
17. A. Lawrence, “Gyro and accelerometer errors and their consequences,” in Modern Inertial Technology, New York, NY, USA: Springer, 1998, pp. 25-42.
18. S. Drake and J. Mclachlan, “Galileo's discovery of the parabolic trajectory,” Scientific American, vol. 232, no. 3, pp. 102-111, Mar. 1975.

Authors Introduction

Jakaria Mahdi Imam

He received his B.Eng. degree from Military Institute of Science and Technology, Bangladesh. Currently, he is pursuing master degree by research from University of Malaya, Malaysia. His research interest includes emerging naval technologies, inertial navigation system and inertial sensors.

Dr. Mohammad Aminul Islam

He received M.Sc. (research) and Ph.D. degree in Electrical, Electronic, and System Engineering from the National University of Malaysia (UKM), Malaysia, in 2012 and 2015. He is currently a Senior Lecturer in the Department of Electrical Engineering, University of Malaya. His research interests are semiconductor materials, optical materials and energy transportation.

Dr. Norrima Mokhtar

She received the B.Eng. degree from University of Malaya, the M.Eng. and the Ph.D. degree from Oita Univerity, Japan. She is currently a Senior Lecturer in the Department of Electrical Engineering, University of Malaya. Her research interests are signal processing and human machine interface.

Dr. S. F. W. Muhammad Hatta

She received her M.Eng. in Electrical Electronics Engineering (University of Sheffield, UK) in 2005, M.Sc. in Microelectronics (University of Malaya) in 2009 and Ph.D. in Microelectronics (LJMU,UK) in 2014. She is currently a Senior Lecturer in the Department of Electrical Engineering, University of Malaya. Her research interests include emerging technologies in wearable electronics, semiconductor reliability, advanced semiconductor modelling and characterisation as well as PUF-based technology in cyber-securities.

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.

On Correcting Luminosity and Contrast of Retinal Images with Reflectance

Mofleh Hannuf AlRowaily, Hamzah Arof, Imanurfatiehah Ibrahim, Wan Amirul Wan Mohd Mahiyiddin, Norrima Mokhtar
Department of Electrical Engineering, Faculty of Engineering,
University of Malaya 50603, Kuala Lumpur Malaysia

Email: ahamzah@um.edu.my, imanurfatiehah@gmail.com, wanamirul@um.edu.my, norrimamokhtar@um.edu.my

Abstract

In this paper, we perform an automatic correction of luminosity and contrast of retina images. One hundred retina images with varying level of reflectance taken from custom and online databases are used to test the effectiveness of the proposed method. The approach is implemented in 4 stages namely pre-processing, lowpass filtering, luminosity equalization, and contrast stretching. In the pre-processing stage, the three components of a color retina image are separated and only the green channel is processed further as it contains the most information. Then the region of interest (which is the eye region) and its border are marked. After that, the eye region (ROI) of the green channel is subjected to lowpass filtering, row by row to create a smooth background luminosity surface without the foreground objects like the optic disc, exudates, blood vessels and blood spots. Three different types of lowpass filters are used and their performances are compared. The resulting background luminosity surface allows the estimation of the background illumination. Then, based on the background surface, the luminosity of the ROI is equalized so that every pixel experiences the same brightness. Finally, the contrast of the ROI is improved by histogram stretching so that the foreground objects appear more clearly. The proposed method was implemented using MATLAB R2021b running on AMD 5900HS processor and the average execution time was less than 1 second. The execution time can be further reduced if the codes are optimized and GPU is used. Overall, the proposed method improves the luminosity and contrast of the images greatly. This technique can be a useful tool to ophthalmologists who perform visual inspection of microaneurysm, exudates and other lesions.

Keywords: Retina Images, Luminosity, Contrast, Reflectance, Lowpass Filters.

1. Introduction

Ophthalmologists examine the foreground and background of retina images to assess the eye condition of a patient. As the process is time-consuming, it becomes daunting when there are many patients to diagnose. Therefore, computer assisted diagnosis is an invaluable tool that expedites the process and increases its effectiveness. The quality of retina images plays an important role in ensuring the accuracy of diagnosis. Retina images may suffer from uneven illumination, blurring, and low contrast. Image enhancement can improve the quality of an image for a better visual perception. For retina images suffering from luminosity and contrast variation, many image enhancement methods have been proposed to overcome these issues [1][2][3][4]. The methods are categorized into two main groups namely spatial and frequency domain approaches. Spatial domain approaches work directly with the pixels of an image where the pixel values are transformed or altered [5][6]. Techniques that manipulate pixels directly include logarithmic transforms, power law transforms, and histogram equalisation. Meanwhile frequency domain uses Fourier transform (FT), discrete cosine transform (DCT), and discrete wavelet transform (DWT) to convert an image to another domain, then manipulate it and then inverse transform the outcome to its original image space [7][8].

Most color retina image enhancement methods work in HSV or LAB color space to improve the luminosity of the images. Then contrast is improved using histogram equalization and other techniques [7][8][9][10]. There are also a few methods that work in RGB space but it is more difficult to improve color images in RGB space since the color tone can change if the color components are not adjusted carefully. The majority of the works are done in greyscale and so does our method.

The work presented here is a spatial domain approach. The method utilizes spatial filtering to improve and enhance the luminosity and contrast of images respectively. The rest of this paper is organized as follows. Section 2 describes the methodology of the proposed method. Section 3 presents the initial experimental results, and Section 4 provides the conclusion.

2. Methodology

Periodic evaluation of retinal images is important to patients suffering from retinopathy. Here, we propose a new method to correct both luminosity and contrast of retina images that are affected by strong reflectance. The process is executed in four stages which are pre-processing, filtering, luminosity adjustment, and contrast stretching. Fig. 1 shows the flow of the process.

In the pre-processing stage, the red, green and blue (RGB) channels of the retina image are separated. At this juncture, the red and blue channels can be dropped. However, they can also be processed if deemed necessary. The green channel G(x,y) is further processed as it is deemed to contain the most information. First, the area of the eye and its border are identified by thresholding. We realize that in the ROI of an image that suffers from reflectance, the centers of reflectance are located in the vicinity of its border. After locating the reflectance centers, measures can be taken to reduce the effects of reflectance before lowpass filtering.

Next, in lowpass filtering stage, each row of the ROI is processed one by one. First, the intensities of all pixels in a row are stored in an array. Then the array is subjected to lowpass filtering at three resolutions to smoothen its values. For this purpose, we perform downsampling to lower the resolution of the image. Then upsampling is performed to increase the size of the images back to its original scale. Three different filters are tested, and they are the median filter, Gaussian and moving average filters. The length of the filter is 11, 13 and 15 for the three resolutions. The image can be further smoothened column by column using the same filters. Once the lowpass filtering step is completed, the result is a smooth estimate of the background surface for the image called the background luminosity surface (BLS). The BLS represents the background luminosity of every pixel in the image without foreground objects. As the ROI is small, it is logical to assume that the background luminosity of every pixel should be equal. Hence, the luminosity of each pixel should be corrected and made the same. In our experiment, our target is to level the average intensity of all pixel in an image to 128. This value is chosen since it is the middle intensity value for normal images represented by eight bit unsigned integer. The intensity of each pixel G(x,y) in the input image can be higher or lower than its background luminosity in the BLS(x,y). So, at every location (x,y) in the ROI, the difference between the input G(x,y) and its background BLS(x,y) is calculated and recorded. This difference shall be maintained even after the background BLS(x,y) is adjusted to 128 to equalize the background brightness for each pixel in the ROI. The result of equalizing the background of the input image G(x,y) is an image with evenly spread luminosity called equalized brightness image (EBI). Keep in mind that the difference between G(x,y) and BLS(x,y) is always added to 128 at every location (x,y).

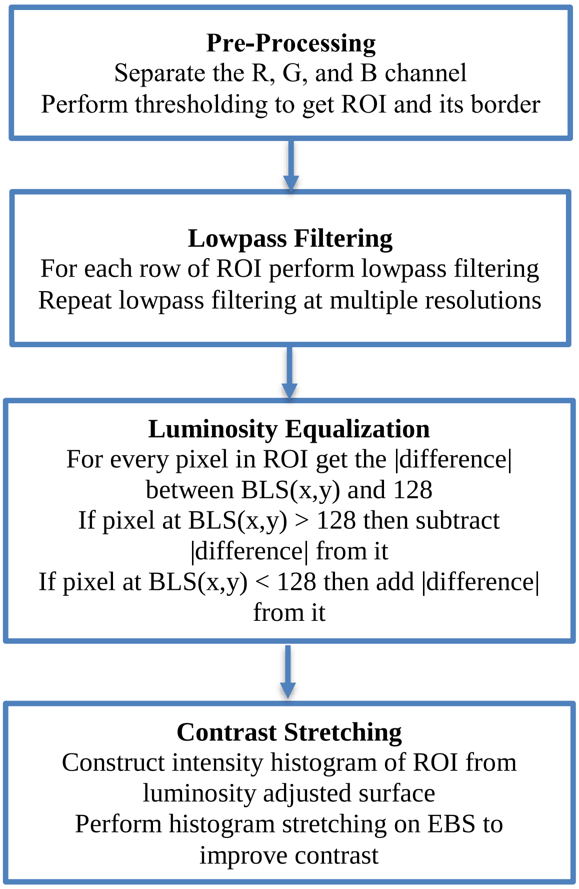

Pre-Processing
Separate the R, G, and B channel
Perform thresholding to get ROI and its border
Lowpass Filtering
For each row of ROI perform lowpass filtering
Repeat lowpass filtering at multiple resolutions
Luminosity Equalization
For every pixel in ROI get the |difference| between BLS(x,y) and 128
If pixel at BLS(x,y) > 128 then subtract |difference| from it
If pixel at BLS(x,y) < 128 then add |difference| from it
Contrast Stretching
Construct intensity histogram of ROI from luminosity adjusted surface
Perform histogram stretching on EBS to improve contrast

Fig. 1 Flow of stages in the process

The next step is to perform histogram stretching on the ROI of EBI(x,y) to improve its contrast. First the histogram of intensity frequency of the image is constructed. Then the histogram spread is calculated. It is the difference between the highest and lowest intensities in the image. Then the histogram is stretched so that its spread occupies the whole range of 0 until 255. Of course, this method only works if the histogram spread is less than 255 to start with. Otherwise, histogram equalization or CLAHE can be used instead of histogram stretching to improve the image contrast.

It is not necessary for the red and blue channels to be processed along with the green channel. However, they can be subjected to the same treatment if required. Details from the red and blue channels can be incorporated into the green channel to enhance its appearance and contrast. Fig. 2 shows the green channel of an input image G(x,y) and its BLS, EBI and contrast stretched EBI (CEBI). As seen, there is little difference between EBI and CEBI of the image. This is because histogram stretching does little on an image which already has a good contrast. It is also possible to combine both histogram stretching and equalization to improve an image. However, the effect is almost similar to histogram stretching or equalization alone. It is also noted that the effect of boundary reflectance at the top border of the image is completely gone in CEBI. It seems that the method is just as effective for images that require luminosity and contrast adjustments without suffering from strong boundary reflectance.

3. Results and Discussion

In this work, 100 retina images taken from custom and online databases were utilized to test the performance of the proposed method. It turned out that the results produced by the Gaussian and moving average filters were almost identical. The results of median filter were not as good as those of the other two filters because some important details of the foreground objects were eliminated by the filter. Perhaps this is because median filter is a nonlinear lowpass filter. However, the execution times of all three filters were less than 1 second. If the codes are optimized and the GPU unit is used, they can be made lower. The best performance belongs to the moving average filter since it records the fastest time in implementation and generates the best result. On average, the proposed method manages to reduce 30% luminosity variation and increase the contrast of the greyscale images by more than 100%. Luminosity variation is implied from the standard deviation of the pixel intensities in the images and contrast is calculated using the metric proposed by Matkovic et. al [12]. It is similar to the sum of absolute difference calculated at three different resolutions. The approach was implemented on a laptop powered by an Intel i7 processor using MATLAB R2021b. This technique can be a useful tool to ophthalmologists who perform visual diagnosis of microaneurysm, fundus and other lesions. Fig. 3 shows five samples of retina images whose luminosity and contrast have been corrected using the proposed method. As seen, all of the boundary reflectance effects are removed from the images.

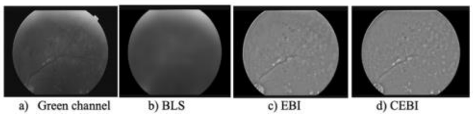


Fig. 2 The green channel of an image and its BLS, EBI and CEBI

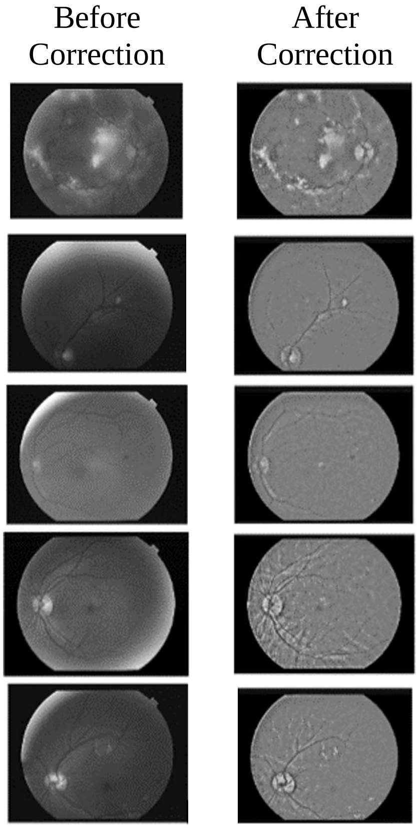


Fig. 3 Five samples of retina images before and after correction.

4. Conclusion

In this work, a new method to rectify luminosity and contrast variation in retina images is introduced. The method successfully removes luminosity variation and improves the contrast of the images even when they suffer intense reflectance. If a post filtering stage is introduced, details of the adjusted green channel can be further enhanced using information from the red and blue channels. The performance of the approach is tested on 100 test images and improvement is noticeable visually and quantitatively for both luminosity and contrast. On average, this method manages to reduce 30% luminosity variation and increase the contrast of the retina images significantly.

References

1. Mathews, M. R., & Anzar, S. M. (2021). A comprehensive review on automated systems for severity grading of diabetic retinopathy and macular edema. International Journal of Imaging Systems and Technology, 31(4), 2093-2122.
2. Xiao, D., Bhuiyan, A., Frost, S., Vignarajan, J., Tay-Kearney, M. L., & Kanagasingam, Y. (2019). Major automatic diabetic retinopathy screening systems and related core algorithms: a review. Machine Vision and Applications, 30(3), 423-446.
3. Kang, Y., Fang, Y., & Lai, X. (2020). Automatic detection of diabetic retinopathy with statistical method and Bayesian classifier. Journal of Medical Imaging and Health Informatics, 10(5), 1225-1233.
4. Gupta, B., & Tiwari, M. (2019). Color retinal image enhancement using luminosity and quantile-based contrast enhancement. Multidimensional Systems and Signal Processing, 30(4), 1829-1837.
5. Qi, Y., Yang, Z., Sun, W., Lou, M., Lian, J., Zhao, W, & Ma, Y. (2021). A comprehensive overview of image enhancement techniques. Archives of Computational Methods in Engineering, 1-25.
6. Vijayalakshmi, D., Nath, M. K., & Acharya, O. P. (2020). A comprehensive survey on image contrast enhancement techniques in spatial domain. Sensing and Imaging, 21(1), 1-40.
7. Sahu, S., Singh, A. K., Ghrera, S. P., & Elhoseny, M. (2019). An approach for de-noising and contrast enhancement of retinal fundus image using CLAHE. Optics & Laser Technology, 110, 87-98.
8. Alwazzan, M. J., Ismael, M. A., & Ahmed, A. N. (2021). A hybrid algorithm to enhance colour retinal fundus images using a Wiener filter and CLAHE. Journal of Digital Imaging, 34(3), 750-759.
9. Cao, L., Li, H., & Zhang, Y. (2020). Retinal image enhancement using low-pass filtering and α-rooting. Signal Processing, 170, 107445.
10. Yang, L., Yan, S., & Xie, Y. (2021). Detection of microaneurysms and hemorrhages based on improved Hessian matrix. International Journal of Computer Assisted Radiology and Surgery, 16(6), 883-894.
11. Zhou, M., Jin, K., Wang, S., Ye, J., & Qian, D. (2017). Color retinal image enhancement based on luminosity and contrast adjustment. IEEE Transactions on Biomedical engineering, 65(3), 521-527.
12. Matkovic, K., Neumann, L., Neumann, A., Psik, T., & Purgathofer, W. (2005). Global contrast factor-a new approach to image contrast. In Computer Aesthetics in Graphics, Visualization and Imaging, (pp. 159-167).

Authors Introduction

Mofleh Hannuf AlRowaily

He is a PhD student at the Electrical Engineering Department, Faculty of Engineering, University of Malaya.

Prof. Ir. Dr. Hamzah Arof

He received the B.Sc. degree from Michigan State University, USA, and the Ph.D. degree from the University of Wales, U.K. He is currently a Professor with the Electrical Engineering Department, Universiti Malaya, Malaysia. He is an established academician. He has published more than 80 articles. His research interests include image and signal processing, robotics, and photonics

Imanurfatiehah Ibrahim

She graduated from International Islamic University of Malaysia and obtained her masters from University of Malaya. Currently she is a PhD student at the Electrical Engineering Department, Faculty of Engineering, University of Malaya. Her research interests are in the areas of big data processing and remote sensing.

Wan Amirul Wan Mohd Mahiyiddin

He received the M.Eng degree from the Imperial College of Science, Technology and Medicine. He received his PhD Degree from Canterbury University, New Zealand. He is currently Senior Lecturer at the Electrical Engineering Department, University of Malaya, Malaysia. His area of expertise is Telecommunications specifically in Multiple Antenna Systems.

Norrima Mokhtar

She received the B.Eng. degree from University of Malaya, the M.Eng. and the Ph.D. degree from Oita Univerity, Japan. She is currently a Senior Lecturer in the Department of Electrical Engineering, University of Malaya. Her research interests are signal processing and human machine interface.

Rate Adaptation for Quality of Service (QoS) Improvement in IEEE 802.11ax Wireless Local Area Network (WLAN)

Hazwani Binti Zawawi
School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA,40450 Shah Alam, Selangor, Malaysia

Dr Wan Norsyafizan W.Muhamad
School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA,40450 Shah Alam, Selangor, Malaysia

Dr Suzi Seroja Sarnin
School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA,40450 Shah Alam, Selangor, Malaysia

Dr Nani Fadzlina Naim
School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA,40450 Shah Alam, Selangor, Malaysia

E-mail: syafizan@uitm.edu.my

Abstract

This paper presents rate adaptation for Quality of Service (QoS) improvement in Wireless Local Area Network (WLAN) specifically IEEE802.11ax. The latest WLAN standard provides better results of performance by improving the WLAN data rate to satisfy the growing number of users' demand. By increasing the data rate, high throughput and low delay will be provided as well as a better network performance can be experienced by users. This study implemented a link adaptation technique which adapts the transmission data rate in IEEE 802.11ax WLAN. The main objectives of this paper are to design link adaptation which adapts the transmission data rate based on radio channel condition and to verify the effectiveness of the proposed algorithm. This technique enables the QoS performance to be improved in terms of its throughput and delay that able to adapt the Modulation Coding Scheme (MCS) based on the radio channel condition. This performance metric is analyzed using OMNeT++ simulator. Simulation results show that the data rate adaptation technique offers better performance in terms of throughput and delay.

Keywords: IEEE 802.11ax, WLAN, QoS, delay, throughput, rate adaptation

1. Introduction

These nearly past two decades, the wireless local area network (WLAN) has become one of the important technology and facilities in our daily life. Any devices that include access point (AP) to the internet and use high frequency radio waves, the method for the wireless distribution is called Wireless Local Area Network (WLAN). During late 1990s, most WLAN solutions and software protocols in different variation versions were replaced based on IEEE 802.11 standard (versions "a" through "ac"). IEEE 802.11 defines as a set of physical layer standards that define communication for WLAN WiFi should not be confused with the Wi-Fi Alliance's Wi-Fi trademark [1]. Wi-Fi is not a technical term, but is described as a superset of the standard IEEE 802.11 and is sometimes used interchangeably with the standard [2]. Fig. 1 shows the evolution of Wi-Fi since its first release until 802.11ax. Process of the development of IEEE 802.11ax is a step of innovation of the current standard. It provides more users in pack environments and support a better network performance for wireless LAN networks [3]. The new standard could also power high technology applications such as Ultra High Definition (HD), high quality video, wireless office, and Internet of Things (IoT). IEEE 802.11 ax has been said could offer 4 times better throughput than IEEE 802.11 ac [4].

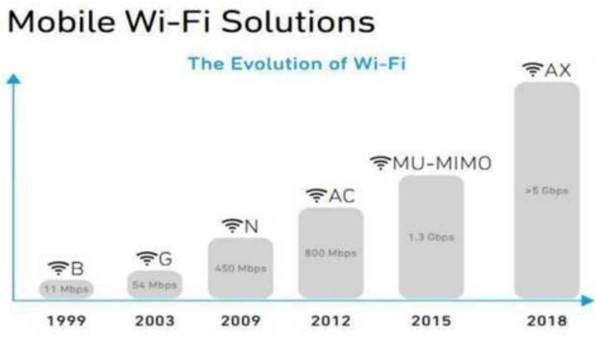


Fig. 1 Wi-Fi evolutions

This modern and technology era has shown the increasing number of smartphone users for more than ten years. The massive number of users give a sign that technology and wireless are growing proportionally with year [5]. Every year more technologies are created for users. Hence, IEEE 802.11ax was developed for maximum ability for one network to work with another, working efficiently with 802.11 devices [6].

Quality of service or also known as QoS, is the characterization or measurement of a service's overall performance, such as a telephony or computer network or cloud technology service, particularly the performance and quality of network users [7]. QoS is the capabilities to prioritize various applications, data flows, clients, or to justify the data flow performance at certain level [8]. QoS has been defined as a quality measures, with many definitions, rather than being referred as reserve capacity. Service quality has also refers to the level of service quality, for instance, the quality of service guaranteed. In this research paper, the QoS is measured on the throughput of the wireless LAN performance.

In wireless communications, the term link adaptation or adaptive coding and modulation (ACM) is used to indicate the adjustment of modulation, coding and other signal and protocol parameters to the radio link conditions. In this situation, WLAN uses a rate adaptation and power adaptation algorithm that adapts the modulation and coding scheme(MCS) to the performance and quality of radio channel, hence the bit rate and data transmission robustness [9] [10]

In this research, a rate adaptation technique is used to determine the throughput performance. Rate adaptation can be defined as a change in data rate based on traffic conditions. This condition has three categories which are low traffic conditions, medium traffic conditions, and high traffic conditions. This technique was chosen because of its ability to improve network performance. Besides, its ability in adapting modulating and coding schemes have an advantage in maintaining the QoS especially in high traffic conditions in IEEE 802.11ax.

Link adaptation has been well studied by many researchers to improve network performance, especially in throughput. Yalda et al introduce a novel machine learning-based approach for dynamically selecting whether to enable or disable the RTS/CTS of IEEE 802.11 DCF [11]. The proposed technique compares the cost of using RTS/CTS or retransmitting data based on the air time, or the ratio between the size of data/control information being broadcast and transmission rate, and network contention. Our algorithm then dynamically switches RTS/CTS on and off as necessary.

Purandare et al [12] presented a Loss Differentiated Channel Aware Rate Adaptation (LD-CARA) for IEEE 802.11n Wireless Links which implements an open-loop per packet for 802.11 ac standard. This research uses block acknowledgment (ACK) in order to differentiate losses as well as optimizing throughput performance for that standard. The algorithm is carried out to observe the behavior of signal to noise ratio at the receiver and based on the measurement will directly related to the bit error rate in the link. Network simulator NS3 is used to run the simulation and compare the adaptations. LD-CARA is a stable algorithm that helps to improve network throughput in 3 situations; static, mobile and interfered channels.

Hence, this research is done based on the objectives of to study the IEEE 802.11ax WLAN standard using rate adaptation technique and analyze the network performance of QoS via OMNeT++ simulator [13] Section 1 briefed a summary of analysis and the case study related with this research. Section II provides an overview of OMNeT++ and performance metric of IEEE

802.11ax standard. Section III will be focusing on the technique of rate adaptation and sector 1V discussed performance analysis based on rate adaptation technique. Lastly section V concludes the paper with suggestion of future works.

2. Background

A. OMNeT++

The OMNeT++ simulator is a simulator used to design a simulation network model in this study. OMNeT++ was used in this project to obtain the performance analysis of IEEE 802.11ax network model. The main elements are modules that have various types of modules, which are basic modules and complex modules. The actual simulation model in OMNeT++ is a complex module called "Network" as stated in [10]. Simulation specified in the configuration file runs with parameter values of IEEE 802.11ax and develops simulation networks using a framework called INET.

B. Performance Metrics

Performance metrics are used to evaluate the behavior, activities, and performance of a service. In this project, performance metrics focuses on Quality of Services (QoS) which are throughput and delay.

Throughput is an important measure in the wireless network which refers to the successful level of information or packet transmission across a communication channel. The throughput is usually being stated in Kilobits per second (kbps), Megabits per second (Mbps) or Gigabits per second(Gbps). On the other hand, throughput refers to the total number of data packets transmitted to the receiver.

Delay is defined as the time required for a packet to reach the receiver site successfully. For good network performance, low delay is required in order to achieve good throughput and better QoS performance.

3. Methodology

A. Network Model

OMNeT++ is used to perform the simulation process.
The example of network model scenario of proposed algorithm in OMNeT++ simulator is shown in Fig 2. The initial model network consists of user terminals and one access point (AP) placed randomly.

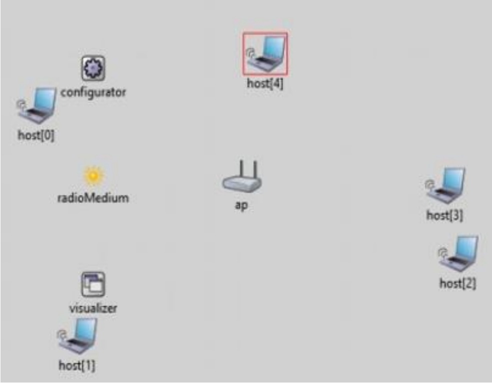


Fig. 2. Network model

Table 1 displays simulation parameters that need to be set in the omnetpp.ini configuration file in order to evaluate the performance of the rate adaptation technique for IEEE 802.11ax. The tables show the parameters that need to be taken into account before running the simulation.

TABLE I. SIMULATION PARAMETER

Parameters	Value
WLAN Standard	IEEE 802.11ax
Carrier Frequency	5GHz
Bandwidth	20MHz
Propagation model	Free space path loss
Spatial Steam	8
Transmission Range	100-200 meters
Simulation Time	120 seconds
Network Load (node)	10,20,30,40,50

B. Block Diagram

Fig. 3 shows a block diagram of rate adaptation in the IEEE 802.11 ax. First of all, the initial data rate value for transmission is set at the maximum rate by each transmitter. Then, after packets are sent to the receiver, and delay will be used as a parameter adaptation which represents the traffic condition. The delay of each transmission will be measured, so the transmission data rate is adjusted to its successive rate transmission. If the delay is more than the threshold value for the delay in IEE 802.11 ax which is 1µs, the transmission data rate will be reduced to one step to another step so it can achieve its target to get minimum delay with high throughput.

Apart from that, if the delay is much smaller than the threshold value, the transmission data rate will be increased since the data rate is inversely proportional to the delay. This process is done in order to gain the advantage of achieving the maximum throughput for users.

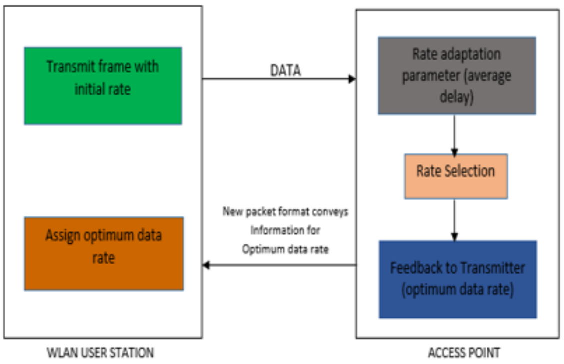

Transmit frame with initial rate
DATA
Rate adaptation parameter (average delay)
Rate Selection
New packet format conveys Information for Optimum data rate
Assign optimum data rate
Feedback to Transmitter (optimum data rate)
WLAN USER STATION
ACCESS POINT

Fig. 3. Block Diagram for Rate Adaptation Technique

Determination of the change of data rate based on traffic conditions in a communication network is known as rate adaptation [13]. Table for transmission data rate and Modulation & Coding Scheme (MCS) for IEEE 802.11ax standard is shown in Table II. Higher MCS values lead to higher SNR values that contribute to higher data rates that provide higher performance.

TABLE II. TRANSMISSION DATA RATE FOR IEEE 802.11ax STANDARD

Modulation	Coding Rate	Data Rate (Mbps)	Minimum SNR (dB)
BPSK	1/4	3.6	2
BPSK	1/2	7.3	5
QPSK	1/4	7.3	9
QPSK	1/2	14.6	11
QPSK	3/4	21.9	15
16-QAM	1/4	14.6	18
16-QAM	1/2	29.3	20
16-QAM	3/8	21.9	25
16-QAM	3/4	43.9	29
64-QAM	2/3	58.5	31
64-QAM	3/4	65.8	33
64-QAM	5/6	73.1	35
256-QAM	3/4	87.8	37
256-QAM	5/6	97.5	39
1024-QAM	3/4	109.7	40
1024-QAM	5/6	121.9	41

C. *Flowchart*

Fig. 4. shows the flowchart of the proposed algorithm. The first flow starts with transmitting a packet, i from transmitter to receiver. Each time one packet is transmitted, the transmitter will wait for the acknowledgment (ACK) packet from the receiver. If there is no indicator of ACK packet received at the receiver, the transmitter will retransmit the information again until the ACK packet is received via the carrier-sense multiple access with collision avoidance (CSMA/CA). While if the packet is successfully received at the receiver, the delay for a packet to arrive at the receiver will be measured. This procedure will be repeated until all of the packets are sent. After all, packets are sent, the average delay of the packets will be recorded.

In this work, for IEEE 802.11ax specification, the delay threshold for this standard is set as 1µs [14]. The value of the average delay that has been recorded is checked whether it has exceeded or not higher than the threshold value. In the case of the recorded average delay obtained is less than the threshold value for example 0.5µs, the transmission data rate will be increased in order to get a better result of throughput. Meanwhile, if the average delay is higher than the threshold value for example 1.5µs, adaptation of the transmission data rate to much lower level will be performed to control the delay. This discussion can be mathematically expressed by the equation (1) below,

$$t = \frac{L}{R} \qquad (1)$$

where t represents delay and R represents the transmission data rate and L represents the packet size. The equation as shown in equation (1) indicates that low delay can be obtained if the transmission data rate is increased. Hence, the new transmission packet will use the new adaptation of the data rate that has been implemented. This process will be repeated until the final packet has been successfully delivered to the receiver as set in simulation.

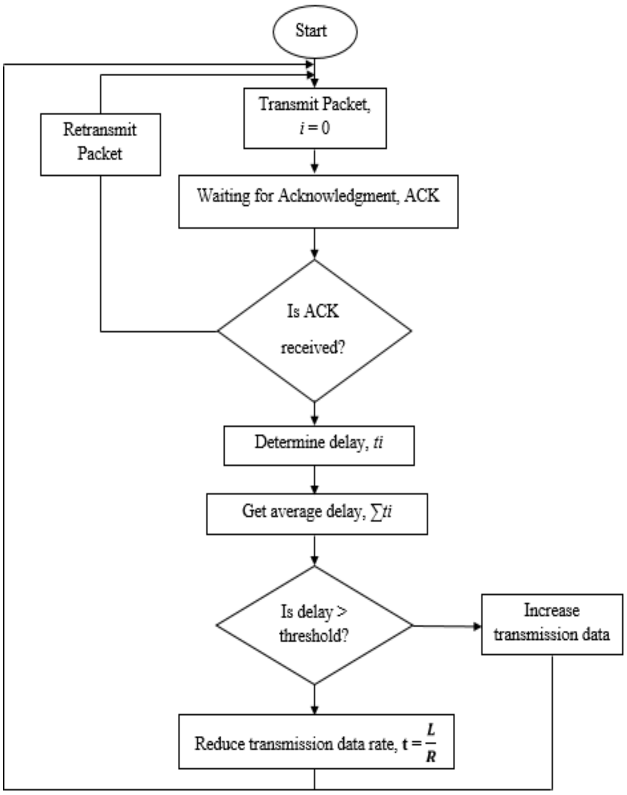

Start
Transmit Packet, $i = 0$
Retransmit Packet
Waiting for Acknowledgment, ACK
Is ACK received?
Determine delay, ti
Get average delay, $\sum ti$
Is delay > threshold?
Increase transmission data
Reduce transmission data rate, $t = \frac{L}{R}$

Fig. 4. Flowchart of rate adaptation

4. Results and Discussion

Results of simulation on performance analysis of rate adaptation algorithm in IEEE 802.11ax WLAN are presented in Fig. 5 to Fig. 7. Transmission rate for three conditions are well illustrated in Fig.5. The blue line indicates low traffic condition, the orange line indicates medium traffic condition and the grey line indicates high traffic condition. The node represents the number of users in the network. At low traffic condition (blue), the transmission rate value is higher than the other two condition and is decreasing when number of user increasing. This can be explained by equation in (1) which applies that delay value is lower than threshold value (1µs) when in low traffic condition. Hence, selection of high level transmission data rate based on MCS will give opportunity of throughput performance increment.

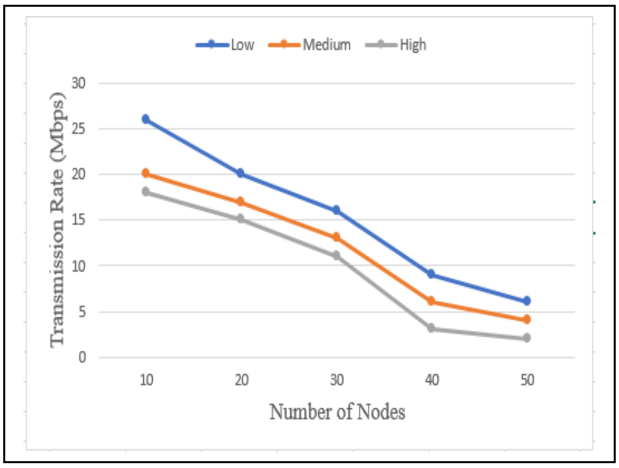


Fig. 5. Transmission rate for three network conditions

Fig. 6. shows the throughput for three conditions of network traffic. As the number of user increase, the throughput value or network performance for high traffic condition does not give satisfy value of throughput. This is because, during high traffic condition, the chances of packet loss during the transmission process is higher due to the congested network there. Hence, it will lead to higher collisions level.

However, at low traffic condition, as the user is increasing, the throughput is also increasing at such high rates. This happens because as the number of users keep expanding, the transmission data rate boots up their data rate in order to let the user experience a better throughput and network performance.

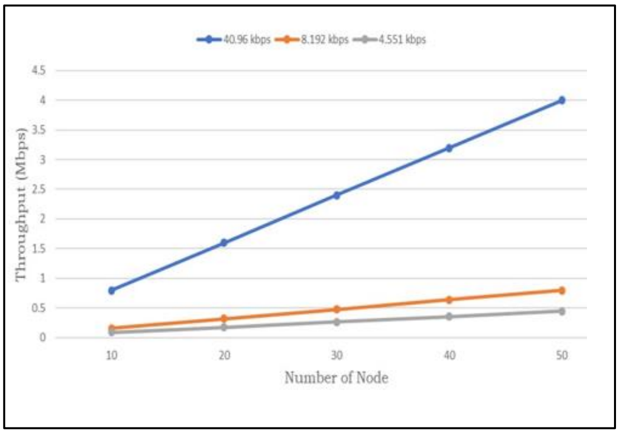


Fig. 6. Throughput for different traffic network conditions

The results of delay for the time taken for the packets to arrive at the receiver is being pictured in Fig. 7. It can be seen that during low traffic condition, when user is increasing, the average delay for the packet transmission is also keeps increasing. This is due to the traffic load has become congested at this condition, making the duration of packets transmission become slower because more packets try to access the network. Delay is proportional to node or number of user. When user in the network increases, the delay will also follow the result. This is also can be supports by higher contention level in the network which leads to lower percentage of successful packet transmission.

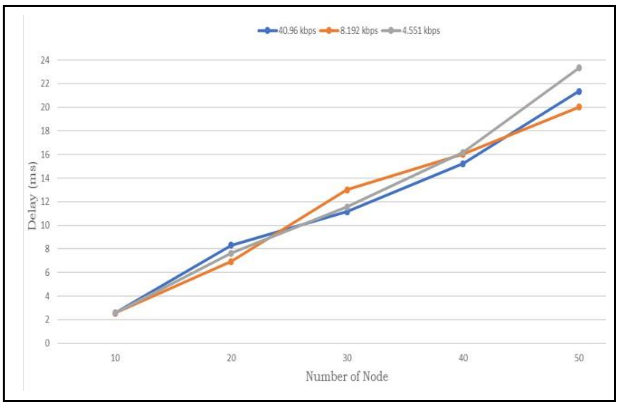


Fig. 7. Average delay for different traffic network conditions

5. Conclusion

In a conclusion, the performance analysis of rate adaptation technique in wireless network for IEEE 802.11ax are being studied. The QoS performance which is throughput and delay for tis standard has been analyzed and obtained based on different network condition and they are high traffic condition, medium traffic condition and low traffic condition. Based on the results, it has proven that by applying rate adaptation technique, it gives a better network QoS performance of IEEE 802.11ax standard. This will be an advantage for the low traffic and medium traffic user because they get to experience better data rates and network performance despite of its traffic condition.

6. Recommendation

For the recommendation, a new rate adaptation algorithm will be designed for the new Wi-Fi standard IEEE 802.11ax WLAN. This new algorithm will be designed in order to give more advantage on QoS performance as well as energy- efficient performance for this standard. This is because the reduction of power consumption on a device can prolong the lifetime of a battery.

References

[1] "IEEE Standard for Information Technology--Telecommunications and Information Exchange between Systems - Local and Metropolitan Area Networks--Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications - Redline," in IEEE Std 802.11-2020 (Revision of IEEE Std 802.11-2016) - Redline , vol., no., pp.1-7524, 26 Feb. 2021.

[2] Shuaib K. Memon, Kashif Nisar, Mohd Hanafi Ahmad Hijazi, B.S. Chowdhry, Ali Hassan Sodhro, Sandeep Pirbhulal, Joel J.P.C. Rodrigues, "A survey on 802.11 MAC industrial standards, architecture, security & supporting emergency traffic: Future directions," Journal of Industrial Information Integration, Volume 24, 2021.

[3] Gokalgandhi B., Tavares M., Samardzija D., Seskar I., Gacanin H, "Reliable Low-Latency Wi-Fi Mesh Networks," (2022) IEEE Internet of Things Journal, 9 (6), pp. 4533 – 4553.

[4] "IEEE Standard for Information Technology--Telecommunications and Information Exchange between Systems Local and Metropolitan Area Networks--Specific Requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 1: Enhancements for High-Efficiency WLAN," in IEEE Std 802.11ax-2021 (Amendment to IEEE Std 802.11-2020) , vol., no., pp.1-767, 19 May 2021, doi: 10.1109/IEEESTD.2021.9442429.

[5] Pan C., Wang Y., Shi H., Shi J., Cai R, "Network Traffic Prediction Incorporating Prior Knowledge for an Intelligent Network", Sensors, 22 (7), art. no. 2674, 2022, doi: 10.3390/s22072674.

[6] Z. Machrouh and A. Najid, "High Efficiency WLANs IEEE 802.11ax Performance Evaluation," 2018 International Conference on Control, Automation and Diagnosis (ICCAD), Marrakech, Morocco, 2018, pp. 1-5, doi: 10.1109/CADIAG.2018.8751296.

[7] Sanan, Hamid & Alam, Khubaib & Rafique, Muhammad & Khan, Bilal, “Quality of Service Enhancement in Wireless LAN: A Systematic Literature Review”, 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics, 2019, 1-8, doi: 10.1109/MACS48846.2019.9024827.

[8] S. Nosheen and J. Y. Khan, "Quality of Service- and Fairness-Aware Resource Allocation Techniques for IEEE802.11ac WLAN," IEEE Access, vol. 9, pp. 25579-25593, 2021, doi: 10.1109/ACCESS.2021.3051983.

[9] Anuar A.S.M., Muhamad W.N.W., Ali D.M., Sarnin S.S., Wahab N.A, “A review on link adaptation techniques for energy efficiency and QoS in IEEE802.11 WLAN, Indonesian Journal of Electrical Engineering and Computer Science, vol.17, no.1, pp. 331 - 339, 2019, doi: 10.11591/ijeecs.v17.i1.pp331-339

[10] Muhamad W.N.W, Sarnin S.S, Idris A, Syahira A, “Link adaptation algorithm for IEEE 802.11 wireless local area networks in fading channel”, Indonesian Journal of Electrical Engineering and Computer Science, vol.12, no.2, pp. 677 - 684, 2018, doi: 10.11591/ijeecs.v12.i2.pp677-684.

[11] Y. Edalat, K. Obraczka, and J. S. Ahn, “Smart adaptive collision avoidance for IEEE 802.11,” *Ad Hoc Networks*, vol. 124, no. November 2020, p. 102721, 2022.

[12] W. Yin, P. Hu, J. Indulska, M. Portmann, and Y. Mao, “MAC-layer rate control for 802.11 networks: a survey,” Wireless Networks, vol. 26, no. 5, pp. 3793–3830, 2020, doi: 10.1007/s11276-020-02295-2.

[13] R. G. Purandare, S. P. Kshirsagar, and S. M. Koli, “Loss Differentiated Channel Aware Rate Adaptation for IEEE 802.11n Wireless Links,” *Wirel. Pers. Commun.*, no. 0123456789, 2019. M. F. Monir and T. A. Ishmam, "Exploiting Link Diversity in IEEE 802.11 WLAN using OMNeT++," 2022 IEEE 10th Region 10 Humanitarian Technology Conference (R10-HTC), Hyderabad, India, 2022, pp. 355-359, doi: 10.1109/R10-HTC54060.2022.9929505.

Authors Introduction

Hazwani Zawawi

She received Diploma in Electrical Engineering (Electronics) in 2016 from University Teknologi Mara. Completed her Bachelor in Electronic Engineering in 2020 from UiTM Shah Alam.Previously worked as Electrical Engineer at Sony EMCS (M) Sdn. Bhd since 2020 until Sept 2022. At Oct 2022 until present,she works at SIRIM QAS International San. Bhd. as an Auditor under Product Certification Scheme for electrical products

Dr Wan Norsyafizan W.Muhamad

She received her Bachelor in Electrical Engineering from the Universiti Malaya, Malaysia in 2002. She completed her Master of Electrical Engineering from Universiti Malaya, Malaysia in 2009. She obtained her Ph. D in 2017 from University of Newcastle, Australia. Currently, she is a Senior Lecturer in Faculty of Electrical Engineering at Universiti Teknologi MARA, Shah Alam, Malaysia. Her current research interests are in the area of wireless communication (Physical and MAC cross layer optimization)

Dr Suzi Seroja Sarnin

She received the B,Eng degree in Electrical and Electronics engineering. She worked as a quality control engineer for two years, from 1999 to 2001, at Memory Tech (M) Sdn. Bhd. She received Master in Microelectronics at University Kebangsaan Malaysia, in 2005. She studied for a PhD, which she obtained in 2018 with a degree in Electrical Engineering. Her research interest includes Wireless Communication, Multiple Access, Space Time Coding and Coding Theories, Signal Processing and Internet of Things.

Dr Nani Fadzlina Naim

She received the B.Eng. degree in electrical and electronics engineering and the M.Eng. degree in electronics and communication engineering from Universiti Teknologi Malaysia, in 2005 and 2007, respectively. She holds a PhD in optical communication from Universiti Kebangsaan Malaysia (UKM) in 2015. Currently, she is a senior lecturer at Universiti Teknologi MARA (UiTM), Malaysia. Her current research interests are in the areas of optical communication, Fiber Bragg Grating based sensors, spectroscopy, optical sensing technology and its application.

Data Transmission by Li-Fi in Coal Mining

Myadadha Goutham Reddy[1], Tan Jia Wei[1], Mandigiri Golla Jagan Kumar[1], Vytla Poojitha[1], Kaki Harshavardhan Goud[1], Wang Ji[1], MHD Amen Summakieh[1], Atefeh Mohamadpour[2], Mastaneh Mokayef[1*]
[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]*Department of Construction Management,*
California State University,
Sacramento, CA, USA

Abstract

There are many dangerous threats in a coal mining industry such as gas explosion, and temperature and humidity variation which can be difficult to be monitored and controlled. In this paper, we proposed and designed a method to utilize the Li-Fi technology for data transmission which are equipped with sensors to provide real time situation stat to the coal miners to ensure their safety. One of the most dangerous gases in coal mining is the carbon monoxide (CO). Hence, specific type of sensor is used and connected to a core component which is the microcontroller and placed at different parts of the coal mine. When the concentration level exceeds a certain threshold value, the buzzer will be triggered, and the notification is sent to the central database to inform the workers. Li-Fi provides the solution for slow data transmission or data loss which enhances the safety and improve the working condition of labors in coal mining sectors.

Keywords: Li-Fi, microcontroller, carbon monoxide sensor, safety.

1. Introduction

Safety is one of the most important aspects in all industries, especially in the coal mining industry. The reason is because we cannot guarantee zero error in the mining process of coal mines, as even a small mistake may lead to disastrous consequences. Communication is the primary key to observing any risk parameter [1]. At present, due to the complexity of coal mining environment around the world, mining workers are affected by many accidents. Therefore, it is essential to monitor the working environment which can be improved by establishing flexible communication with the workers. In the process of mining, the wire communication system is inefficient as the installation cost is high, and it is easily damaged under the influence of natural disasters such as earthquakes. The wireless communication used in this paper which is Li-Fi is a good way to minimize these problems.

2. Literature Review

Li-Fi technology is an optical wireless communication (OWC) technology proposed by German physicist Harald Haas in 2011 [2]. Li-Fi uses the visible part of the electromagnetic spectrum to transmit information at a very high speed. The working principle of Li FI begins with visible light as the emission source of the signal by controlling the ‘on’ and ‘off’ state of the LED light. When the light is ‘on’, the digital signal represents "1", and when the light is ‘off’, the digital signal represents "0". The LED light emits high-speed light and dark flashing encoded information that cannot be detected by the naked eye. The photosensitive sensor receives these changes and uses the decoding chip to recover the same data information as the sender, so as to complete the transmission and reception of wireless data and translate the optical signal into ordinary electrical signals. Currently, the improved LED light is equivalent to a Li-Fi hotspot. The optical signal emitted by LED lights is similar to the electromagnetic signal emitted by AP (Wi Fi hotspot) devices [3]. To put it simply, it uses light to transmit the network, and when there is light, there is network.

3. Proposed System

The entire system is divided into two parts which are the transmitter and receiver and both plays a crucial role in communication between the components. Fig. 1 shows the block diagram of the transmitter and Fig. 2 shows the block diagram of the receiver.

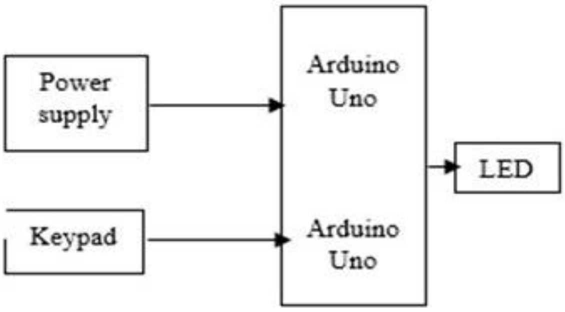


Fig. 1. Block diagram of device at transmitter side

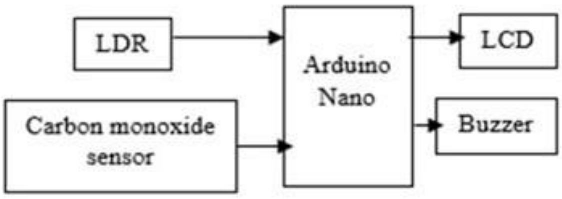


Fig. 2. Block diagram of device at receiver side

The main components that were used to ensure the functionality of the operation design system are listed below:

3.1. *Arduino uno*

UNO
ARDUINO

Fig. 3. ATmega328P Arduino Uno

The ATmega328P-based Arduino Uno shown in Fig. 3 is a microcontroller board. It has 14 digital I/O pins, 6 analogue inputs, a ceramic resonator operating at 16 MHz, a USB connection, a power jack, an ICSP header, and a reset button. The microcontroller can start to function by simply connecting it to a computer via a USB cable or power it via an AC-to-DC adapter or battery. As an output, this board is capable of controlling relays, LEDs, servos, and motors and can be interfaced with other Arduino boards, Arduino shields, and Raspberry Pi boards.

3.2. *Arduino nano*

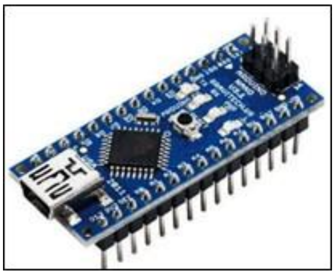


Fig. 4. ATmega328 Arduino Nano

The Arduino Nano illustrated in Fig. 4 is a small, complete, and breadboard-friendly ATmega328-based board. It has similar functionality to the Arduino Duemilanove, but in a different package. It only lacks a DC power jack and uses a Mini-B USB cable rather than a standard one. The Arduino Nano includes a crystal oscillator with a frequency of 16 MHz. It is used to generate a clock with a precise frequency by using constant voltage.

3.3. *MQ-7 sensor*

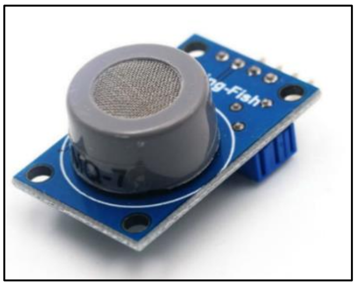


Fig. 5. MQ-7 Gas Sensor

The MQ-7 shown in Fig. 5 is a simple carbon monoxide sensor that can detect CO concentrations in the air. The MQ-7 is capable of detecting CO-gas concentrations ranging from 10 to 500ppm. This sensor is highly sensitive and has a quick response time. The output of the sensor is an analogue resistance. The drive circuit is very simple: we simply supply 5V to the heater coil, add a load resistance, and connect the output to an ADC.

3.4. *Keypad*

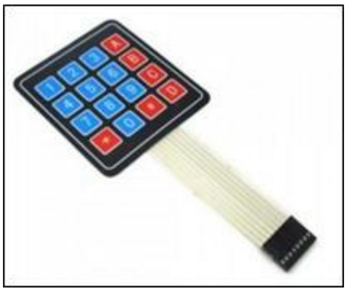


Fig. 6. 4×4 4×4 Matrix Array/Matrix Keyboard

This keypad has 16 buttons organised in a telephone-line 3×4 grid, plus four extra keys labelled A, B, C, and D as shown in Fig. 6. It's made of a thin, flexible membrane material with an adhesive backing that allows it to be attached to almost anything. Because the keys are connected in a matrix, scanning through the pad requires only 8 microcontroller pins (4 columns and 4 rows).

3.5. *Buzzer*

A buzzer, also known as a beeper, is a mechanical, electromechanical, or piezoelectric audio signalling device. Alarm devices, timers, train and confirmation of user input such as a mouse click or keystroke are common applications for buzzers and beepers. Fig. 7 shows the Piezo Buzzer used in this prototype.

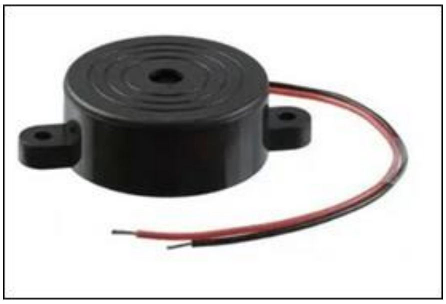


Fig. 7. MCKPR3-G4210-4136 Piezo Buzzer

4. Development of Li-Fi transmission

This section highlights the procedure used to establish a Li-Fi connection for data transmission usage in the mining sector. The project is divided into four different stages which are planning, designing, execution and monitoring to produce the best possible outcome.

4.1. *Planning the suitability of materials and equipment*

There are two types of sources which are widely used, Light Emitting Diode (LED) or lasers depending on the types of propagation mode [4]. Since our project is on a smaller scale, the basic material used as the light source is the LED. Next, Light Dependent Resistor (LDR) is used for light pulses detection on the receiver side. The core of the entire system relies on the microcontroller which is the Arduino Uno and Nano as it is more cost effective and capable of interpreting data such as reading inputs from a sensor and providing an output to another specific device.

Next, carbon monoxide which is produced from incomplete combustion of carbon dioxide is an important risk parameter in the mining environment because it is extremely toxic. It may even cause the death of a person in one to two hours if the level gets as high as 0.2% because it blocks the hemoglobin from absorbing and carrying oxygen. It is also flammable and difficult to extinguish compared to other gases [5]. MQ-7 is used to detect the carbon monoxide level and a buzzer is fixed at the receiver to inform the workers once the concentration exceeds the threshold value. There is also a keypad at the transmitter side to ease communication during crucial moments.

Besides that, a Liquid Crystal Display (LCD) is used at the output to display the current carbon monoxide level of the surrounding environment as well as the important data from the keypad input.

4.2. *Designing data transmission via Li-Fi*

For such a small-scale design, all the devices are fixated onto breadboard. When the user presses a number on the keypad, the Arduino Uno interprets the information and sends the data via the output port to the LED. The LED is connected in series with a resistor to form a voltage divider circuit and to avoid short circuit. The microcontroller controls the blinking of LED which is detected by the LDR as visible light pulses and converts it to interpretable electrical pulses [6]. Then, the LDR is connected to the input port receiver, Arduino Nano, which receives the pulses and converts it to actual data. The MQ-7 sensor which is used to detect the surrounding carbon monoxide collects the data and is connected at the receiver side of the Arduino Nano. The buzzer will be switched on once the sensor detected a value above 100ppm. The LCD will also display the value of carbon monoxide concentration and the keypad value from the input.

4.3. *Execution of devices in mining sectors*

The project was designed to be implemented as a real-life application in the mining sector. First, the carbon monoxide sensor is located at different places and levels in the mining operations to have a wider coverage. Next, the sensors are connected to the embedded system which is fixated at the central database for data analysis and interpretation. LEDs are replaced with white light and built along the mining routes. LDR are fixated onto the workers' helmets to ease the detection of light. The data will be displayed on a wrist band worn by the workers which comes with an alarm that produces sound and notify the workers if the concentration of carbon monoxide exceeds the threshold value. Besides, any important announcements that needs to be transmitted to deeper underground levels can be done by pressing the numbers on the keypad from 1-9 at the transmitter side which will be received by all the workers immediately.

4.4. *Monitoring*

The device needs to be monitored from time to time because some environmental issue may cloud the judgement of the MQ-7 sensor and may require maintenance to ensure high accuracy as well as to protect the lives of the workers in the mining sector. Besides that, nowadays the use of LED is replaced with lasers because they are more powerful, operate at high speeds due to its monochromatic characteristic, allows further transmission of light and has fewer errors.

4.5. *Flowchart of the device*

The procedure of the working principle of Li-Fi data transmission follows a designated flow as shown in Fig. 8 below.

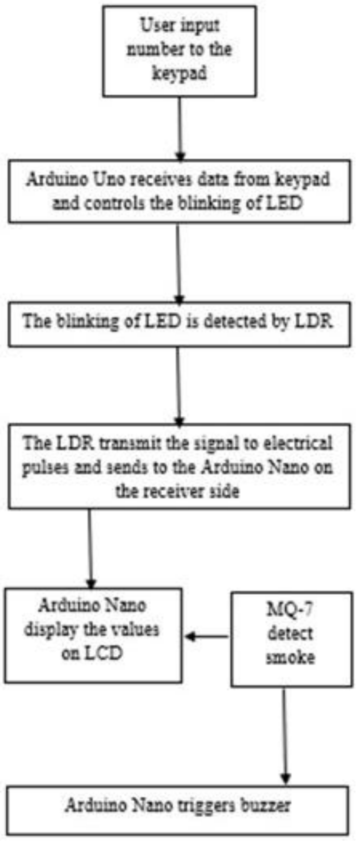

User input number to the keypad
Arduino Uno receives data from keypad and controls the blinking of LED
The blinking of LED is detected by LDR
The LDR transmit the signal to electrical pulses and sends to the Arduino Nano on the receiver side
Arduino Nano display the values on LCD
MQ-7 detect smoke
Arduino Nano triggers buzzer

Fig. 8. Flow chart of data transmission

5. Results

The result of the experimental setup from the transmitter to receiver was recorded.

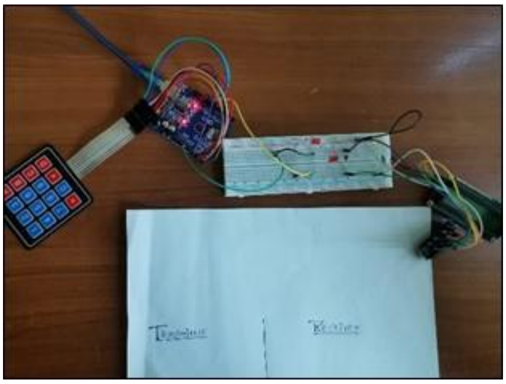


Fig. 9. Hardware setup of data transmission via Li-Fi

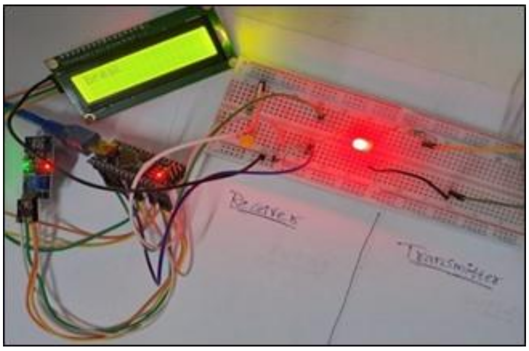


Fig. 10. Data transmission via LED to LDR when keypad is pressed

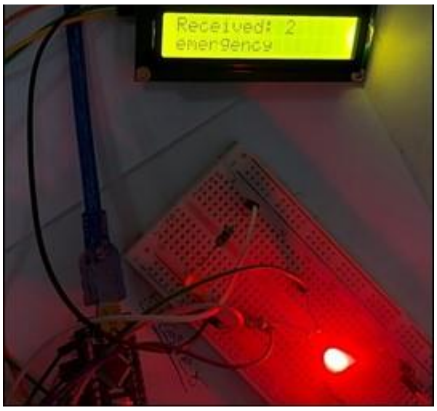


Fig. 11. Input data displayed on LCD screen

6. Discussion

Li-Fi is a cutting-edge communication that has the potential to be 100 times quicker than Wi-Fi and uses visible light sources to transmit data which acts as a communication medium. The photodiode serves as a receiver that receives the light signals, while an LED works as a light source. We can communicate specific data patterns by managing the light pulse on the transmitter side. The data is then transformed into meaningful information at the receiver side by the photodiode or Light-dependent resistor (LDR) [7].

As depicted in Fig. 9 the hardware setup of data transmission via Li-Fi has the following steps; when the numbers in the keypad are pressed, the LED lights up. As the numbers in the keypad goes higher, the intensity of light also increases. LDR at the receiver side detects the change in light intensity and sends the signal to Arduino Nano which is connected to an LCD. The LCD module displays the number input in the keypad based on the difference in light intensity detected by LDR. Different time delay was set for the LED to remain at the 'on' state as shown in Fig. 10. For instance, when the user input the number '2' from the keypad, the LED lights up dimly and only for a short amount of time. The LCD module display the number '2' at receiver side as shown in Fig. 11. If the user input the number '8', then the LED lights up with much brighter intensity and remains 'on' for a longer time compared to the first number. This is because an increase in incident light's intensity translates into an increase in photon content. When the quantity of incident photons rises, more photo- electrons are likewise released, which raises the photoelectric current.

Carbon monoxide sensor, MQ-7 is used at receiver side to detect the concentration of this particular gas emitted at the surroundings. This sensor has a small heater inside with electrochemical sensor to measure different kinds of gas combinations. The working principle of MQ-7 depends on chemiresistor which has free electrons. These free electrons are attracted to the oxygen molecules and push to the surface of the chemiresistor. This prevents the free electrons from conducting current which helps in the detection of CO [8]. As the gas concentration increases, so does the sensor's conductivity. The sensor may be used to determine which gases contain carbon monoxide and protects the coal miners' health by triggering a buzzer when a threshold value is surpassed.

The keypad serves as input for Li-Fi communications at the transmitter portion. This implies that the keypad will be used to select the text to be conveyed. The control unit, Arduino Uno, processes the input and transforms the data into binary pulses that can be transmitted via an LED source. The LED light uses these data to transmit pulses of visible light to the receiver side. The LDR sensor in the reception part absorbs the visible light pulses and turns them into comprehensible electrical pulses that are fed to the Arduino Nano (Control unit) which transforms it into data and displays it on a 16x2 LCD screen.

6.1. *Advantages of Li-Fi*

First, Li-Fi can be expanded in places where Wi-Fi technology is not yet widely used, such as in hospitals, power plants, and other establishments. Besides, Li-Fi only uses light, making it safe to use in places like aeroplanes and hospitals where Wi-Fi is prohibited due to its propensity to interact with radio frequencies. There is also no health concern because optical bands are not harmful unlike the Radio Frequency spectrum. Devices that use Li-Fi have lower power consumption for operation and hence is used in a lot of IoT applications. Lastly, due to line-of-sight data exchange, it offers a high level of security (LOS). Additionally, the Li-Fi signal cannot penetrate through walls. This will prevent unauthorised individuals from gaining access to the signal [9].

6.2. *Advantages of Li-Fi*

In order to receive data, there must be a perfect line of sight. This means that the signal is immediately lost if the receiver is occluded by any opaque object. Furthermore, currently, only a few gadgets are compatible with Li-Fi because it is a relatively new technology. It's doubtful that we will see Li-Fi-enabled personal gadgets in the next several years because the majority of the devices we use now still utilise hardware for Wi-Fi networking. Although Li-Fi has a quicker data transmission rate, the best feature of this technology is still unimportant if service providers' Internet speeds are still poor. Deploying a Li-Fi network would be useless in nations with slower internet connections. Coordination between a number of industries will be necessary to promote widespread adoption of this technology [10, 11].

7. Further Improvement

The present work focuses mainly on unidirectional data transmission. However, it can be improved further by using another LDR at the transmitter module and LED at the receiver side which will lights up when the buzzer is triggered. With this, the Li-Fi system can now be bi-directional. In the bidirectional communication system, each microcontroller acts as a transmitter and receiver simultaneously. Considering this case, the user at transmitter module (coal miners working closer to surface ground) can send data to the user at receiver module (deep underground) and vice versa.

8. Conclusion

Li-Fi has evolved into a ubiquitous system technology that has a wide range of application. Due to its abundance advantages, Li-Fi is capable of improving the safety in many industries especially in coal mine sectors. This paper proposed a method of incorporating such technology to help in efficient monitoring and data transmission which reduce life risks of the coal miners. Carbon monoxide is the prime focus among all other risk factors because it is highly poisonous and can cause death in extreme cases. Li-Fi technology enables the coal miners to understand about the condition of environment at high reliability. It is also cost effective compared to Wi-Fi because it offers significant savings in the long run. However, since the technology is still new and under development, various research and development methods have to be conducted to ensure that the Li-Fi can meet the demand for connectivity.

References

1. Naidu, K. P. S. S. V., Visalakshi, P., & Chowdary, P. C, “Coal mine safety system using Li-Fi technology”, International Journal of Advance Research, Ideas and Innovations in Technology, Vol 5, Issue 1288-1291, 2019
2. Professor Harald Hass, Chair of Mobile Communications”, The University of Edinburgh, 2018.
3. J. Condliffe, “Is Li-Fi Ready to Establish Itself as the New Wi-Fi?” *New Scientist*, Vol. 211, No. 2822. p.18, July 2011. [Online].
4. N. Yeasmin, R. Zaman, I.J. Mouri, “Traffic Control Management and Road Safety using Vehicle to Vehicle Data Transmission Based on Li-Fi Technology”, International Journal of Computer Science, Engineering and Information Technology, IJCSEIT, Vol 6, Issue ¾, Aug 2016.
5. “The Most Dangerous Gases in Mining”, Howden Articles, 2022.
6. P. Jadhav, S. Khatib, and K. Maner, “Data Transmission Through Li-Fi”, International Research Journal of

Engineering and Technology, IRJET, Vol 04, Issue 02, Feb 2017.

7. Jayarajan, P., Gayathri, K. V., Gowshikha, S., & Harini, N, “Improved Cost Effective IoT Based Coal Mining Safety System”, IOP Publishing, Journal of Physics: Conference Series, Vol. 1916, No. 1, p. 012060, May 2021.
8. “MQ7 Carbon Monoxide (CO) Gas Sensor Module”, Circuits DIY, 2022.
9. “Benefits of Li-Fi”, University of Strathclyde Glasgow, 2022.
10. “Advantages and Disadvantages of Li-Fi”, GeeksforGeeks, July 2022.
11. Othman, R. A. A., ap Sagaran, D., Mokayef, M., binti Wan, W. I. I. R., & Nasir, M. (2018, December). Effective LiFi communication for IoT applications. In 2018 IEEE 4th International Symposium in Robotics and Manufacturing Automation (ROMA) (pp. 1-4). IEEE.

Authors Introduction

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. Atefeh Mohammadpour

She received her Ph.D. degree in Architectural Engineering (Construction Management Option) from the Pennsylvania State University, Master of Science degree in Construction Management from the University Technology Malaysia, and Bachelor of Engineering degree in Civil Engineering from Persian Gulf University. She has over seven years of experience in the manufacturing and construction industry. She is currently an Assistant Professor in the Department of Construction Management at California State University, Sacramento. Her interdisciplinary research focuses on innovative approaches to using artificial intelligence, various aspects of sustainability, and preventive measures to improve safety in the construction industry.

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.

Embedded Table Tennis Ball Launcher with a Trajectory Path Analyser for Junior Players

Mastaneh Mokayef[1*], **Lee Qi Jian**[1], **Bushra Naeem**[2], **Miad Mokayef**[1], **M.K.A Ahamed Khan**[1], **MHD Amen Summakieh**[1]
[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]*Faculty of ICT, Balochistan University of Information Technology, Engineering,*
and Management Sciences (BUITEMS), Pakistan

Abstract

Table Tennis as one of the most popular sports in Asia that its training requires a close supervision of a coach to analyse the strength and the weakness of the player. In this project, a pitcher machine as a personal Table Tennis trainer has been designed and the trajectory path analysis of the practical performance of player is analysed accordingly. In the proposed system, a camera is presented to track the table tennis ball. The obtained images from the camera are processed by OpenCV software and the flying trajectory is predicted based on the X and Y coordinates of a ball position to analyse the performance of player in each direction. The pitcher machine is then set via smart phone to the direction with weak performance to shoot more balls in that direction. However, the spinning serve is yet to be added in the proposed system, this project offers an acceptable platform for the early stage of the table tennis training. Experimental results show the acceptable performance analysis in complex mess environment and background.

Keywords: Table tennis launcher, performance analysis, trajectory path, ball recognition.

1. Introduction

Table Tennis is one of the most popular physical activities that involves two or four players who hit a ball in sequence. Like any other sport, in table tennis also coaching has an important role in the performance of the player before and during the competition. In this regard, several research works have been done on the performance improvement of the ping pong players. The integration of sensors and devices provided more efficient analysis of the athlete performance in different aspects such as fitness, mental health, gender, and blood lactate level [1]-[6]. Moreover, with the rapid rate of change in integrated technology and development of robotics, the table tennis robots have been designed and fabricated and the trajectory generation of the ball have been developed and optimised by time [2, 7-9]. The aggressive trajectory generator was introduced in Ref. 10 to control the robot in which the sensors have vital role. The trajectory prediction of the ball spinning is evaluated in Ref. 11 to derive the physical behaviour of the spinning ball. Other trajectory prediction and generation methods were discussed and proposed [12-14]. In this paper however, the table tennis launcher and performance analyser are proposed, designed and fabricated. The motivation of this work is to improve the solo ping pong trainers by analysing the trajectory path analysis and shoot more balls in the weak performed direction.

2. Hardware Architecture and Experimental Setup

The design of the proposed embedded Table Tennis ball launcher with trajectory path analyser is based on the operation of a ball launcher attached at the end of the table tennis table and the camera attached to the laptop beside the table for the path analysis. The entire launching system is placed on a wooden frame stand. Hence this part is a backbone of the system and need a proper stable design. The actuator mechanism is based on the accurate shoot of only one ball at each time and the shooting trigger. To meet these requirements, a modified plastic basket is used as a ball container. The basket with half-circle shaped ball holders' disc is then mounted on a motor. The combination of motor and the disc allows the full 360-degree rotation of the launcher. A hole has been made at the bottom of the wall-side of the basket to attach the long handlebar to. the second shorter handlebar with a degree of freedom is attached to the longer handlebar to provide ease of attainability and manoeuvrability. As motor rotates, the propeller shaped disk that holds the balls will rotate and places the balls to the track of shooting. The speed of shooting can be controlled by varying the motor speed. The ping pong balls travel from the hopper propeller shaped disc into the first horizontal track. The second track is attached with 30 degrees of freedom to the first track to provide different shooting direction. The side, front, and top views of the prototype are shown in Fig. 1.

Due to the cost restrictions, and getting align with the Sustainable Development Goals (SDG), all recycled materials have been used in the proposed prototype.

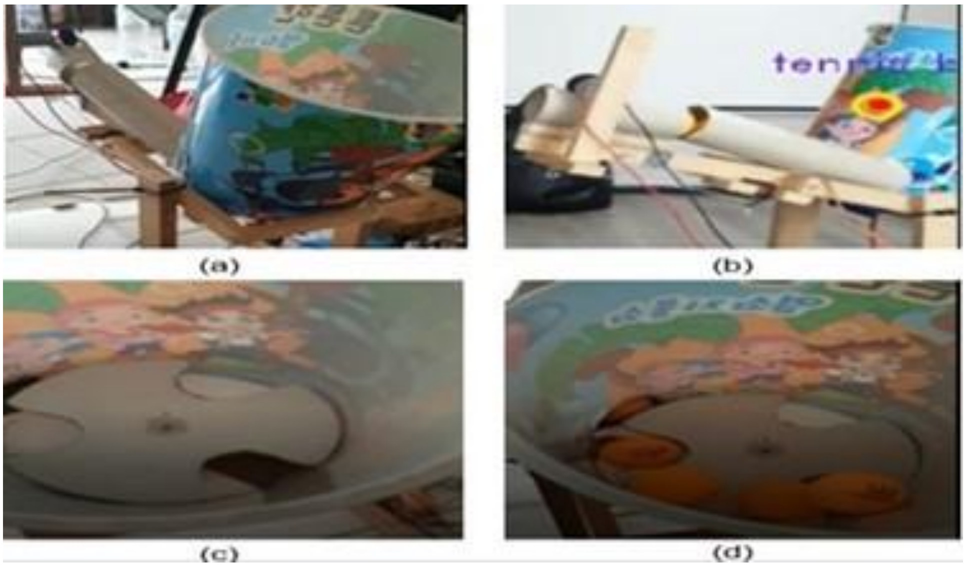


Fig. 1. The (a) Side view, (b) Front view, (c) Top view of the ball (d) Top view with the balls preparing to shoot

The first and the most important step in this project is to detect and track the ball. This process requires some additional steps to boost up the accuracy of the detection technique. For this purpose, the RGB color has been converted to HSV to describe the colors similar to those human eyes can detect. For this purpose, the manual tracking or Detection-free trackers has been used in OpenCV. This has been decided to prevent pre-training of the object detection. To filter the required results, the trackbarPos code is used. This helps to generate the track bar and adjust the required HSV levels. The result of track bar code on changing the image threshold in runtime is illustrated in Fig. 2.

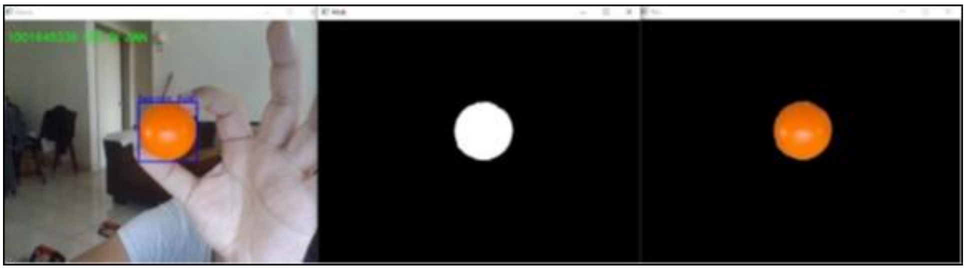


Fig. 2. The HSV level adjustment based on the TrackbarPos

The location estimation of the ball requires a technique to measure the relationship between the physical dimensions of the ball and the practical area of counter, and the pixels that the ball appears in. For this, the contour area measurements have been setup as shown in Fig. 3.

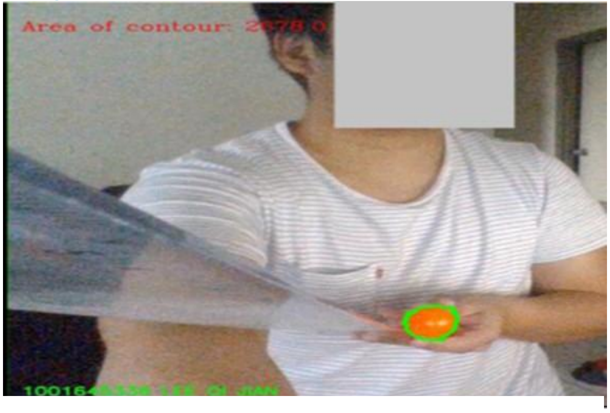


Fig. 3. The contour area measurements

In practice, there are several ways to estimate the location of ping pong ball. In this project, we have derived an equation to measure and detect the position of the ball based on the counter area. This equation is described as per Eq. (1):

$$y = -4.963453 + \frac{5296.735 + 4.963453}{1 + \left(\frac{x}{0.09193972}\right)^{0.4415413}} \quad (1)$$

To achieve that, the camera is placed at a fixed position and the separation distance between the area of the contour and the camera is plotted in excel. With the help of excel function and the Python code, equation (1) is generated to estimate the separation distance. The Trendline between area and distance of contour is plotted in Fig. 4.

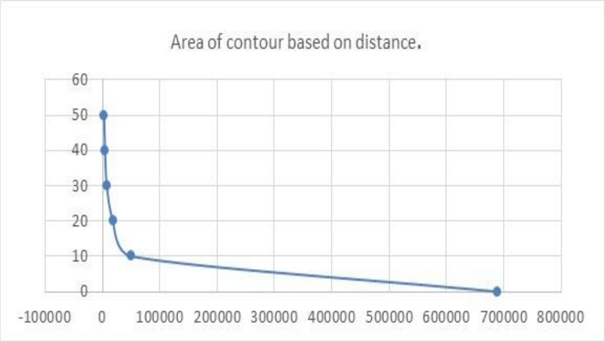


Fig. 4. The Trendline between area and distance of contour

To test the accuracy of the measurement and the contour measurement function using OpenCv, a physical measurement has been conducted as shown in Fig. 5. The comparison illustrates the acceptable accuracy.

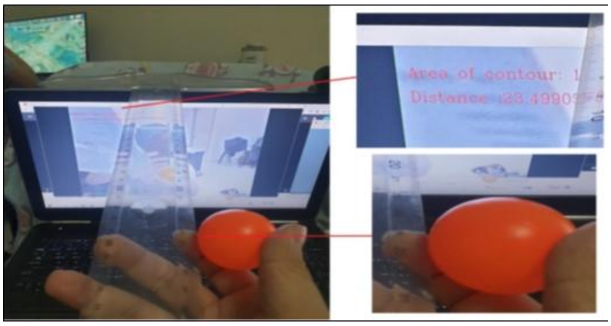


Fig. 5. Distance measurement accuracy check

The accurate result has been proved by comparison between physical and virtual measurement done by ruler and OpenCV measurement. To efficiently store and trace the movement of the ball in different points, the ball has been first detected and each tracked point is saved by the deque datatype in Python is used. Hence the history points of, f(x,y) has been recorded and saved and longer trajectory paths has been created considering more ball points with processing and resizing the frames to achieve the most suitable width and height of the moving ball that is later used for the simulation analysis. In the next step, the image blurring or image smoothing has been implemented in which, the Gaussian blur filter has been applied to the detected ball image. In this stage, the image is convolved with Low Pass Filter (LPF) Gaussian kernel, that resulted to the noise reduction of the image. With the help of Gaussian blur, the Laplacian of the source image is calculated by calculating the gradient of image intensity at each pixel within the image utilizing soble operator; hence, the computed response is stored in the destination image at the same location (x,y) and is calculated as:

$$dest = \frac{\partial^2 src}{\partial x^2} + \frac{\partial^2 src}{\partial y^2} \tag{2}$$

where src is a source image and dest is the destination image.

In this project, the HSV value of the source image has been extracted from OpenCV trackbar and then filtered out. Moreover, a mask has been constructed for the ball color detection. In this project, an orange ball mask has been implemented. The biggest contour of the mask determines the ping pong ball.

To determine that if the ball is hit or not, the value of the y-coordinate is considered. If the value of y-coordinate is negative, it shows that the ball has been shoot towards the player, and if the y-coordinate gets positive, this means that the player has hit back the ball and consequently, the trajectory path is turned back. The result of the y-coordinate value has been tested and shown in Fig. 6. Based on the result of the Y-coordinate value, the performance of the player can be determined. To have a reasonable performance measurement, a simple analysis on the X and Y coordinates of the ball tracking has been implemented. In this case, the Y-coordinate 150 values have been considered, if the overall value is negative, this means that the player did not counterplay the ball and hence the performance will be marked as weak. Once the weak performance is detected, the trainer will be set to shoot the ball to the poor performed direction. This is done by using the pitcher machine for launch the ping pong ball and shoot on the direction that is controlled by trainer, the coach, or even the player. Moreover, the IR remote control for controlling the pitcher machine and next is draw the trajectory of ping pong ball has been implemented to monitor the player's playing style.

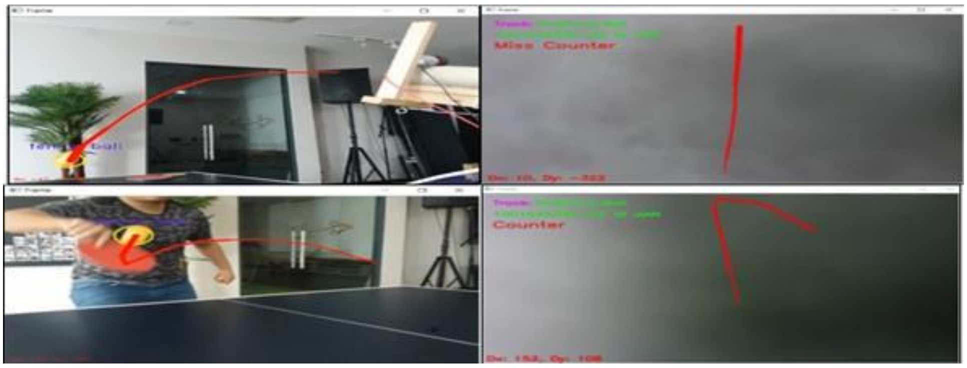


Fig. 6. Trajectory path calculation based on Y-coordinate values

3. Results and Discussion

As a pre-processing step towards the result analysis, the image thresholding is utilized in this paper. This technique is used to remove the un-necessary parts of image and consequently, produces a more accurate image of a ping pong ball. The image thresholding is known as the simplest image segmentation technique in which, the image is divided into two groups of black and white to categorize the background and foreground respectively. For this step to be done, the regions with saturation of greater than threshold level is translated to white region and the rest remain as black. The threshold selection is done manually in this project. The result of threshold selection for both background/foregrounds, and the threshold value are illustrated in Fig. 7.

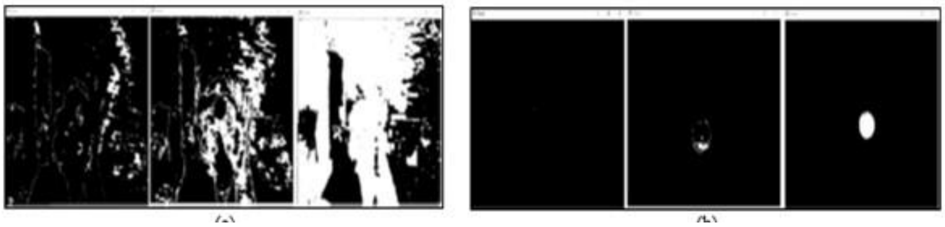


Fig. 7. Threshold value selection for (a) background/foreground, (b) level adjustment from 0 to 255

The object overlapping scenario is also considered as a false condition in this project. In case of any overlapping of ping pong balls, at the current stage of our work, the ball detection is failed, and the result is not counted in the analysis. Consequently, there will be an issue of ball tracking that will be tackled in our future work.

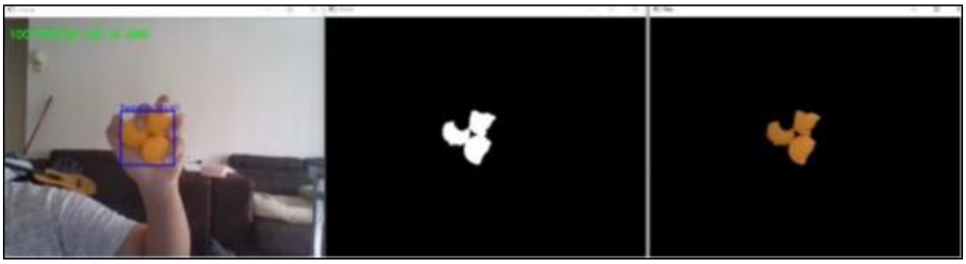


Fig. 8. Tracking failure due to the objects overlap

The third Python file has been created to draw the trajectory and provide the movement. To increase the accuracy of the ball detection, the counter number is limited to one counter at each time. For instant, if two ping pong balls are detected in the screen, the one with the larger radius would be detected as the main ball. The tracking failure due to the objects overlap has been implemented and illustrated in Fig. 8. This auto ball selection and the tracking of only one object at the time has been illustrated in Fig. 9. Plenty of stuffs has been added to the screen to ensure the accuracy of the ball detection in a crowded environment.

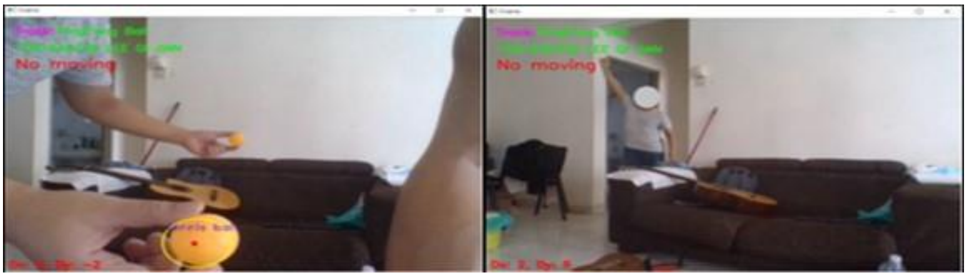


Fig. 9. Object detection in presence of two balls in a crowded area

The color-based object detection has some disadvantages such as the background tracking interruption in which, the ping pong ball tracking will be lost due to the similarity of the background color with the ping pong ball. This will reduce the accuracy and therefor, affect the analysis dramatically. Fig. 10 is a proof of the tracking interruption due to the similarity of the background and the ball. As per depicted in Fig. 10, the yellow ping pong ball tracking is lost at the moment the ball is passing by the container with similar color background. This issue is considered in this project and the project enhancement in the object detection will be tackled in our future work.

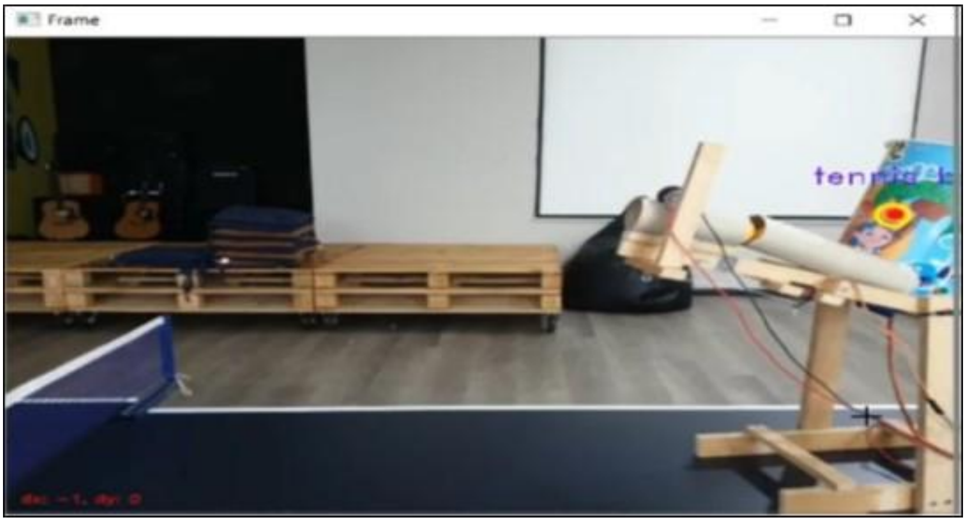


Fig. 10. Tracking interruption due to the color similarity with the background

4. Conclusion

The table tennis launcher with a real time ping pong ball tracking to train the player based on their performance analysis, has been designed and fabricated. This project is based on the operation of a servo motor that is located under the modified basket and shoots the balls on the table tennis board upon its rotation. The color-based segmentation has been used to detect the ball. After the segmentation is done, the shape, location and the motion of the ball are the factors considered to identify the performance of the player. The OpenCV software is used for the ping pong ball tracking. The player performance is then analysed based on the players trajectory path calculation in each point based on the ball traveling distance. Accordingly, the weakness of the player can be found, and the smart table tennis trainer uses a pitcher machine to launch the ping pong ball set to the direction in which the player performed weak and shoots more balls in that specified direction. This in other hand will help the player to enhance his performance in the specific direction and monitor the weakness of the player and help the improvement in his performance.

References

1. Arnold Baca (2006) Innovative diagnostic methods in elite sport., International Journal of Performance Analysis in Sport, 6:2, 148-156, DOI: 10.1080/24748668.2006.11868380
2. Z. Zhang, D. Xu and M. Tan, "Visual Measurement and Prediction of Ball Trajectory for Table Tennis Robot," in *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 12, pp. 3195-3205, Dec. 2010, doi: 10.1109/TIM.2010.2047128.
3. Pradas, Francisco, et al. "Analysis of Specific Physical Fitness in High-Level Table Tennis Players—Sex Differences." International Journal of Environmental Research and Public Health 19.9 (2022): 5119.
4. Picabea, Jon Mikel, Jesús Cámara, and Javier Yanci. "Physical fitness profiling of national category table tennis players: Implication for health and performance." International Journal of Environmental Research and Public Health 18.17 (2021): 9362.
5. Mokayef, Miad, Mehrzad Moghadasi, and Reza Nuri. "Effect of cold water immersion on blood lactate levels of table tennis players." Int. J. Curr. Res. Aca. Rev 2.9 (2014): 115-123.
6. L. Acosta, J. J. Rodrigo, J. A. Mendez, G. N. Marichal, and M. Sigut, “Ping-pong player prototype,” IEEE Robot. Autom. Mag., vol. 10, no. 4, pp. 44–52, Dec. 2003.
7. Yang, Luo, et al. "Ball motion control in the table tennis robot system using time-series deep reinforcement learning." IEEE Access 9 (2021): 99816-99827.
8. Zhao, Hongtu, and Fu Hao. "Target tracking algorithm for table tennis using machine vision." Journal of Healthcare Engineering 2021 (2021).
9. Zhang, Bo, Bingqiang Chen, and Yansong Deng. "Robot arm trajectory planning study for a table tennis robot." E3S Web of Conferences. Vol. 260. EDP Sciences, 2021.

10. Cong, Vo Duy. "Real-time measurement and prediction of ball trajectory for ping-pong robot." 2020 5th International Conference on Green Technology and Sustainable Development (GTSD). IEEE, 2020.
11. Wang, Yilei, and Ling Wang. "Machine Vision-Based Ping Pong Ball Rotation Trajectory Tracking Algorithm." Computational Intelligence and Neuroscience 2022 (2022).
12. Wang, Y. "Path trajectory prediction of rapidly rotated ping pong ball after hitting." The International Journal of Multiphysics 13.4 (2019): 351-360.
13. Andersson, Russell L. "Aggressive trajectory generator for a robot ping-pong player." IEEE Control Systems Magazine 9.2 (1989): 15-21.
14. Huang, Yanlong, et al. "Trajectory prediction of spinning ball for ping-pong player robot." 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2011.

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.

Dr. Bushra Naeem

She received the bachelor's degree in telecommunication engineering from the Balochistan University of Information Technology, Engineering and Management Sciences (BUITEMS), Quetta, Pakistan, in 2009, the master's degree in engineering and management from the University of Exeter, Exeter, U.K., in 2010, and the Ph.D. degree from the University of Technology Malaysia, Johor Bahru, Malaysia, in 2016. She is working in academia since 2010 and currently serves as an Associate Professor and the Head of the Software Engineering Department, BUITEMS. She published various scientific papers and delivered keynotes in international conferences and holds continued professional development programs at UAE, Pakistan, and Oxford, U.K. Her subjects of interest include wireless communications, heterogeneous networks, cognitive radios, and 5G. Dr. Naeem is a registered member of the Pakistan Engineering Council, IEEE, WIE, and the editorial board in several international journals.

Dr. M.KA. Ahamed Khan

He is now attached to UCSI University. His research interests are power system, power electronics and microprocessor/microcontroller.

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.

Table Tennis Tournament Scores and Statistics Web Application

Belal Khaled[1], **Mastaneh Mokayef**[1*], **Chin Hong Wong**[2], **Sew Sun Tiang**[1], **Wei Hong Lim**[1], **MHD Amen Summakieh**[1], **Miad Mokayef**[1]
[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]*Maynooth International Engineering College, Fuzhou University*
Fujian, China

Abstract

Table tennis is a popular sport that involves two or four players to hit a lightweight ball back and forth across a tennis table using small rackets. The scores are generated based on the tournament results. These scores were not digitized in the early days, therefore, the match referees had to record them manually while most of the competitors provided only basic information and real-time scores online. Over the years of technological advancements, the modern development of software applications has helped users perform useful tasks and retrieve useful data based on the requests. The objective of this project is to develop a functional web application that will retrieve or store table tennis tournament statistical data and visualize them using tables, bar charts, pie charts, histograms as a medium of demonstration for the user. Moreover, the application will perform specific analysis on the scores and display insightful data about the tournament for league structures investigation. The developed system will include a database to store and retrieve data for display on the user interface. The development of the project is fully stacked (front-end and back-end), so it is built with the appropriate web technologies to function in the background (PHP, MySQL, Apache) while displaying results (HTML, CSS, JavaScript) on the page for the user. The project is managed using effective project management methods to plan, design, implement, develop, and maintain the application.

Keywords: Table Tennis, Statistics, Web Application, Tournaments, Scores.

1. Introduction

The primary objective of this project is to develop a web-based application that provides statistical insights of table tennis tournaments scores by utilizing score values and converting them into useful data for the user. When users launch the application, the system will display tabulated data that are retrieved from the database. The application contains a built-in control panel for which administrators can log in and input tournament and match details such as their name, scores and the players who were involved. When browsing the application, there are several pages that perform specific functions. For instance, when navigated to a “tournaments” page, the system will display all tournaments that are held in the current year and in the previous years. On the tabulated results being displayed, the user can click on available hyperlinks located in each row of data to be redirected to another page that will show detailed results of the tournament. The results will include the date for which it was held, the tournament’s name, the players involved and the scores of each match.

Every table tennis match consists of variables that can be considered as data [1]. The web application utilizes MySQL database to store them as records which then can be added, modified, or deleted. The system is dynamic so the data being displayed on the web application will be updated automatically. The application uses PHP to communicate with the database.

In this project, an opensource software called "phpMyAdmin" will be used to help create the database and its corresponding tables [2]. It will also allow developers to execute MySQL operations such as importing data into the database [3]. Fig. 1 shows a sample of phpMyAdmin interface.

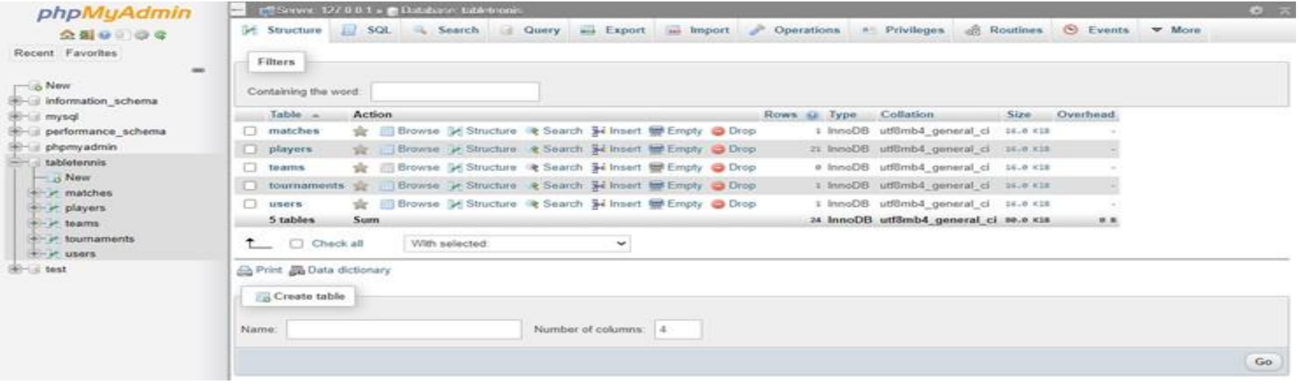
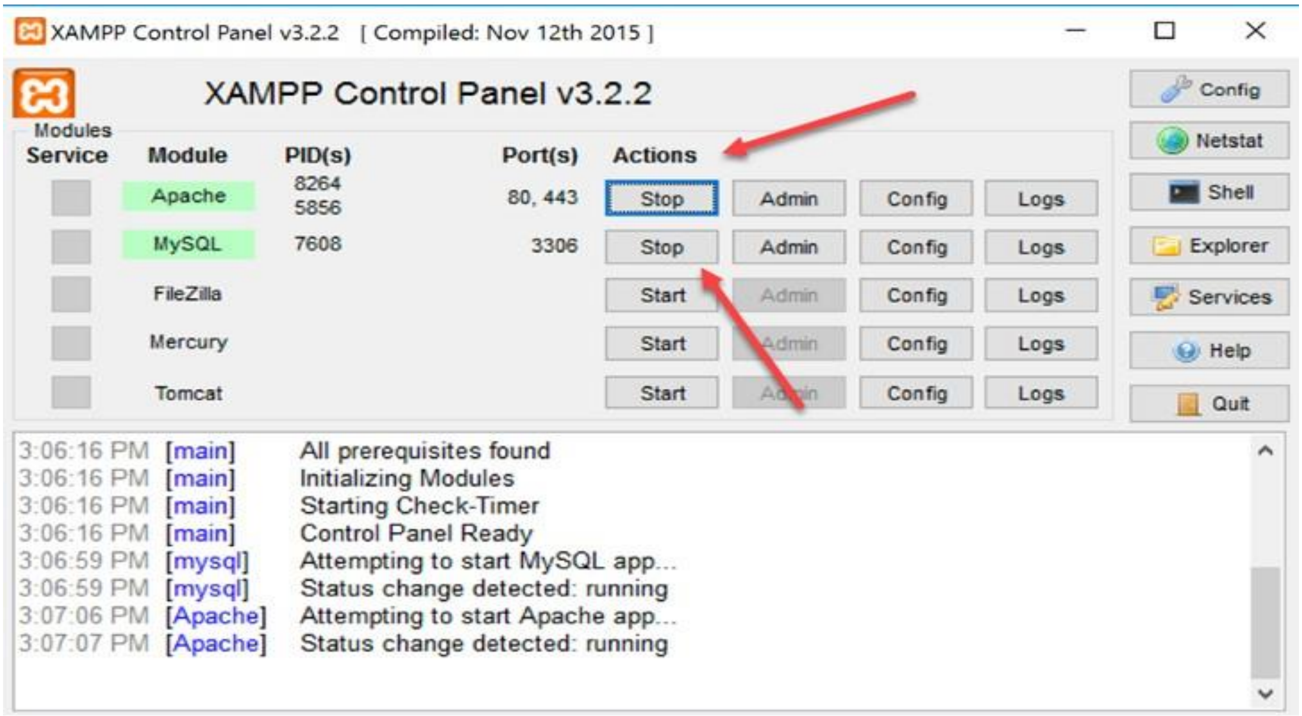

phpMyAdmin
Recent Favorites
New
information_schema
mysql
performance_schema
phpmyadmin
tabletennis
New
matches
players
teams
tournaments
users
test
Server: 127.0.0.1 » Database: tabletennis
Structure
SQL
Search
Query
Export
Import
Operations
Privileges
Routines
Events
More
Filters
Containing the word:
Check all
With selected:
Print
Data dictionary
Create table
Name:
Number of columns: 4
Go

	Table	Action	Rows	Type	Collation	Size	Overhead
☐	matches	Browse Structure Search Insert Empty Drop	1	InnoDB	utf8mb4_general_ci	16.0 KiB	-
☐	players	Browse Structure Search Insert Empty Drop	21	InnoDB	utf8mb4_general_ci	16.0 KiB	-
☐	teams	Browse Structure Search Insert Empty Drop	0	InnoDB	utf8mb4_general_ci	16.0 KiB	-
☐	tournaments	Browse Structure Search Insert Empty Drop	1	InnoDB	utf8mb4_general_ci	16.0 KiB	-
☐	users	Browse Structure Search Insert Empty Drop	1	InnoDB	utf8mb4_general_ci	16.0 KiB	-
	5 tables	Sum	24	InnoDB	utf8mb4_general_ci	80.0 KiB	0 B

Fig. 1. phpMyAdmin Interface

HTML is used as a markup language to setup the layout of the web page and CSS was used for styling each web component of the page. By using a combination of these web technologies, a table can be created to display the retrieved data to the user. While HTML and CSS are responsible for the front-end development, PHP facilitates in executing back-end operations by communicating with the database.


XAMPP Control Panel v3.2.2 [Compiled: Nov 12th 2015]
XAMPP Control Panel v3.2.2
Modules
Config
Netstat
Shell
Explorer
Services
Help
Quit

Service	Module	PID(s)	Port(s)	Actions			
	Apache	8264 5856	80, 443	Stop	Admin	Config	Logs
	MySQL	7608	3306	Stop	Admin	Config	Logs
	FileZilla			Start	Admin	Config	Logs
	Mercury			Start	Admin	Config	Logs
	Tomcat			Start	Admin	Config	Logs

```
3:06:16 PM  [main]    All prerequisites found
3:06:16 PM  [main]    Initializing Modules
3:06:16 PM  [main]    Starting Check-Timer
3:06:16 PM  [main]    Control Panel Ready
3:06:59 PM  [mysql]   Attempting to start MySQL app...
3:06:59 PM  [mysql]   Status change detected: running
3:07:06 PM  [Apache]  Attempting to start Apache app...
3:07:07 PM  [Apache]  Status change detected: running
```

Fig. 2. XAMPP Control Panel

2. Methodology

The importance in the usage of methodology in this project is the identification of dependent and independent variables based on the requirement of the user and the system.

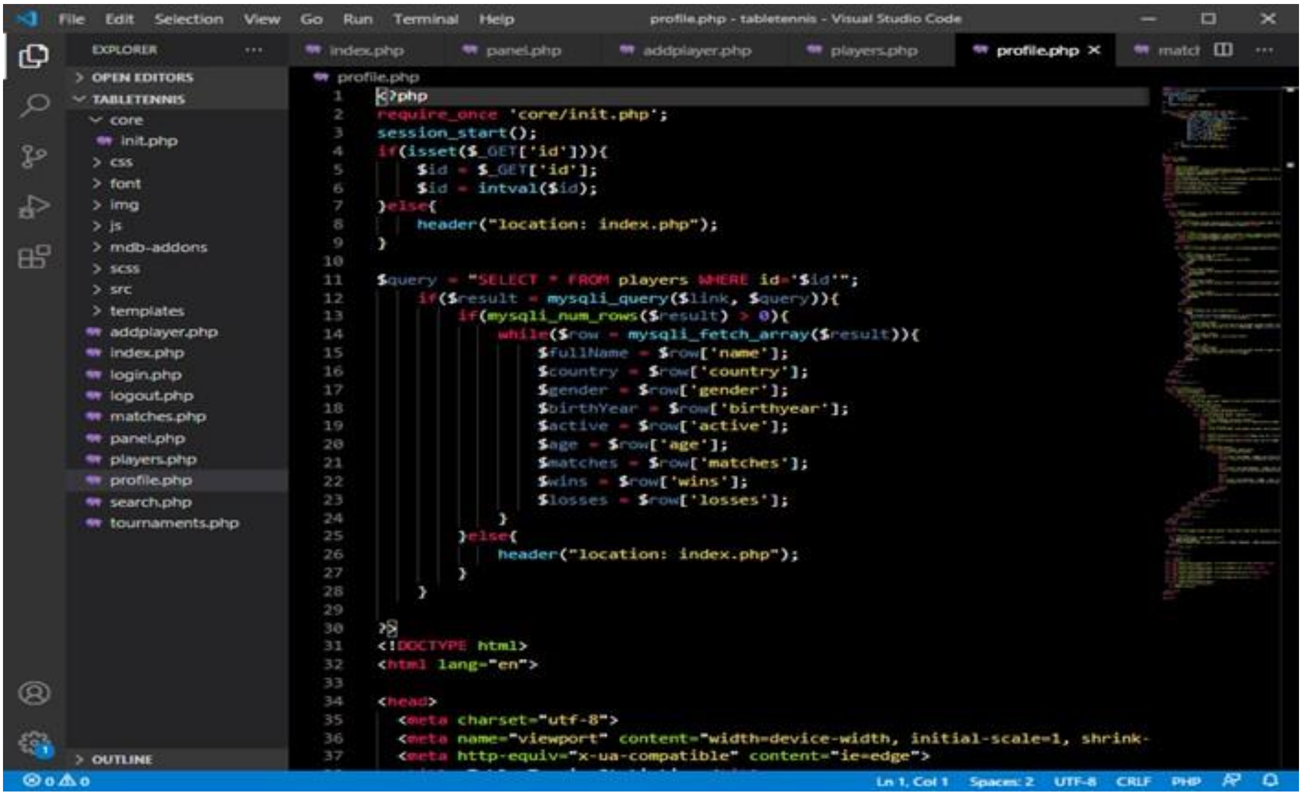
2.1. *Developer Tools*

For the ability to run the web application on a device (desktop/laptop), a software called "XAMPP" is utilized [4]. XAMPP is a free and open-source cross-platform web server solution stack package developed by Apache Friends, consisting mainly of the Apache HTTP Server,

An additional tool that will aid in the programming process is the installation of an integrated development environment software. For this project, we will be using Visual Studio Code since it supports code auto-completion, proper indentation, and software debugging.

MySQL database, and interpreters for scripts written in the PHP and Perl programming languages [2]. XAMPP contains necessary packages and dependencies that will make the system operational. It consists of Apache which will be used as a web server. A web server is a computer that stores web server software and a website's component files [5]. (for example, HTML documents, images, CSS stylesheets, and JavaScript files). A web server connects to the Internet and supports physical data interchange with other devices connected to the web. Fig. 2 shows a sample of XAMPP control panel that will allow the develop to activate the Apache server and MySQL database.

Visual Studio Code, like Visual Studio software, is essentially a text editor with pre-installed plugins to aid in application development [6]. Fig. 3 illustrates an example of how the IDE organizes code and indents them properly to aid developers.


File
Edit
Selection
View
Go
Run
Terminal
Help
profile.php - tabletennis - Visual Studio Code
EXPLORER
index.php
panel.php
addplayer.php
players.php
profile.php
matc
OPEN EDITORS
TABLETENNIS
core
init.php
css
font
img
js
mdb-addons
scss
src
templates
addplayer.php
index.php
login.php
logout.php
matches.php
panel.php
players.php
profile.php
search.php
tournaments.php
OUTLINE
Ln 1, Col 1
Spaces: 2
UTF-8
CRLF
PHP

```
<?php
require_once 'core/init.php';
session_start();
if(isset($_GET['id'])){
    $id = $_GET['id'];
    $id = intval($id);
}else{
    header("location: index.php");
}

$query = "SELECT * FROM players WHERE id='$id'";
    if($result = mysqli_query($link, $query)){
        if(mysqli_num_rows($result) > 0){
            while($row = mysqli_fetch_array($result)){
                $fullName = $row['name'];
                $country = $row['country'];
                $gender = $row['gender'];
                $birthYear = $row['birthyear'];
                $active = $row['active'];
                $age = $row['age'];
                $matches = $row['matches'];
                $wins = $row['wins'];
                $losses = $row['losses'];
            }
        }else{
            header("location: index.php");
        }
    }

?>
<!DOCTYPE html>
<html lang="en">

<head>
  <meta charset="utf-8">
  <meta name="viewport" content="width=device-width, initial-scale=1, shrink-
  <meta http-equiv="x-ua-compatible" content="ie=edge">
```

Fig. 3. Visual Studio Code Editor

2.2. *Selected Development Methodology*

The Agile Model of Systematic & Iterative Development was applied. The primary reason for selecting this approach is because of the complexity of the framework that will be developed. The process focuses on breadth rather than depth. We will not concentrate on the complexity of the feature because of the significance of providing few functional features that the user could utilize until the initial version of the framework has been created. This will ensure that the system operations are designed so that users can use them [6]. There are 7 phases in the Incremental & Iterative Development Model, which include Planning, System Analysis, Design, Implementation, Application Testing, Assessment and Delivery [4]. Fig. 4 illustrates the title of each stage in the development.

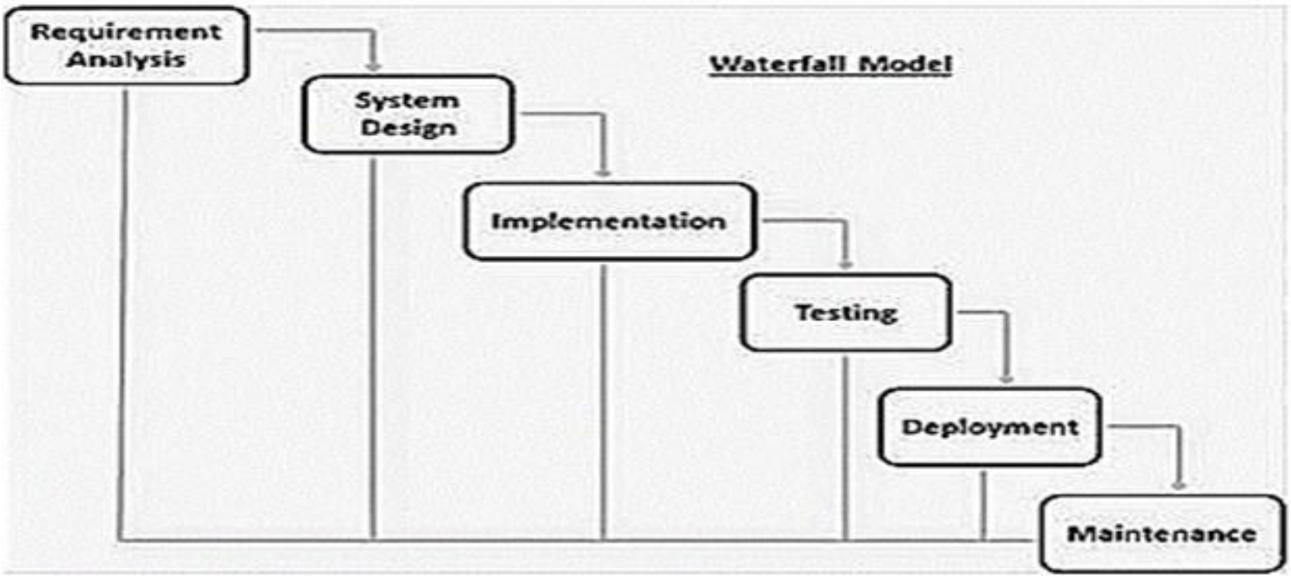


Fig. 4. Incremental & Iterative Development Stages

2.3. *Selected Development Methodology*

The general hardware requirements, microcapsules physical properties, and software specifications for both the user and the developer are tabulated in Table 1, Table 2, and Table 3 respectively.

Table 1. Basic hardware requirements

Component	Minimum Requirements
Processor	1.5GHz and above
Memory (RAM)	512MB
Hard-disk Space	5GB
Video Graphics Card	DirectX 9 and above
Screen Resolution	1024 x 768 or higher

Table 2. The process yield and physical properties of the microcapsules

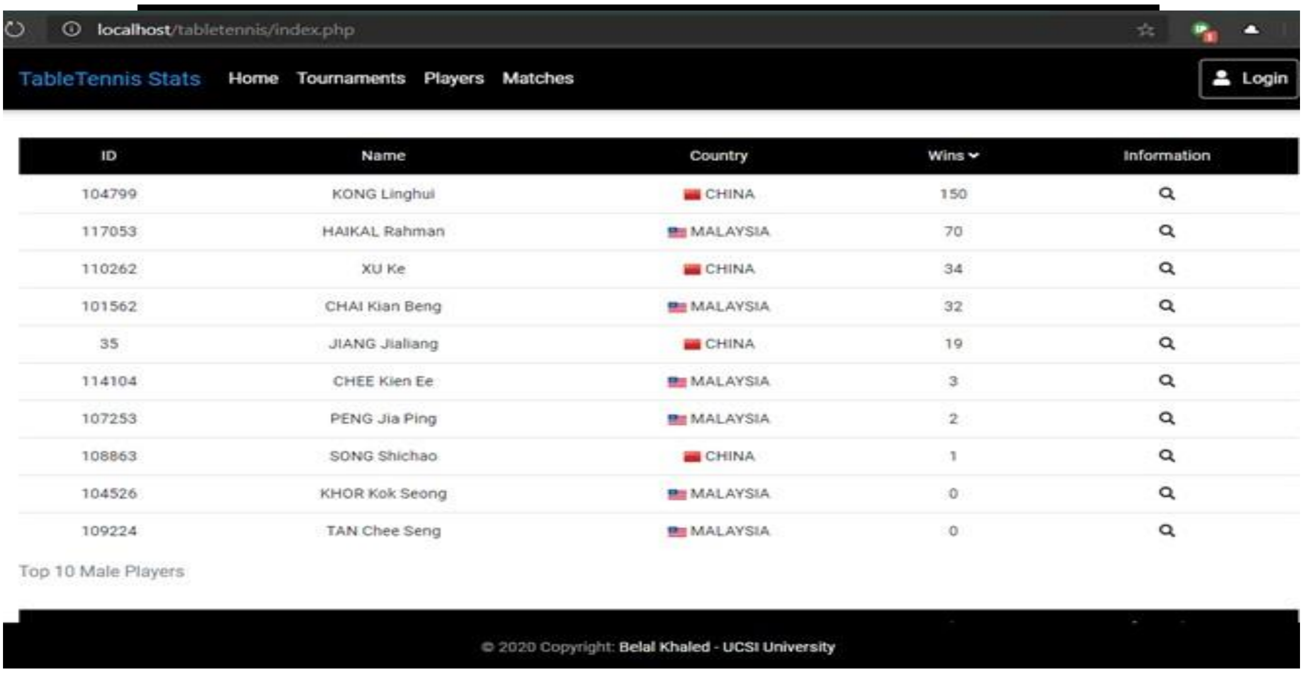
Component	Minimum Requirements
Operating System	Windows XP/7/8/10
Web Browser	Chrome/Firefox
Server Software	Apache 2.x or above
Database Software	MySQL 5.0 or above
Integrated Development Environment	Visual Studio Code
PHP Editor	Visual Studio Code
Web Server	XAMPP

Table 3. User software requirements

Component	Minimum Requirements
Operating System	Windows XP/7/8/10
Web Browser	Chrome/Firefox

3. Results and Discussion

The following demonstrations were results obtained from the web application when launched by the user. As per depicted in Fig. 5, the web application home page includes the Id, name, Country name, number of winning cases and the last but not the least, information search has been provided to cover the in depth information of the athlete.



ID	Name	Country	Wins	Information
104799	KONG Linghui	CHINA	150	🔍
117053	HAIKAL Rahman	MALAYSIA	70	🔍
110262	XU Ke	CHINA	34	🔍
101562	CHAI Kian Beng	MALAYSIA	32	🔍
35	JIANG Jialiang	CHINA	19	🔍
114104	CHEE Kien Ee	MALAYSIA	3	🔍
107253	PENG Jia Ping	MALAYSIA	2	🔍
108863	SONG Shichao	CHINA	1	🔍
104526	KHOR Kok Seong	MALAYSIA	0	🔍
109224	TAN Chee Seng	MALAYSIA	0	🔍

Top 10 Male Players

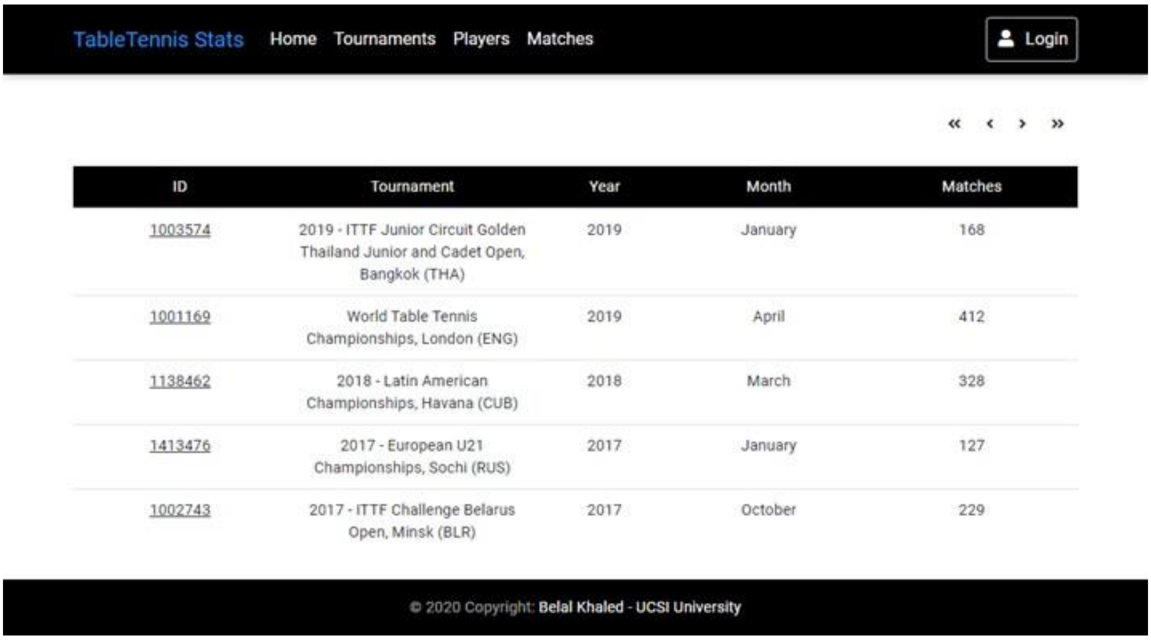
© 2020 Copyright: Belal Khaled - UCSI University

Fig. 5. Web Application Homepage

The homepage of the web application demonstrated in Fig. 5, displays basic statistics of the top 10 table tennis players of the year for both men and women separated into two tables. Each table will contain the unique identification number of the player for which can be used to retrieve their designated records from the database using PHP and MySQLv. Additionally, the name of the player described in the second column of the table. The table will also display the player's association as well as their total scores for the year. All data are sorted by the players' total scores in descending order. The table will also contain a final column which represents a magnifying glass that will redirect the user to a separate profile page which will describe the information of the player in detail when clicked.

An administrator can log into the system by using the login form as shown in Fig. 6. The form includes two input fields: username and password. Once the user enters their credentials, the web application will cross-check the details with the user records stored in the database. If the credentials are correct, the user will be redirected to the homepage where the navigation bar will display the full name of the user. If the input values are invalid, the application will display an error message to the user. Once logged in, the user can click on their full name located at the top right corner of the application and they will be redirected to the control panel. The control panel will allow the user to execute administrative operations such as adding, modifying, or deleting players.

This is applicable to table tennis matches and tournaments as well. The administrative operations are essentially handled at the back end where PHP will validate the user's input and make changes to the database accordingly.



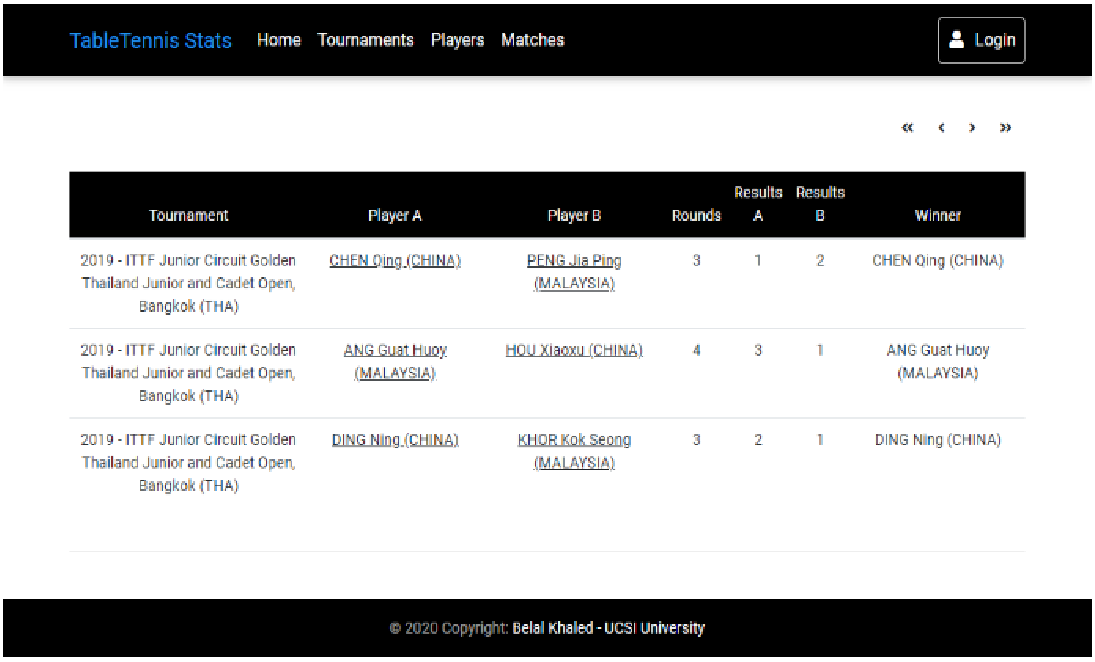
ID	Tournament	Year	Month	Matches
1003574	2019 - ITTF Junior Circuit Golden Thailand Junior and Cadet Open, Bangkok (THA)	2019	January	168
1001169	World Table Tennis Championships, London (ENG)	2019	April	412
1138462	2018 - Latin American Championships, Havana (CUB)	2018	March	328
1413476	2017 - European U21 Championships, Sochi (RUS)	2017	January	127
1002743	2017 - ITTF Challenge Belarus Open, Minsk (BLR)	2017	October	229

© 2020 Copyright: **Belal Khaled - UCSI University**

Fig. 6. Tournaments List Page

Fig. 6 is an illustration of the web application that will display a list of tournaments sorted in descending order by time (in years). For each row of data, the table contains a unique identification number for reference with the database followed by the name of the tournament being held. The user will be able to see the number of total matches that was held during the tournament which can be viewed in detail when the user clicks on the identification number. They will be redirected to a separate page that will list all matches and their details as shown in Fig. 8. When viewing the list of matches in a tabulated form, the user is expected to find the name of players who were involved including the number of points each player has allocated. The total rounds displayed is calculated based on the total number of points collected by the match's corresponding players.

There is always a need for the exact information of the competition lists to have a better understanding and more reliable analysis in each tournament and hence, in this web app, the matches list page has been introduced in which, the tournament, players name and the result of the game has been stated clearly. This webpage is shown in Fig. 7.

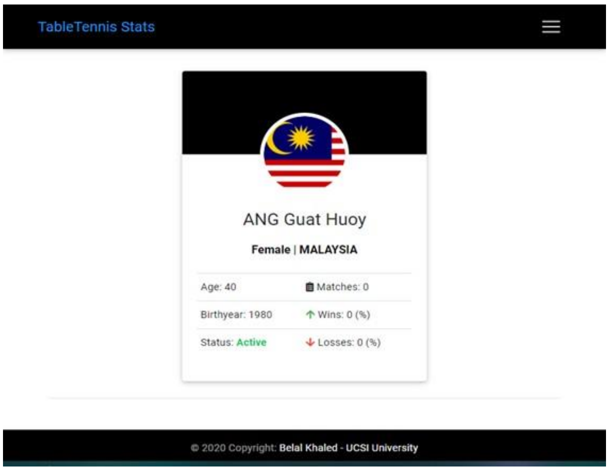


Tournament	Player A	Player B	Rounds	Results A	Results B	Winner
2019 - ITTF Junior Circuit Golden Thailand Junior and Cadet Open, Bangkok (THA)	CHEN Qing (CHINA)	PENG Jia Ping (MALAYSIA)	3	1	2	CHEN Qing (CHINA)
2019 - ITTF Junior Circuit Golden Thailand Junior and Cadet Open, Bangkok (THA)	ANG Guat Huoy (MALAYSIA)	HOU Xiaoxu (CHINA)	4	3	1	ANG Guat Huoy (MALAYSIA)
2019 - ITTF Junior Circuit Golden Thailand Junior and Cadet Open, Bangkok (THA)	DING Ning (CHINA)	KHOR Kok Seong (MALAYSIA)	3	2	1	DING Ning (CHINA)

© 2020 Copyright: Belal Khaled - UCSI University

Fig. 7. Matches List Page

To have a clear analysis on the athletes, the player profile page is also considered in this web page. The athlete’s name, gender, age, number of wins and loses followed by the status of the player as either active or retired is determined.



ANG Guat Huoy

Female | MALAYSIA

Age: 40
Matches: 0
Birthyear: 1980
Wins: 0 (%)
Status: Active
Losses: 0 (%)

© 2020 Copyright: Belal Khaled - UCSI University

Fig. 8. Player Profile Page

4. Conclusion

This paper summarizes the development of a web application for managing and viewing table tennis tournament scores and statistics. It is a demonstration of applying efficient methodologies to implement the system based on the user requirements as well setting up the developer’s workspace for application development. Minimum specifications were stipulated to ensure that the system is optimal for development, implementation, and execution to the developer as well as the user. The objective of this paper was to map out the guidance for setting up a web server and install the necessary components to successfully run the software.

The application is usable in any device across several platforms and to promote responsiveness and flexibility, the application can also run on mobile devices with the ability to work in almost any web browser.

By running the web application, the user can view valuable statistical insights about table tennis tournament scores as well as view additional information about the players involved in matches that were held in those tournaments. Additionally, this paper also describes the other features that was built in the web application including the dedicated user control panel for executing administrative operations with the records stored in the database.

References

1. Abdullah, R.A., Mokayef, M., Mokayef, M., Tiang, S.S., & Lim, W.H. (2020). Table Tennis Using Arduino For Seniors' Healthcare.
2. Amann, Sven & Proksch, Sebastian & Nadi, Sarah & Mezini, Mira. (2016). A Study of Visual Studio Usage in Practice. 124-134. 10.1109/SANER.2016.39
3. Stobart S., Vassileiou M. (2004) MySQL Database and PHPMyAdmin Installation. In: PHP and MySQL Manual. Springer Professional Computing. Springer, London. https://doi.org/10.1007/978-0-85729-404-3_34
4. Cvitić, I., Peraković, D., Periša, M., & Sekondo, M. (2022). Exploring the Applicability of Open-Source Tools for Web Application Cybersecurity Improvement. In International Conference on Future Access Enablers of Ubiquitous and Intelligent Infrastructures (pp. 64-79). Springer, Cham.
5. Chan R, Lok K, Woo J. Prostate cancer and vegetable consumption. Mol Nutr Food Res. 2009;53(2): 201-216.
6. Weintraub, Eli. (2010). Weintraub, E. "A New System Development Life Cycle Model: Vertical, Integrative and Dynamic", IRM - International Conference on Information Resources Management, USA, 5, 2010.

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.

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. 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.

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.

Small Target Detection Based on YOLOX

Keying Ren
School of Electronic Information and Automation, Tianjin University of Science and Technology,
1038 Dagu Nanlu, Hexi District, Tianjin, China

Xiaoyan Chen
School of Electronic Information and Automation, Tianjin University of Science and Technology,
1038 Dagu Nanlu, Hexi District, Tianjin, China

Zhihui Chen
School of Electronic Information and Automation, Tianjin University of Science and Technology,
1038 Dagu Nanlu, Hexi District, Tianjin, China

E-mail: renkeying@mail.tust.edu.cn, cxywxr@tust.edu.cn, 1594838831@qq.com
www.tust.edu.cn

Abstract

With the development of drone, small target detection has become a hotspot of current research. In this paper, the network of small target detection is based on YOLOX is studied. There are a lot of small targets in the images taken by UAV, which brings great difficulty to the detection task. The main improvement of the proposed network is that, based on the original YOLOX network, four-scale adaptive spatial feature fusion pyramid is added to filter the conflicting information between different scales and improve the expressive accuracy of the small target features. Experiments show that the proposed method has average accuracy of 32.83% in the self-built dataset, which is 2.53% higher than that of YOLOX-s, 1.53% higher than that of Yolov6-s, and 5.73% higher than that of YOLOv5-s.

Keywords: Drone shooting, Small target detection, YOLOX, Adaptively Spatial Feature Fusion

1. Introduction

With the development of artificial intelligence technology, the field of computer vision has made a great breakthrough. Object detection is one of the main tasks of computer vision, which has been applied to pedestrian detection, face detection and other tasks. Object detection algorithms include SSD (Single Shot MultiBox Detector)[1], CornerNet[2], YOLO (You Only Look Once)[3]series, etc. These algorithms can directly classify and locate objects. The one-stage algorithm has faster detection speed than the two-stage algorithm, but the error rate and missing rate are relatively high. In MS COCO[4], the accuracy of small-target detection is less than half in the large/medium target detection. Therefore, it is an urgent problem to improve the accuracy of small targets in UAV aerial images.

In 2020, Nayan et al. proposed a small target detection algorithm based on YOLOv3, which used upsampling and residual joint to extract multi-scale features of different convolution levels in learning tasks[5]. This method significantly improves the ability of small target detection. In the same year, an Enhanced Context Model (ECM) was proposed, which used a double-dilation convolution structure to reduce the number of parameters, expand the effective receptive field, strengthen the context information, and apply to the prediction layer of the network. However, this method depends on the

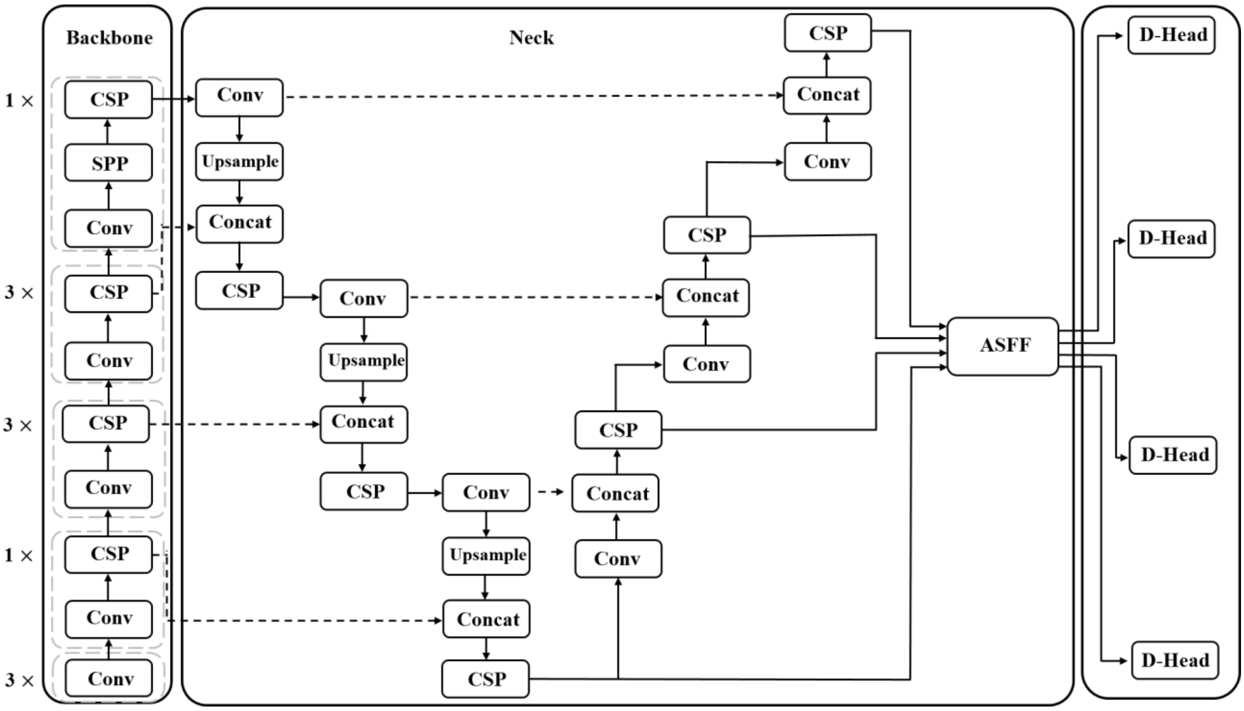


Fig.1 Structure comparison of YOLOX

design of the context window or the size of the receptive field, which may result in the loss of important context information. Wang Jianjun et al. studied an improved YOLOv3 small target detection algorithm, It increases the depth of shallow convolution in backbone networks to enhance the feature extraction capability of backbone. The RFB (Receptive Field Block) structure is used to enlarge the receptive field of the shallow feature map. Megvii Technology's YOLOX algorithm uses anchor free method and does not need to set anchors in target detection tasks. It combines data enhancement, simOTA label allocation strategy and so on. YOLOX has become a mainstay in recent years.

In order to improve the accuracy of small target detection, this paper proposes a small target detection algorithm based on YOLOX, which has the following characteristics:

1.Add a small target detection layer based on the original YOLOX structure.

2.A four-scale adaptive spatial feature fusion pyramid is used to filter the conflicting information between different scales.

2. Method

The overall structure of this method is shown in Figure 1, which mainly includes backbone, Neck and decoupling detection head.

2.1. *backbone*

The backbone is used to take the input image and extract useful features from it. This is a very important step in the object detector because it is the main structure for extracting context information. Most of the backbone of traditional YOLO algorithm adopts residual connection [6]. Identity mapping is used to solve the problem of network training difficulty

Our backbone uses the structure of CSPNet. its small model uses a basic building block ratio of 1,3,3,1. Four stages were used in the backbone to obtain different scale features. At the end, the SPP acquires features at different scales and merges them.

2.2. *Neck*

In convolutional networks, deep networks are sensitive to semantic features, while shallow networks are sensitive image features. This feature of convolutional networks often brings a lot of trouble in object detection.

The high-level network can respond to semantic features. However, because the size of feature map is too small to get detailed geometric information, which is not conducive to target detection.

To solve this problem, YOLOX introduced PANet[7], which uses top-down feature fusion followed by bottom-up feature fusion. Through this operation, we can achieve better detection results.

2.3. *Decoupled head*

The existing target detector still uses coupling head to realize object detection. YOLOX introduced the decoupling detection head, whose structure is shown in Figure 2. In this structure, 1×1 convolution was used to compress the feature channels to 128, and two 3×3 convolution were used respectively to extract the features required for classification and regression tasks.

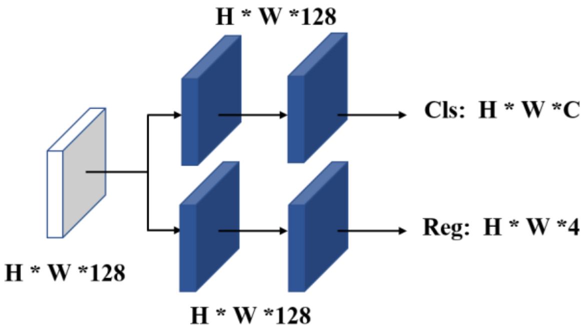


Fig.2 decoupling Head

2.4. *Adaptively spatial feature fusion*

In object detection, feature pyramid is used for feature fusion of different scales. However, this operation results in inconsistent information between different scale features. This phenomenon increases the difficulty of small target detection. Therefore, adaptive spatial feature fusion pyramid (ASFF) is adopted in this paper to filter the spatial conflict information between different scale features, so as to improve the inconsistency between different scale features.

3. Experiments

3.1. *Dataset*

The images used for the experiment are taken by drones at different locations and altitudes. There are 9,563 images, 11 categories, and more than 560,000 pre-labeled anchors.

3.2. *Experiments and Results*

Stochastic gradient descent (SGD) was used in training, and the parameters were set as follows: momentum=0.9, weight decay=5e-4. Cosine annealing learning rate was used, initial learning rate was set to zero, warm-up training for 5 epoch, the basic learning rate was 0.01, and the learning rate was adjusted by linear scaling. In the training process, the exponential moving average (EMA) strategy was used, and its attenuation was 0.9998.

Common Objects in Context (COCO) evaluation criteria are most commonly used in target detection, which can best reflect the performance of a target detection algorithm. COCO evaluation criteria are divided into average precision (AP) and average recall (AR). AP and AR are subdivided into indexes under different parameters. Among them, AP is divided into mean average precision (mAP), AP50, AP75. mAP is the most commonly used index among these evaluation indexes and can best reflect the performance of the detector.

In order to compare the effectiveness of the proposed algorithm, a comparative experiment is conducted in this paper. COCO evaluation index is used to evaluate. The results are shown in Table 1.

Table 1 Detection results of different algorithm

Method	**mAP（%）**	**AP50（%）**	**AP75（%）**
YOLOv5-s	27.1	47.6	28.4
YOLOX-s	30.3	59.5	32.5
YOLOV6-s	31.3	60.7	33.6
our	**32.83**	**62.36**	**34.54**

Experiment results reported in Table 1 is the average of multiple experiments. Experiments show that the proposed method has a mAP average accuracy of 32.83% in the self-built dataset, which is 2.53% higher than that of YOLOX-s, 1.53% higher than that of Yolov6-s, and 5.73% higher than that of YOLOv5-s.

4. Discussion

In order to effectively detect small targets in aerial photography, this paper proposes a small target detection algorithm based on YOLOX. A four-scale adaptive feature fusion module (ASFF) was used to filter the conflicting information between different scale features, so as to improve the efficiency of the detector. Although the method proposed in this paper has improved the performance index, there is still room for improvement in practical application. For example, optimize the detection effect on complex scene data set.

Acknowledgment

This work was supported by "Tianjin University of Science and Technology – softsz Intelligent edge computing Joint Laboratory"

References

1. Liu W , Anguelov D , Erhan D , et al. SSD: Single Shot MultiBox Detector, *European Conference on Computer Vision. Springer, Cham*, 2016: pp.21-37.
2. Law H, Deng J. Cornernet: Detecting objects as paired keypoint, *Proceedings of the European conference on computer vision (ECCV*), 2018: pp 734-750.
3. Redmon J , Divvala S , Girshick R , et al. You Only Look Once: Unified, Real-Time Object Detection, *IEEE conference on computer vision and pattern recognition*, 2016: pp. 779-788.
4. Lin T Y, Maire M, Belongie S, et al. Microsoft coco: Common objects in context, *European conference on computer vision*, 2014: pp 740-755.
5. Nayan A A, Saha J, Mozumder A N, et al. Real time detection of small objects[J]. International Journal of Innovative Technology and Exploring Engineering, 2020, 29(5): pp 14070-14083.
6. He K, Zhang X, Ren S, et al. Deep residual learning for image recognition, *Proceedings of the IEEE conference on computer vision and pattern recognition.* 2016: pp 770-778.
7. Liu S, Qi L, Qin H, et al. Path aggregation network for instance segmentation, *Proceedings of the IEEE conference on computer vision and pattern recognition.* 2018: pp 8759-8768.

Authors Introduction

Mr. Keying Ren

He is an Master of Control Science and Engineering, Tianjin University of Science and Technolog. The research topic is target detection and tracking based on deep learning.

Prof Xiaoyan Chen

She has been working on the application of deep learning network models to the field of computer vision for many years. She is currently a Professor with the Tianjin University of Science and Technology.

Mr. Zhihui Chen

He received his bachelor's degree from the school of electronic information and automation of Tianjin University of science and technology in 2021. He is acquiring for his master's degree at Tianjin University of science and technology.

Super Resolution Reconstruction Model Based on Attention Mechanism and Generative Adversarial Network

Xia Miao
School of Electronic Information and Automation, Tianjin University of Science and Technology,
1038 Dagu South Road, Tianjin, China

Kunzhi Yang
School of Electronic Information and Automation, Tianjin University of Science and Technology,
1038 Dagu South Road, Tianjin, China

Xiaoyan Chen
School of Electronic Information and Automation, Tianjin University of Science and Technology,
1038 Dagu South Road, Tianjin, China

E-mail: miaoxia@mail.tust.edu.cn, cxywxr@tust.edu.cn
www.tust.edu.cn

Abstract

The purpose of image super resolution reconstruction is to recover high frequency details from low resolution images containing little information, so as to improve the visual effect of images. Based on the traditional SRGAN algorithm, this paper by importing the residual network structure and combining the channel attention mechanism constructs the basic network module, and this method is named CA-SRGAN. The proposed CA-SRGAN replace the base block of SRGAN with a residual dense block based on the channel attention mechanism. Test in Set5, Set14 and BSD100 data sets, the results of peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) of CA-SRGAN algorithm are better than those of the traditional SRGAN algorithm, and the details of the reconstructed image are clearer. The whole show better robustness and comprehensive performance.

Keywords: Super resolution reconstruction, Attention mechanism, Generative adversarial network

1. Introduction

With the rapid development of computer technology and the popularization of intelligent equipment, people's requirements for image clarity and resolution are constantly improving. However, in the real world scene, the image information obtained by the acquisition equipment often contains noises results in blur, phantom, sharpening. Sometimes this kind of image information cannot be directly applied to actual work, and cannot better complete the work of target recognition and detection, which increases the difficulty and cost of the work.

With the penetration of smart phones and mobile Internet in daily life, the cost of small high-resolution screens is decreasing year by year, and the computing power of miniaturized devices is also increasing rapidly. Benefiting from such external environment and the rapid development of deep learning, image super resolution reconstruction technology has been applied in many fields[1], such as security monitoring[2], satellite imaging, medical imaging, and mobile terminal image transmission. Image Super Resolution (SR) technology can recover detailed information. The improved reconstructed image by SR can be convenient for target recognition, target tracking and other work. Due to the fitting ability of superior neural network, image reconstruction method based on deep learning can extract the feature information of image, and generate images with richer texture details compared with traditional image reconstruction methods. The main purpose of this paper is to apply deep learning theory and method to, so as to obtain higher quality images.

2. Image super resolution reconstruction and generation of adversarial networks

2.1. *Image super resolution reconstruction*

Digital images stored in a computer usually contain attributes such as brightness, color space, resolution, etc. Image resolution reflects how many pixels an image has within a certain range, which determines how fine the detail of the image is. Generally, the higher the resolution, the more pixels are included in the image and the better the visual effect the image will present.

The optical signal in nature is continuous before entering the imaging system. In the process of being captured and entering the imaging equipment, it may go through various kinds of noise superposition, and finally get a digital image after discrete sampling. In this process, there may be some problems, such as imperfect transmission medium, noise interference, the object or the shooting equipment cannot achieve absolute stillness, and the degradation of the image is bound to be caused by the equipment performance, storage space limitation and other reasons, so that the output of the image is relatively low resolution. The process of image super resolution is exactly the opposite of the above imaging process, aiming to recover a higher resolution image with more detailed information from the degraded low resolution image.

2.2. *The adversarial generation network*

Inspired by the two-person zero-sum game in game theory, adversarial generative network was first proposed by Goodfellow et al. in 2014[3]. With the rapid development of adversarial generative network in theory and model, various improved variants keep emerging, which are not limited to the original research field of computer vision, but have extended and applied in natural language processing, human-computer interaction and other fields.

Adversary to generate network contains two models, which are generator and discriminator. The generator is used to capture a distribution network and generate a noisy image. A discriminator is a network that determines whether a picture is generated image or a real image. For the generator, the generated image tries to deceive the discriminator so that the discriminator cannot tell whether the source of the input image is a generator or real data[4]. For the discriminator, it can make use of prior knowledge in the training process to continuously acquire more powerful discriminant ability. The two models form competition and confrontation, promote each other, and finally reach a state of equilibrium maintained at a higher level.

Ledig et al.[5] first used the idea of adversarial generative network to solve the problem of image super resolution reconstruction and proposed the SRGAN model. Because of the particularity of the image hyper division problem, the input of the generator in SRGAN is the low-resolution image and the output is the super-resolution reconstructed image[6].

3. Super resolution reconstruction algorithm combined with channel attention mechanism

Based on SRGAN, we proposed an improved network named CA-SRGAN. In this paper, the residual network structure is introduced first, and the basic network module is constructed by combining the channel attention mechanism. Secondly, three key components of SRGAN, network architecture, loss resistance and loss perception, are studied and improved. This article then describes the experimental study environment, and the analyzes of the experimental results is introduced.

According to the different functions of the channels in the training process, the channel attention mechanism uses an adaptive way to assign different weights to each channel. The attention mechanism introduced into the generative network can make the network focus on facial features during the training process. In general, if an image has C channels, low resolution image (LR) can be represented as the real valued tensor of size $C \times W \times H$, high resolution image (HR) and super resolution image (SR) can be represented as the real valued tensor of $rH \times rW \times C$. Where H is the height of the image, W is the width of the image, is the number of channels in the image, C is the number of image channels and r is the down-sampling factor.

First, the backbone of the generated network is improved using residual channel attention blocks (RCAB) based on the channel attention mechanism to replace the base blocks[7] of SRGAN. RCAB is composed of residual channel attention blocks (RB) and channel attention mechanisms (CA). Each of these residual in residual (RIR) groups contains g residual blocks (RG) and long skip connections (LSC) each residual block contains b RCABs.

In the field of super resolution technology and image deblurring technology, it has been proved that the network performance and generalization ability will be enhanced when BN layer is removed, and the computational complexity of the network model will also be reduced. The generated network in this paper also enhances the generalization ability of the network model by removing the BN layer. By considering the dependency between channels, an attention mechanism for adaptive re-estimation of channel characteristics is introduced. The reconstructed face image has a more natural overall contour and local regional features.

The experimental results show that using this residual structure can train very deep CNN without the phenomenon of gradient disappearing, and the high-frequency information extracted by feature will not missing, which can better reconstruct the low-resolution image and improve the performance of the whole network model. The long jump connection can bypass the rich low frequency information and process the high frequency information, so as to improve the reconstruction effect. The overall network model structure is shown in Figure 1.

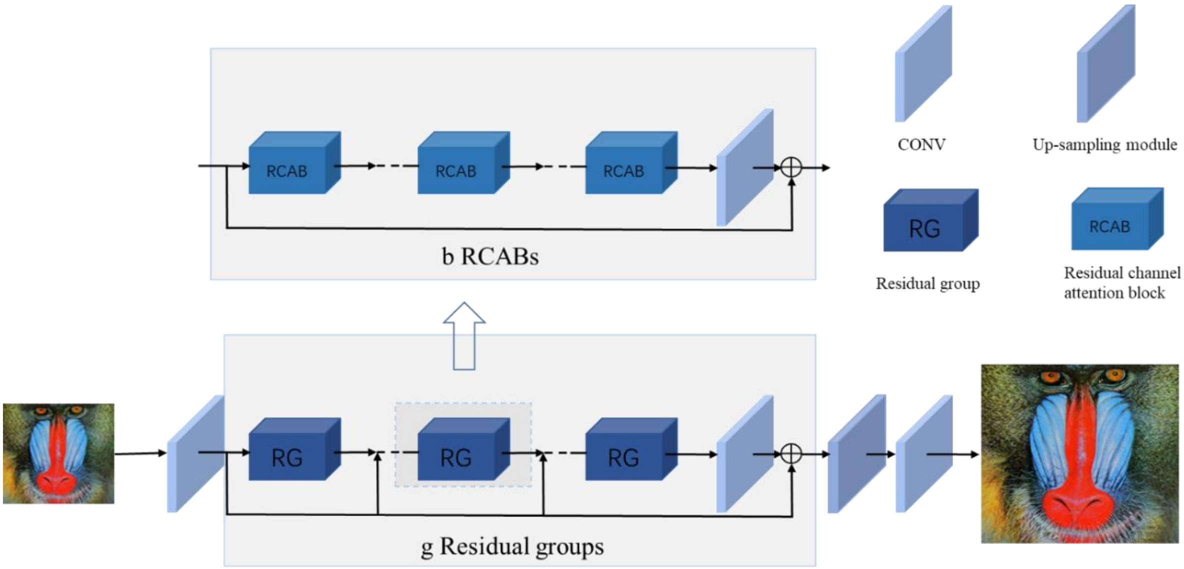

RCAB
RCAB
RCAB
b RCABs
CONV
Up-sampling module
RG
Residual group
RCAB
Residual channel attention block
RG
RG
RG
g Residual groups

Fig. 1 Generation network improvements

The formula for Group i RGS is shown in Eq. (1)

$$F_i = H_i(F_{i-1}) = H_i(H_{i-1}(\dots H_1(F_0)\dots))$$

(1)

The RCAB block for group i can be expressed as Eq. (2)

$$F_{i,b} = H_{i,b}(H_{i,b-1}(\dots H_{i,1}(F_{i-1})\dots)) \qquad (2)$$

Where $F_{i,b}$ and $F_{i,b-1}$ are the input and output of the b RCAB in group i, respectively. The structure of RCAB is expressed as Figure 2,

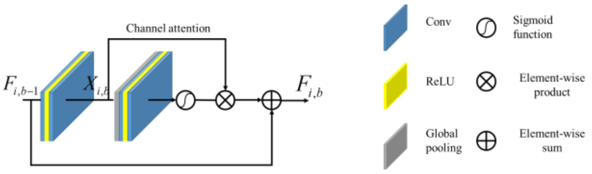

Channel attention
$F_{i,b-1}$
$X_{i,b}$
$F_{i,b}$
Conv
Sigmoid function
ReLU
Element-wise product
Global pooling
Element-wise sum

Fig. 2 The structure of RCAB

4. Experiment and Analysis

4.1. *Hardware Conditions and experiment set*

The training environment is Ubuntu 18.04 operating system, CPU is Intel Xeon E5-2360, graphics card is two NVIDIA GE-FORCE GTX1080Ti 11 GB, development language is Python, Pytorch framework. DIV2k data set was used for the training set of the network model, and SET5, SET14 and BSD100 were used for the test set.

4.2. *Experimental evaluation*

In order to further verify the detection performance of proposed CA-SRGAN, SRCNN, VDSR and SRGAN were used in this paper to compare the detection performance with CA-SRGAN proposed in this paper. The test data sets were Set5, Set14 and BSD100. The results are shown in the Table 1.

Table 1 Experimental results

	Set5		Set14		BSD100	
	PNSR	SSIM	PNSR	SSIM	PNSR	SSIM
SRCNN[8]	24.987	0.650	23.243	0.529	23.231	0.559
VDSR[9]	26.054	0.698	24.984	0.636	24.235	0.613
SRGAN[5]	27.971	0.740	25.318	0.682	24.237	0.631
CA-SRGAN	28.662	0.860	26.356	0.744	25.795	0.698

It can be seen from the table that compared with the original SRGAN model, the PSNR and SSIM of the improved CA-SRGAN model on SET5 data set are increased by 0.691 and 0.12 respectively. In SET14 data set, PSNR increased by 1.038, SSIM increased by 0.062; On BSD100 data set, PSNR increased by 1.558 and SSIM increased by 0.067. Compared with other models, the index also has a certain degree of improvement. To sum up, CA-SRGAN algorithm is superior to SRGAN algorithm.

5. Conclusion

In this chapter, an image super resolution reconstruction method based on attention mechanism and generative adversarial network is studied. The CA-SRGAN algorithm is compared with several classical super resolution reconstruction algorithms. The experimental results show that the CA-SRGAN algorithm proposed in this chapter has obvious improvement in the objective evaluation index and subjective visual effect, which shows the effectiveness and superiority of the proposed algorithm.

Acknowledgment

This work was supported by Shenzhen Softsz Co. Ltd., China.

References

1. Nasrollahi, K. , & Moeslund, T. B. . (2014). Super-resolution: a comprehensive survey. Machine Vision & Applications, 25(6), 1423-1468. https://sci-hub.st/10.1007/s00138-014-0623-4
2. Xiang, X. , Liu, W. , & Ling, L. . (2014). Low resolution face recognition in surveillance systems. Journal of Computer & Communications, 02(2), 70-77. https://www.scirp.org/journal/PaperInformation
3. Goodfellow, I. J., “Generative Adversarial Networks”, 2014. https://arxiv.org/abs/1406.2661
4. Mao, X. , Li, Q. , Xie, H. , Lau, R. , & Smolley, S. P. . (2017). Least squares generative adversarial networks. IEEE. https://sci-hub.st/10.1109/ICCV.2017.304
5. Ledig, C. , Theis, L. , Huszar, F. , Caballero, J. , Cunningham, A. , & Acosta, A. , et al. (2016). Photo-realistic single image super-resolution using a generative adversarial network. IEEE Computer Society. https://arxiv.org/abs/1609.04802
6. Ma, C. , Jiang, Z. , Rao, Y. , Lu, J. , & Zhou, J. . (2020). Deep face super-resolution with iterative collaboration between attentive recovery and landmark estimation. IEEE. https://arxiv.org/abs/2003.13063v1
7. Zhang, Y., Li, K., Li, K., Wang, L., Zhong, B., and Fu, Y., “Image Super-Resolution Using Very Deep Residual Channel Attention Networks”, 2018. https://arxiv.org/abs/1807.02758
8. C. Dong, C. C. Loy, K. He and X. Tang, "Image Super-Resolution Using Deep Convolutional Networks," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 38, no. 2, pp. 295-307, 1 Feb. 2016, doi: 10.1109/TPAMI.2015.2439281. https://arxiv.org/abs/1501.00092
9. J. Kim, J. K. Lee and K. M. Lee, "Accurate Image Super-Resolution Usinrence on Computer Vision and Pattern Recognition (CVg Very Deep Convolutional Networks," 2016 IEEE ConfePR), 2016, pp. 1646-1654, doi: 10.1109/CVPR.2016.182. https://ieeexplore.ieee.org/document/7780551

Authors Introduction

Ms. Xia Miao

She is currently studying in Tianjin University of Science and Technology, majoring in electronic information, and main research field is computer vision.

Mr. Kunzhi Yang

He received his bachelor's degree in Engineering from Lanzhou University of Technology in 2019. He is currently studying control science and engineering at Tianjin University of Science and Technology.

Prof. Xiaoyan Chen

She, professor of Tianjin University of Science and Technology, graduated from Tianjin University with PH.D (2009), worked as a Post-doctor at Tianjin University (2009.5-2015.5). She had been in RPI, USA with Dr. Johnathon from Sep.2009 to Feb.2010 and in Kent, UK with Yong Yan from Sep-Dec.2012. She has researched electrical impedance tomography technology in monitoring lung ventilation for many years.

An improved network for pedestrian-vehicle detection based on YOLOv7

Zhihui Chen
School of Electronic Information and Automation, Tianjin University of Science and Technology,
1038 Dagu Nanlu, Hexi District, Tianjin, China

Xiaoyan Chen
School of Electronic Information and Automation, Tianjin University of Science and Technology,
1038 Dagu Nanlu, Hexi District, Tianjin, China

Keying Ren
School of Electronic Information and Automation, Tianjin University of Science and Technology,
1038 Dagu Nanlu, Hexi District, Tianjin, China
E-mail:1594838831@qq.com, cxywxr@tust.edu.cn, renkeying@mail.tust.edu.cn
https://www.tust.edu.cn/

Abstract

With the continuous development and improvement of information technology, target detection has gradually attracted people's attention. Therefore, object detection has become more important. In this paper, a large number of experiments have been made to improve the detection accuracy of pedestrians, vehicles and license plates in cities. The improved algorithm based on YOLOv7 was used to conduct a large number of experiments on the urban pedestrian vehicle dataset. Experiments show that new improvements have improved detection accuracy.

Keywords: YOLOv7, target detection, Swin Transformer, CNeB

1. Introduction

With the sequential development of computer technology, the task of target detection in the transportation system is gradually transformed from manual to machine. Object detection is a basic research direction in the field of computer vision technology, depth learning and image recognition. It is the technical premise of more complex high-level computer vision tasks, such as predicting the behavior of objects in images after the detection task. Traditional detection methods mainly use HOG[1] or vector machine algorithm. They mainly use sliding Windows to detect targets. This method has the disadvantages of long time redundancy and poor robustness, so it is not suitable for large-scale applications. In recent years, target detection algorithm based on deep learning is gradually used in all walks of life. The algorithm based on deep learning has the characteristics of strong real-time and high accuracy in detection. Therefore, target detection algorithm based on depth learning has become the mainstream algorithm in target detection. This paper uses the depth learning algorithm to detect vehicles and pedestrians.

2. Detection algorithm

The main goal of pedestrian and vehicle detection algorithm is to find the target from a series of images or videos, and determine the size and location of the target according to the feature information. But in the actual scene of the city, there will be a lot of interferences. For example, occlusion, illumination and other conditions lead to false detection and missing detection, resulting in inaccurate detection. At present, the detection algorithms based on depth learning are mainly divided into two categories, namely, the two-stage target detection model based on region extraction and the one-stage target detection model for direct position regression.

2.1 *Two-stage object detection algorithm*

The core of the two-stage target detection model based on region extraction is the CNN, which has the advantages of local connection and weight sharing, and has good robustness in object classification applications. Firstly, the region extraction operation uses the CNN backbone to extract image features, then finds out possible foreground objects (candidate regions) from the feature map, and finally performs sliding window operation on the candidate regions to further judge the target category and location information, which greatly reduces the time complexity of calculation. The classical algorithms include R-CNN, SPP-NET, Fast R-CNN, Faster R-CNN and Mask R-CNN. R-CNN (Region CNN) was proposed by Ross-Girshick and was the earliest algorithm based on depth learning target detection. The mAP (mean average precision) of the algorithm on the VOC2007 dataset reaches 66%, breaking through the bottleneck that R-CNN has been unable to overcome for many years, and significantly improving the detection rate. But it has some disadvantages. The intermediate data generated by this algorithm will consume too much storage resources, and the input image must be forcibly scaled to a fixed size, which will lead to deformation of the target object and affect the accuracy of detection. SPP-NET solves the problems of repeated calculation and fixed size based on R-CNN.

Fast R-CNN uses the feature pyramid in SPP-NET for reference to improve the R-CNN network, greatly reducing the training time and improving the model detection performance. However, Fast-CNN generation of candidate boxes is time-consuming, and can not be used for end-to-end training and GPU acceleration. To solve this problem, Ren et al. designed RPN (Region Proposal Network,) in 2015 and proposed Faster R-CNN[2] model. Faster R-CNN achieves end-to-end training, but because of its structure, it has the disadvantages of large parameters, slow detection speed and poor real-time performance. The above target detection methods based on region extraction generate the frame of the region of interest, and then classifies and regresses the frame. Although the detection accuracy has been continuously improved, the detection speed is generally slow, which is not suitable for application scenarios with high real-time requirements.

2.2 *One stage target detection*

Different from the two-stage target detection algorithm, the one-stage detection algorithm pays more attention to the improvement of detection speed. Among the single-stage detection algorithms, SSD (Single Shot Multi Box Detector)[3] and YOLO (You Only Look Once) series are the main methods.

The one-stage target detection algorithm does not use the middle layer to extract candidate regions, but performs feature extraction, target classification and position regression in the entire convolution network, and then obtains the target position and category through a backward calculation. The accuracy is slightly lower than that of two-stage detection, but its speed is greatly improved, so that the detection algorithm based on One stage target detection can be used in many scenes that require reasoning speed.

This experiment uses YOLOv7 algorithm. YOLOv7 is faster and more accurate than most known target detectors. Among the known real-time target detectors above 30 frames per second of GPU V100, YOLOv7 has the highest detection accuracy. According to the different code running environments (edge GPU, common GPU and cloud GPU), YOLOv7 has designed three models respectively: YOLOv7 tiny, YOLOv7 and YOLOv7-W6. Compared with other YOLO series network models, the detection idea of YOLOv7 is similar to that of YOLOv4 and YOLOv5[4], and its network structure is shown in the Figure 1 below.

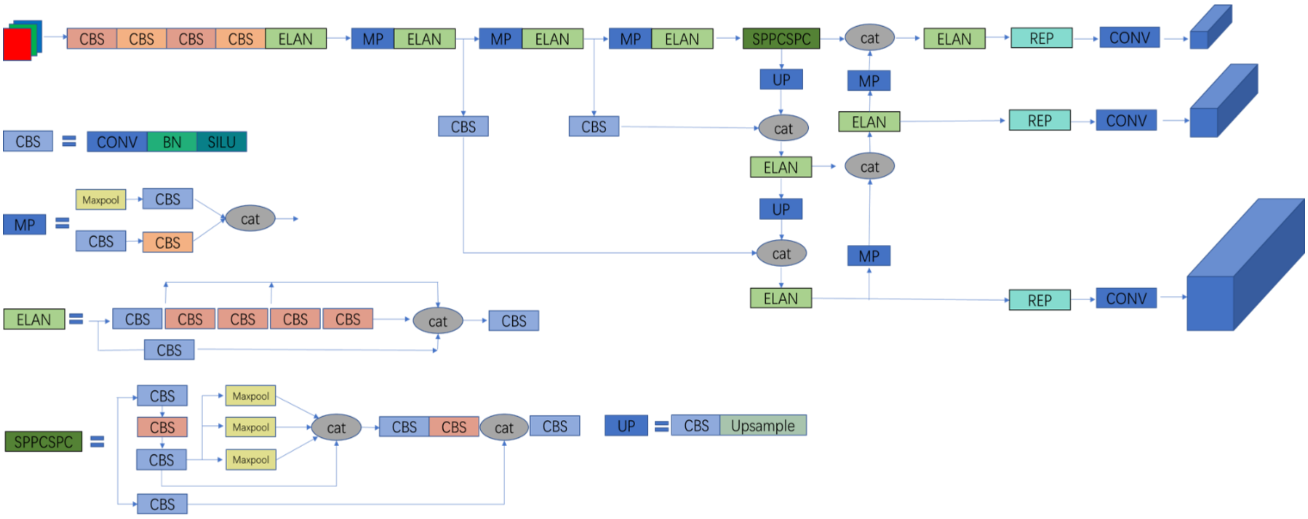

CBS
CBS
CBS
CBS
ELAN
MP
ELAN
MP
ELAN
MP
ELAN
SPPCSPC
cat
ELAN
REP
CONV
UP
MP
CBS
CBS
cat
ELAN
REP
CONV
ELAN
cat
UP
cat
MP
ELAN
REP
CONV
CBS
=
CONV
BN
SILU
Maxpool
CBS
cat
MP
=
CBS
CBS
ELAN
=
CBS
CBS
CBS
CBS
CBS
cat
CBS
CBS
CBS
Maxpool
CBS
Maxpool
cat
CBS
CBS
cat
CBS
SPPCSPC
=
CBS
Maxpool
CBS
UP
=
CBS
Upsample

Fig. 1 YOLOv7 network structure diagram

2.3 *YOLOv7*

YOLOv7[5] network is divided into input, backbone and head. The input picture is preprocessed by data enhancement. The preprocessed images enter the backbone for feature extraction. Then, the extracted features are fused by the head to obtain features of large, medium and small sizes. Finally, the fused features are sent to the detection head and the results are output. The backbone network of YOLOv7 network model is mainly composed of convolution, ELAN module, MPConv module and SPPCSPC module. In addition to architecture optimization, the method proposed in this study also focuses on the optimization of the training process, focusing on some optimization modules and optimization methods.

2.3.1 *Re-parameterization*

The model re-parameterization technology combines multiple calculation modules into one in the reasoning stage. The weights of the models under different iterations are weighted and averaged. This method divides a module into several module branches during training, and integrates multiple branch modules into a completely equivalent module during reasoning.

2.3.2 *Label assignment*

At the same time, the YOLOv7 proposes a new label allocation method, called coarse to fine. The researchers call the head responsible for the final output the lead head, and the head used for auxiliary training the auxiliary head. The coarse to fine hierarchy of tags is generated, which are used for auxiliary head and lead head learning, respectively. YOLOv7 proposed a deep supervision label allocation strategy, as shown in the Figure 2 below.

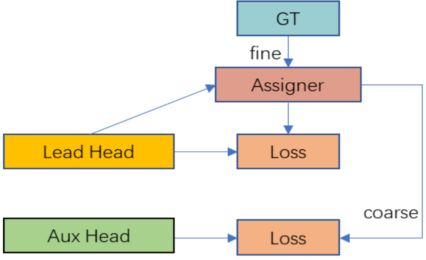

GT
fine
Assigner
Lead Head
Loss
coarse
Aux Head
Loss

Fig. 2 Label assignment

3. Improvements

3.1 *Backbone improvement*

In this paper, The first and fourth ELANs in the backbone are replaced by the combination of C3 structure and CNeB (ConvNeXt Block). Experiments show that CNeB has reached the limit of pure Conv network design. CNeB uses large core 7*7 DWCONV to replace ordinary convolution. At the same time, although BN layer can improve convergence and reduce over fitting, it still has many complexities that will adversely affect the performance of the model, so LayerNorm is used to replace BN. Finally, replacing ReLU with GELU makes FLOPs reduce to 2.66G during training while improving accuracy. The structure diagram of CNeB is shown in Figure 3.

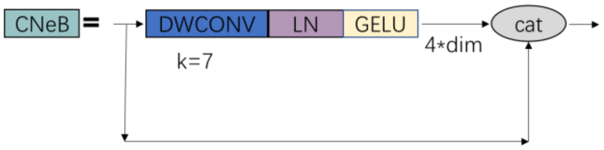


Fig. 3 CNeB structure diagram

3.2 *Head improved*

Attention mechanism has been widely used in the research of deep learning in recent years. This article uses Swin Transformer Block[6] to replace ELAN structure in the head part of YOLOv7. For the target detection task, many studies have shown that the addition of attention modules in the network can improve the representation ability of the network model, effectively reduce the interference of invalid targets, improve the detection effect on the target of interest, and achieve high effect of the network model. When the traditional Vision Transformer detects the whole image, it will face the problem of too many parameters and too high video memory usage. Therefore, the Swin Transformer Block method is adopted. Divide the incoming data into several windows, and pay attention to the small windows to reduce the video memory occupation of Q * K operation, and reduce the computational complexity from O(N^2) becomes O(N). The specific implementation will be divided into n 7*7 patches when the data is transferred to the W-MSA. So before passing through the W-MSA, data must be filled to make it an integer multiple of 7. Swin Transformer uses window self-attention to reduce computational complexity. In order to ensure the connection between non overlapping windows, it uses shifted window self-attention to recalculate the self-attention after window offset. The Swin structure diagram is shown in Figure 4.

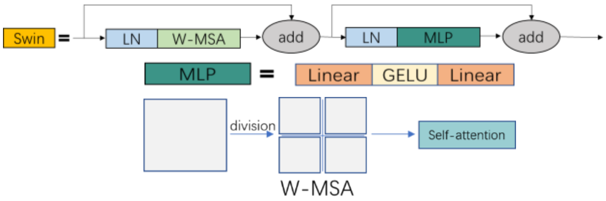


Fig. 4 Swin structure diagram

4. Experiment and software platform

4.1 *Experimental equipment*

The experimental platform is configured in Table 1.

Table. 1 Device-related configuration

Name	Version
Ubunt	20.04.3 LTS
CUDA	11.2
Graphics	GeForce GTX 1080 Ti
Frame	Torch

4.2 *Data collection*

The dataset used in the experiment was collected by ourselves. Each image contains multiple targets types. The dataset contains four types of target. They are people, faces, cars and license plates in cities. A labeling software is used for manual marking. There are 20,000 pictures in the experimental dataset. After selected, 5000 training sets and 500 verification sets are finally set.

4.3*Experimental setup*

The experimental parameters are set as shown in Table 2.

Table. 2 Experimental parameters

Name	Parameter
Pre-training	Yolov7_training.pt
hyp	hyp.scratch.p5.yaml
batch_size	16
epochs	300

4.4 *Experimental result*

In this paper, mAP (mean Average Precision) is selected as the evaluation index of the model. mAP is the average of all categories of AP (Average Precision). It usually reflects the performance of the model.

In this paper, under the same hardware conditions, we set the same parameters to conduct a comparative experiment between YOLOv7, YOLOX-s[7], YOLOv5-s and the improved YOLOv7. The experimental results are shown in Table 3.

Table. 3 Detection results of different algorithm

Method	mAP@.5	mAP@.5:.95
YOLOv5-s[4]	0.832	0.502
YOLOX-s[7]	0.885	0.511
YOLOv7[5]	0.909	0.516
our	0.949	0.537

Experiments show that the accuracy of the improved model is improved. After adding CNeb and Swin Transformer Block, the accuracy of the model is improved from 0.516 to 0.537. The inspection effect is shown in the Figure 5 below.

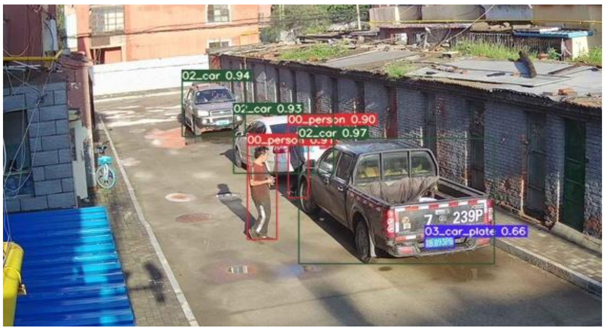


Fig. 5 detection results

5. Concluding

In order to improve the detection accuracy of people, faces, cars and license plates in the city, this paper chooses YOLOv7 as the main framework in its CNeB structure is added to Backbone to reduce model size and improve feature extraction capability, and Swin Transformer Block is added to head to ensure model parameters and improve detection accuracy. The experiment shows that the improved YOLOv7 has good performance in all aspects.

Acknowledgment

This work was supported by “Shenzhen Softsz Co. Ltd., China".

References

1. T. Zhang et al., "HOG-ShipCLSNet: A Novel Deep Learning Network With HOG Feature Fusion for SAR Ship Classification," in IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-22, 2022, Art no. 5210322, doi: 10.1109/TGRS.2021.3082759.
2. X. Xiao and X. Tian, "Research on Reference Target Detection of Deep Learning Framework Faster-RCNN," 2021 5th Annual International Conference on Data Science and Business Analytics (ICDSBA), 2021, pp. 41-44, doi: 10.1109/ICDSBA53075.2021.00017
3. V. R. A G, M. N and D. G, "Helmet Detection using Single Shot Detector (SSD)," 2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC), 2021, pp. 1241-1244, doi: 10.1109/ICESC51422.2021.9532985.
4. A. Albayrak and M. S. Özerdem, "Gas Cylinder Detection Using Deep Learning Based YOLOv5 Object Detection Method," 2022 7th International Conference on Computer Science and Engineering (UBMK), 2022, pp. 434-437, doi: 10.1109/UBMK55850.2022.9919478.
5. S. Liu, Y. Wang, Q. Yu, H. Liu and Z. Peng, "CEAM-YOLOv7: Improved YOLOv7 Based on Channel Expansion and Attention Mechanism for Driver Distraction Behavior Detection," in IEEE Access, vol. 10, pp. 129116-129124, 2022, doi: 10.1109/ACCESS.2022.3228331.
6. A. Yueyuan and W. Hong, "Swin Transformer Combined with Convolutional Encoder For CephalometricLandmarks Detection," 2021 18th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), 2021, pp. 184-187, doi: 10.1109/ICCWAMTIP53232.2021.9674147.
7. L. Feng and Y. Jia, "Traffic sign recognition based on YOLOX in extreme weather," 2022 Global Conference on Robotics, Artificial Intelligence and Information Technology (GCRAIT), 2022, pp. 299-303, doi: 10.1109/GCRAIT55928.2022.00070.

Authors Introduction

Mr. Zhihui Chen

He received his bachelor's degree from the school of electronic information and automation of Tianjin University of science and technology in 2021. He is acquiring for his master's degree at Tianjin University of science and technology.

Prof. Xiaoyan Chen

She, professor of Tianjin University of Science and Technology, graduated from Tianjin University with PH.D (2009), worked as a Post-doctor at Tianjin University (2009.5-2015.5).

Mr. Keying Ren

He is an Master of Control Science and Engineering, Tianjin University of Science and Technolog. The research topic is target detection and tracking based on deep learning.

A lightweight low-light image enhancement network

Yutao Jin
College of Electronic Information and Automation, Tianjin University of Science and Technology, No.1038 Dagu Nanlu,
Hexi District, Tianjin, China, 300222

Xiaoyan Chen
College of Electronic Information and Automation, Tianjin University of Science and Technology, No.1038 Dagu Nanlu,
Hexi District, Tianjin, China, 300222

Xiwen Liang
College of Electronic Information and Automation, Tianjin University of Science and Technology, No.1038 Dagu Nanlu,
Hexi District, Tianjin, China, 300222
E-mail: Charlesjyt@163.com, cxywxr@tust.edu.cn, 1540227453@qq.com
www.tust.edu.cn

Abstract

Existing methods based on deep learning have achieved great success in low-light image enhancement. However, the existing methods generally have a large amount of computation and poor generalization ability for low-light images in different scenes. In order to solve this problem, this paper explores a feasible way to introduce auxiliary blocks in the training process. These blocks can connect the feature map inputs of each stage to the inputs of the first stage to explore whether the model converges. In the prediction phase, due to the superior capabilities of the network, we use only one basic block for inference, so as to greatly reduce the computational cost. Owe to the flexibility of the network, the model can easily handle low-light images in a variety of complex environments.

Keywords: deep learning, low-light image enhancement, auxiliary blocks

1. Introduction

With the increasing availability of devices to capture images and the development of hardware such as impact sensors, people can take pictures in a variety of environments. However, it is difficult to obtain clear and reliable images in poor lighting conditions such as indoors and outdoors at dusk, night, and so on. The pictures taken in these conditions are called low-light images. Therefore, it is necessary to design a method to cope with the brightness of different scenes and enhance the low-light image. In the past decades, many low-light image enhancement algorithms have been proposed, which mainly go through three categories: histogram equalization, Retinex theory and convolutional neural network. Retinex-Net(Retinex Network)[1] is an image enhancement algorithm that combines convolutional neural network with Retinex theory. MBLLEN(Multi-Branch low-light Enhancement Network)[2] achieves low-light image enhancement by extracting the features of multiple branches and fusing them. EnlightenGAN(Enlighten Generative Adversarial Network)[3] is based on unsupervised learning GAN Network to achieve good effect. Compared with traditional algorithms, these methods based on deep learning have made a great leap in performance, but they still have some shortcomings, such as obvious noise and serious distortion, which can not achieve satisfactory results.

In this paper, we extract shallow features through the improved InceptionV2, and then design a sub-network as an auxiliary block to help the network to train. The feature fusion of the auxiliary block is used to guide the network learning and make the training more stable.
After this, we design a joint loss function to further consider the effects of image structure, brightness, detail and noise. In the last section, we prove that the network has excellent performance by comparing subjective and objective evaluation indicators with existing methods.

2. Method

Inspired by the work of Ma et al.[4], our network is composed of shallow feature extraction and three modules in a cascading manner. The image obtained through feature extraction needs to be spliced and transmitted to the middle module. Figure 1 shows the overall framework of our network. The input is a low-light image, from which we first obtain an shallow feature map using an effective strategy(see Section 2.1 for details). In addition, these intermediate modules share the same internal network structure(see Section 2.2 for details) and share weights during the learning process. The method we adopted is based on Retinex theory. This theory holds that the color of an object is determined by the object itself, and has nothing to do with the intensity and color of the light. The original image is obtained by multiplying the reflection and illumination images, as shown in Eq. (1)

$$S = R \times L \tag{1}$$

L is the irradiation component of solar illumination. R is the reflected component of the object itself. S is the imaging result. We modified the Retinex theory and put it at the end of each stage of the network. Furthermore, we define an auxiliary block that accelerates convergence by converging the results of each phase to the same state. The details of the network will be explained next.

2.1. *Feature extraction*

For the input image, we first extract its shallow features, which is very important for low-light image recovery. The extracted shallow feature contains rich details such as edges and textures of images. The network uses an improved InceptionV2 to extract features(as shown in Figure 2). Based on InceptionV2, we delete the average pooling layer and use dilated convolution to avoid the loss of feature information generated in the downsampling process and enlarge the receptive field of the network. The improved InceptionV2 has four branches, and each branch has dilated convolution of different sizes to obtain different scale receptive fields. Finally, shallow features are fused by feature splicing operation.

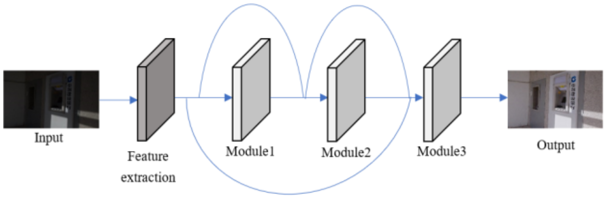


Fig. 1 Overall network framework

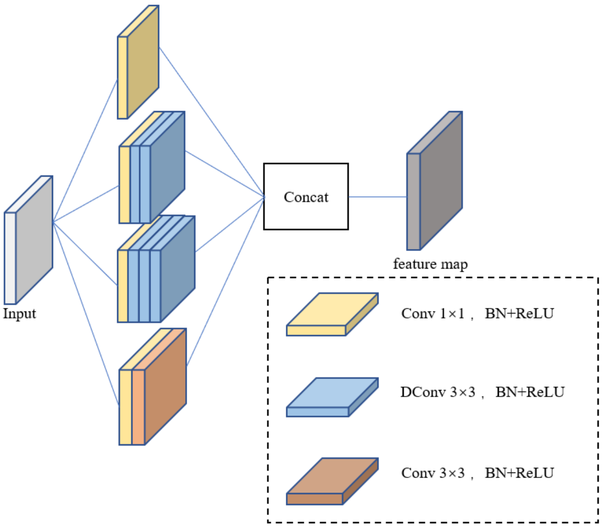


Fig. 2 Feature extraction module

2.2. *Intermediate module*

In the middle module, part of Retinex theory is used as the illumination adjustment block, and the other part is the auxiliary block. According to previous studies by scholars[4], there is a linear relationship between the irradiation component of solar illumination in Retinex and most regions in low-light images, and the irradiation component L is a core variable for restoring normal illumination. Thus, the exposure component is learned by introducing a model H_θ with a learnable parameter θ, as shown in Eq. (2)

$$F(L^t): \begin{cases} u^t = H_\theta(L^t), L^0 = S, \\ L^{t+1} = L^t + u^t, \end{cases} \tag{2}$$

u^t is the residual difference between the irradiation component of solar illumination at stage t and the low-light image. L^t is the component of the illumination at t stage. Since the weight of the network is shared, the model H_θ will not change significantly in each stage, so that the model H_θ learning residuals does not need to record the number of stages. The network of this model is composed of three convolutional layers, and the structure is shown in Figure 3.

In fact, the enhanced normal light image can be obtained by learning residuals from Retinex theory. However, the reasoning speed of the network will be slowed down because the network is weight sharing. Ideally, the first block will output the desired result, and the subsequent blocks will be similar to the first, so that only one block needs to be used for inference. This is then achieved by building auxiliary block.

The approach here is to link the inputs of each block to the first block for minimizing the differences between stages through a learnable parameter ξ. The auxiliary block is represented by Eq. (3)

$$G(L^t): \begin{cases} R^t = S/L^t, \\ E^t = K_\xi(R^t), \\ V^t = S + E^t, \end{cases} \tag{3}$$

V^t is the input processed by the auxiliary block at each stage. t is always greater than or equal to 1. Then, combining these two blocks (as shown in Figure 4) together as a Module in Figure 1.

2.3. *Joint loss function*

Low-light image has many problems, such as low brightness, color distortion and so on. It is not enough of using simple L1 or L2 loss in regression task to average all pixel differences. Therefore, a joint loss function is introduced to further consider the structure, brightness, detail and noise of the image. The total loss function is defined as Eq. (4)

$$L_{total} = \gamma_f L_f + \gamma_s L_s + \gamma_v L_v \tag{4}$$

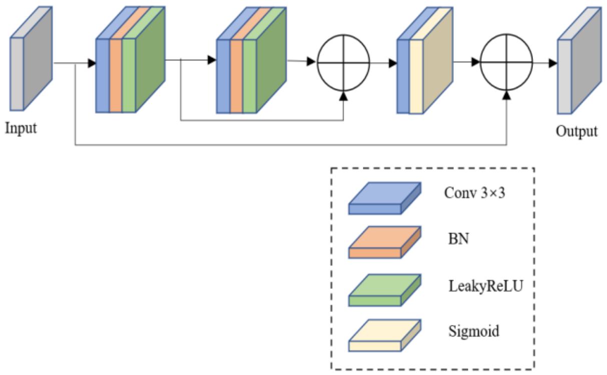


Fig. 3 Model $\boldsymbol{H_\theta}$

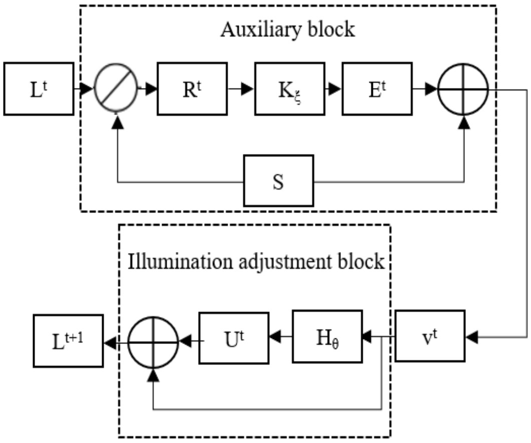


Fig. 4 An intermediate block

Where, L_f、L_s and L_t are fidelity loss, smoothing loss and perceptual loss respectively. γ_f, γ_s and γ_v are the weight balance coefficients, which are 0.8, 0.65 and 0.5 respectively. The fidelity loss is to ensure pixel-level consistency between the enhanced image and the input of each stage, as shown in Eq. (5)

$$L_f = \sum_{t=1}^{T} ||L^t - (s + E^{t-1})||^2 \tag{5}$$

T is the total number of stages. In the formula, the illumination component L^t is limited by $s + E^{t-1}$ to represent the normal illumination image. Real normal light images are not used in order to enhance the generalization of the model.

In addition, we used smoothing losses in our study, as shown in Eq. (6). The enhanced image noise is reduced and the image quality is better.

$$L_s = \sum_{i=1}^{H}\sum_{j=1}^{W}\sqrt{(p_{i,j}-p_{i+1,j})(p_{i,j}-p_{i,j+1})} \qquad (6)$$

Where H and W are the width and height of the image, and $p_{i,j}$ is the pixel value of the image.

In order to measure whether the enhanced image is real, image perception loss is introduced here. The basic method is to use the trained VGG16[5] as the future extractor. If the enhanced illuminance component is similar to $s+E^T$, the output from the VGG16 feature extractor should also be similar. The definition is shown in Eq.(7)

$$L_v = \frac{1}{W_{i,j}+H_{i,j}+C_{i,j}}\cdot\sum_{x=1}^{W_{i,j}}\sum_{y=1}^{H_{i,j}}\sum_{z=1}^{C_{i,j}}||\varphi_{i,j}(s+E^T)_{x,y,z}-\varphi_{i,j}(L^T)_{x,y,z}|| \qquad (7)$$

$\varphi_{i,j}$ is the feature graph of the j-th convolution layer of block i in VGG16 network. $W_{i,j}$, $H_{i,j}$ and $C_{i,j}$ are the dimensions of each feature graph in VGG16 network, respectively.

3. Experimental results and analysis

3.1. *Experimental detail*

The hardware platform of this experiment is configured with AMD Ryzen 7 5800H CPU and NVIDIA GeForce RTX 3060 GPU. Adam optimizer was used in the training stage, the parameters were $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\alpha = 10^{-8}$, and the learning rate was 0.0001. Iteration epoch is set to 800 times. The LOL-v2 dataset[6] is selected for training in this paper, which is larger and more diverse than the LOL dataset[1]. The dataset consists of 1589 image pairs, which are divided into two parts: synthetic image pairs and real image pairs. Most low-light images are obtained by changing the camera exposure time and ISO.

3.2. *Experimental result*

To test the effect of the model, MBLLEN and EnlightenGAN are selected as a comparison, and the subjective and objective evaluation index are analyzed. In order to show the algorithm effect more clearly, the original diagram is attached.

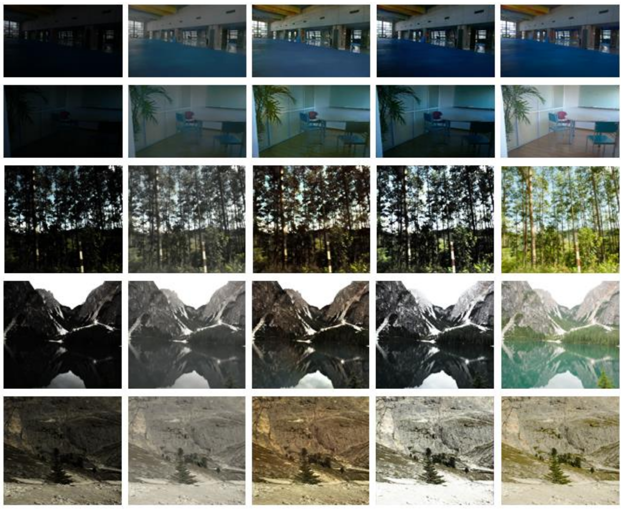


Fig. 5 A comparison of different enhancement algorithms

Tab. 1 Performance index

Methods	PSNR	SSIM
MBLLEN	16.88	0.719
EnlightenGAN	14.04	0.667
Ours	**17.02**	**0.732**

Tab. 2 Parameter quantity

Methods	Parameters(M)
MBLLEN	5.9527
EnlightenGAN	8.636
Ours	**0.0013**

The subjective effects of vision are evaluated. The comparison results of different algorithms are shown in Figure 5. We selected 5 pictures of indoor and outdoor with different scenes. From left to right, they are original low-light images, MBLLEN, EnlightenGAN, our method and the normal light images.

As can be seen from the Figure 5, the overall clarity of the image enhanced by MBLLEN is insufficient, and it loses the image color. Besides, the visual effect is poor, and there are a lot of noise. EnlightenGAN image color distortion is serious, and enhanced boundaries look not very true. The processing result of the algorithm in this paper has clear outline, less noise and moderate brightness. Comparison results show that the improved algorithm in this paper is more reliable and the overall visual effect is close to natural.

As can be seen from Table 1 and Table 2, on the LOL-v2 test set, the algorithm in this paper achieves optimal values in both PSNR and SSIM indexes. It indicates that the quality of low-light images enhanced by the algorithm in this paper is improved. It also shows that the algorithm parameters in this paper are small and it is a lightweight network.

4. Conclusion

In this paper, we design a low-light image enhancement algorithm using auxiliary block to help network training. The Retinex theory makes the reconstructed image better in brightness, and prominent in detail with less color distortion and closer to nature. The enhanced network is not only efficient but also lightweight with excellent performance in terms of image quality. In the follow-up study, we will continue to optimize the algorithm to further improve its ability to maintain color enhancement and other aspects.

Acknowledgment

This work was supported by "Tianjin University of Science and Technology – Softsz Technology Intelligent edge computing Joint Laboratory".

References

1. Wei, Chen et al. "Deep Retinex Decomposition for Low-Light Enhancement." *British Machine Vision Conference* (2018).
2. LV F, LU F, WU J, et al. MBLLEN : Low-light image/video enhancement using CNNs［C］//BMVC. 2018, 220（1）: 4.
3. JIANG Y, GONG X, LIU D, et al. Enlightengan : Deep light enhancement without paired supervision［J］. IEEE Transactions on Image Processing, 2021, 30 : 2340-2349.
4. Ma, Long et al. "Toward Fast, Flexible, and Robust Low-Light Image Enhancement." 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2022): 5627-5636.
5. Simonyan, Karen and Andrew Zisserman. "Very Deep Convolutional Networks for Large-Scale Image Recognition." CoRR abs/1409.1556 (2014): n. pag.
6. Yang, W. , et al. "Sparse Gradient Regularized Deep Retinex Network for Robust Low-Light Image Enhancement." IEEE Transactions on Image Processing 30(2021):2072-2086.

Authors Introduction

Mr. Yutao Jin

He received his bachelor's degree from the School of Information and intelligence Engineering of Ren'ai College of Tianjin University in 2022. He is acquiring for his master's degree at Tianjin University of science and technology.

Dr. Xiaoyan Chen

She, professor of Tianjin University of Science and Technology, graduated from Tianjin University with PH.D (2009), worked as a Post-doctor at Tianjin University (2009.5-2015.5). She had been in RPI, USA with Dr. Johnathon from Sep.2009 to Feb.2010 and in Kent, UK with Yong Yan from Sep-Dec.2012. She has researched electrical impedance tomography technology in monitoring lung ventilation for many years. Recently, her research team is focus on the novel methods through deep learning network models.

Mr. Xiwen Liang

He received his bachelor's degree from the school of electronic information and automation of Tianjin University of science and technology in 2021. He is acquiring for his master's degree at Tianjin University of science and technology.

An Improved Landweber Method for Electrical Impedance Tomography

Qian Wang

School of Electronic Information and Automation, Tianjin University of Science and Technology, No. 1038 Dagu Nanlu, Hexi District, Tianjin, China

Xiaoyan Chen

School of Electronic Information and Automation, Tianjin University of Science and Technology, No. 1038 Dagu Nanlu, Hexi District, Tianjin, China

E-mail: 1248265505@qq.com, cxywxr@tust.edu.cn

www.tust.edu.cn

Abstract

Electrical impedance tomography (EIT) is a non-destructive monitoring technique. The inverse problem of EIT shows serious nonlinear and ill-posed nature, which leads to the low spatial resolution of the reconstructed images. The iterative idea is an effective method to deal with the inverse problem for imaging, the existing iterative imaging methods, however, have many drawbacks, such as needing huge computational resources and unstable convergence properties. To solve the above-mentioned problems, this paper proposes an improved Landweber iterative image reconstruction framework. By carefully designing a regularization term in the optimization function of the iterative framework, this method improves the convergence speed. The tank experiment results show that the improved Landweber method is superior to the traditional Landweber method in convergence speed.

Keywords: EIT, Landweber, improved Landweber, regularization term

1. Introduction

Electrical Impedance Tomography (EIT) is a new imaging technique that has been developed in recent years. The alternating current with safe amplitude is injected into the observation domain, and then the boundary voltage responded to the surface is measured through electrodes placed on the surface of the object. The conductivity is calculated for each pixel inside the field to show the distribution of the different media. Some traditional imaging modalities, such as computed tomography (CT) and magnetic resonance imaging (MRI), have some limitations, for example, ionizing radiation, high price, and inconvenience to patient movement. These factors make it limited in practical clinical applications and prevent the implementation of long-term clinical monitoring. In contrast, EIT has the advantages of low cost, high temporal resolution, portability, and no radiation, which can effectively compensate for the shortcomings of traditional medical imaging modalities, and therefore has become an important research direction in the medical field with broad application prospects.

Currently, the core problem of solving the inverse problem of EIT, i.e., image reconstruction, is to construct a suitable image reconstruction algorithm to obtain an accurate image to describe the conductivity parameters and boundary features. In order to solve this problem, many image reconstruction methods have been studied. The image reconstruction methods researched over world are mainly divided into non-iterative and iterative categories. Among them, non-iterative methods have linear back projection (LBP)[1] and Tikhonov regularization[2], etc. The iterative methods have the conjugate gradient method[3] and Landweber method[4], etc. The LBP method is characterized by simple structure and fast imaging, but its spatial resolution is poor. It is suitable for online and fast qualitative imaging, but cannot provide accurate quantitative information. The regularization method is an effective method to overcome the ill-posed characteristic of EIT inverse problem. Nevertheless, the selection of the parameters and patterns of the penalty function is more complex, and it is usually based on an empirical setting. So, these methods have great limitations. The conjugate gradient method (CG) requires the system matrix which is a strict square matrix, and the EIT system matrix is an approximate result, so this approximation error will be magnified in the solution, and the reconstruction result contains artifacts, which cannot be applied to quantitative analysis. Landweber iterative method[5], based on the steepest descent principle, is the most commonly used iterative method for solving the problem of EIT ill-posed problems, but it requires multiple iterations and has a slow convergence speed.

Aiming at the problems of Landweber iterative algorithm requiring multiple iterations and slow convergence, this paper proposes an improved Landweber iterative algorithm. The method increases the convergence speed by adding regularization term to the objective function of EIT inverse problem. The experimental results of the tanks show that the improved Landweber iterative method outperforms the Landweber iterative method in terms of convergence speed.

2. Method

2.1. *Forward problem*

The construction of mathematical model is the basis of electrical impedance tomography reconstruction. The EIT simulation model measurement system is shown in Figure 1. This paper adopts the data acquisition method of adjacent current excitation and adjacent voltage measurement[6]. The specific acquisition method is to randomly select a pair of adjacent electrodes to inject current excitation as the initial excitation electrode, and then successively select two adjacent electrodes to measure the voltage as the measurement electrode. Using 16 pairs of adjacent electrodes as excitation electrodes in turn, 208 measurement data can be collected in total.

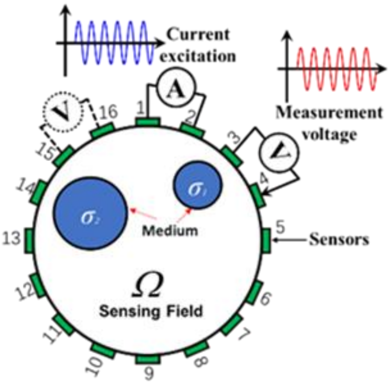


Fig. 1 EIT simulation model measurement system

In this paper, according to the complete electrode model (CEM)[7], the mathematical expression of the forward problem of CEM is as follows Eq. (1)

$$\begin{cases} \nabla\cdot(\sigma(x)\nabla\varphi(x)) = 0, x\in\Omega, \\ \varphi(x) = u(x), x\in\partial\Omega\setminus\bigcup_{L=1}^{16} e_L, \\ \sigma(x)\dfrac{\partial\varphi(x)}{\partial n} = 0, x\in\partial\Omega\setminus\bigcup_{L=1}^{16} e_L, \\ \int_{e_L}\sigma(x)\cdot\dfrac{\partial\varphi(x)}{\partial n}dS = I_L, x\in e_L, L=1,2,\cdots,16, \\ \varphi(x)+\rho_L\sigma(x)\cdot\dfrac{\partial\varphi(x)}{\partial n} = U_L, x\in e_L, L=1,2,\cdots,16. \end{cases} \tag{1}$$

where, $x\in\Omega$ is the spatial position, $\sigma(x)$ is the conductivity distribution in the measurement field, $\varphi(x)$ represents the potential distribution in the measurement field, Ω represents the measurement field, n represents the external normal vector on the boundary, $\partial\Omega$ represents the boundary information of the measurement area, and ρ_L represents the contact impedance between the electrode and the contact surface, U_L is the response voltage on the L-th electrode.

In order to solve the Laplace equation numerically with CEM boundary conditions, numerical techniques are needed. Generally, finite element method[8] is used to solve this problem. Assuming that the measurement noise is additive Gaussian noise, the observation model of EIT can be written as follows Eq. (2)

$$V = U(\sigma) + e \tag{2}$$

where, V is the vector of the measured voltage, $U(\sigma)$ is a forward solution based on the finite element method, and e is Gaussian noise.

2.2. *Inverse problem*

The mathematical model of EIT inverse problem can be expressed as Eq. (3)

$$y = Ax + b \tag{3}$$

where, y (matrix $y \in R^{m*1}$) represents the difference in boundary voltage values, and x (matrix $x \in R^{n*1}$) represents the difference in conductivity distribution. A is the sensitivity matrix, representing the mapping between conductivity change and boundary voltage. b represents additive noise.

Because the mathematical model of the EIT inverse problem is ill-conditioned, Eq. (2) cannot be solved uniquely and is usually calculated with the least square error. Namely Eq. (4)

$$\hat{x} = \arg\min_{x} \frac{1}{2}\left\| Ax - y \right\|_2^2 \tag{4}$$

2.3. *Landweber method*

Now we analyze the principle of Landweber method from a mathematical perspective[9]. Assume that matrix A_0 is the approximate matrix of matrix A^{-1}, and matrix P is the residual matrix of matrix A^{-1} and matrix A_0. Then there is formula Eq. (5)

$$P = I - A_0 A \tag{5}$$

Eq. (6) can be obtained from Eq. (5)

$$A^{-1} = (I - P)^{-1} A_0 \tag{6}$$

If the radius $\rho(P) < 1$ of the residual matrix P, it will be expanded. Step k is expanded as Eq. (7)

$$A_k = (I + P + P^2 + \cdots + P^{k-1}) A_0 \tag{7}$$

Because Eq. (8)

$$(I - P)(I + P + P^2 + \cdots + P^{k-1}) A_0 \tag{8}$$

According to Eq. (5), Eq. (6) and Eq. (8), we can get Eq. (9)

$$P^k A_0 = A_0 (I - AA_k) \tag{9}$$

If A_{k+1} is very close to A^{-1} in step k+1, we can get Eq. (10)

$$x_{k+1} = x_k + P^k A_0 y \tag{10}$$

Then by Eq. (9) and Eq. (10), we can get the iterative equation Eq. (11) of Landweber method

$$x_{k+1} = x_k + A_0 A^T (y - Ax_k) \tag{11}$$

In practical applications, the gain factor A_0 selection in the iterative method is very important because both the measurement voltage error and the bias brought by linearizing the nonlinear problem in the positive and inverse problems affect the final calculation results. In 2019, Han Guanghui demonstrated that the convergence of the function is faster and the imaging of the image is better when the gain factor is $A_0 = 2/(\lambda_{\max} + \lambda_{\min})$ (λ_{max} and λ_{min} are respectively the maximum and minimum eigenvalues of the matrix A^TA)[10].

2.4. *Improved Landweber method*

In order to improve the convergence speed of Landweber, a regularization term is added to the iteration Eq. (10) of Landweber method. The improved function is expressed as Eq. (12)

$$\hat{x} = \arg\min_{x} \frac{1}{2}\left\{ \alpha \left\| Ax - y \right\|_2^2 + \lambda \left\| x_k - x_{k-1} \right\|^2 \right\} \tag{12}$$

where, first item is the difference between the calculated voltage value and the measured voltage value. The second item is the difference between the conductivity distribution at the current time and the previous time. α is a gain factor. λ is a regularization term coefficient.

Combined with Eq. (11), the iterative formula of the improved Landweber iterative method proposed in this paper is as follows Eq. (13)

$$x_{k+1} = x_k + \alpha A^T (y - Ax_k) + \lambda \left\| x_k - x_{k-1} \right\|^2 \tag{13}$$

Step length in equation (11) α and λ It can be selected as a fixed value according to experience. Based on human experience $\alpha = 2/(\lambda_{\max} + \lambda_{\min})$ ($\lambda_{\max}$ and $\lambda_{\min}$ are the maximum and minimum eigenvalue of matrix A^TA respectively), $\lambda = 0.6$.

3. Experiment

3.1. *Evaluation metrics*

To quantitatively describe the imaging quality of the Landweber method and the improved Landweber method, the Relative Error (RE), Correlation Coefficient (CC) and Structure Similarity Index Measure (SSIM) of the reconstructed image are used as objective evaluation indicators. The calculation formula is as follows Eq. (14), Eq. (15) and Eq. (16).

$$RE = \frac{\left| x - \hat{x} \right|}{\left| x \right|} \tag{14}$$

$$CC = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(\hat{x}_i - \bar{\hat{x}})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \sum_{i=1}^{n} (\hat{x}_i - \bar{\hat{x}})^2}} \tag{15}$$

$$SSIM = \frac{2 * \frac{1}{n}\sum_{i=1}^{n} x_i * \frac{1}{n}\sum_{i=1}^{n} \hat{x}_i * \operatorname{cov}(x, \hat{x})}{\left[\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)^2 + \left(\frac{1}{n}\sum_{i=1}^{n} \hat{x}_i\right)^2\right] * \operatorname{cov}^2(x, \hat{x})} \qquad (16)$$

where, $\hat{x}$ represents the conductivity distribution of the reconstructed image, x represents the conductivity distribution of the real image, $\bar{\hat{x}}$ represents the mean value of the conductivity distribution of the reconstructed image, and $\bar{x}$ represents the mean value of the conductivity distribution of the real image.

3.2. *Experimental device*

The boundary voltage signal acquisition equipment used in this paper is composed of a high-precision and fast EIT data acquisition system (EIT-DAS), an image reconstruction system and a circular tank model. The schematic diagram[11] of the device is shown in Figure 2.

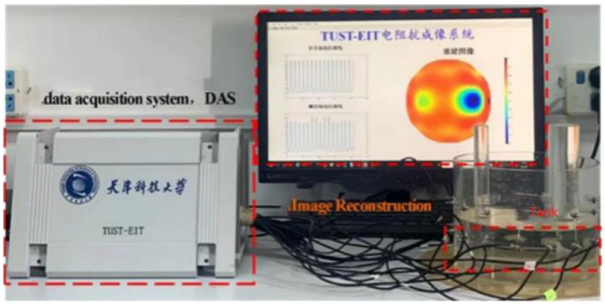


Fig. 2 TUST-EIT electrical impedance imaging System

The DAS data acquisition system consists of a host computer for data processing and image reconstruction, a data acquisition and signal processing module, and a tank with 16 titanium electrode sensors. A tank with a radius of 0.095m is used in the experiment, and 16 electrodes are distributed at the same interval and height on the tank wall. Uniform inclusions (glass rods) are placed in the tank, and the experimental data is obtained by changing the size and position of the inclusions. In the experiment, appropriate NaCl is added to tap water to obtain the background conductivity of the experiment (the conductivity is set to be about 1S/m), and the conductivity of the glass rod model can be approximately 0S/m. During the experiment, the current is excited through the hardware circuit. The amplitude of the exciting current is 4.5mA, and the frequency of the exciting current is 100kHz. Then the voltage between adjacent electrodes is measured to obtain the boundary voltage of the field.

3.3. *Tank Experiments*

In order to verify the performance of the improved Landweber method proposed in this paper, the Landweber method and the improved Landweber method will be used for image reconstruction. Four groups of tank experiments are set up in this paper. Tank 1 is inclusions with a radius of 0.02m and symmetrical position, tank 2 is inclusions with a radius of 0.015m and symmetrical position, tank 3 is inclusions with a radius of 0.02m and asymmetrical position, and tank 4 is inclusions with a radius of 0.015m and asymmetrical position. The image reconstruction results are shown in Figure 3。

As can be seen from Figure 3 and Table 1, when the boundary clarity, visualization effect and evaluation index values of reconstructed images are similar, iterations of the improved Landweber method decreased by more than 50% compared with the Landweber method. Therefore, the improved Landweber method proposed in this paper can greatly reduce iterations of image reconstruction and improve the speed of image

Table 1 Evaluation metrics and iterations of tank experiment

		Tank 1	Tank 2	Tank 3	Tank 4
Landweber	RE	0.4119	0.4331	0.3979	0.4115
	CC	0.6922	0.6137	0.7141	0.5849
	SSIM	0.6139	0.5379	0.6370	0.5229
	Iterations	**2500**	**3500**	**2400**	**2550**
Improved Landweber	RE	0.4107	0.4309	0.3965	0.4013
	CC	0.6944	0.6149	0.7173	0.7085
	SSIM	0.6162	0.5400	0.6399	0.6311
	Iterations	**1010**	**1420**	**1020**	**1020**

reconstruction.

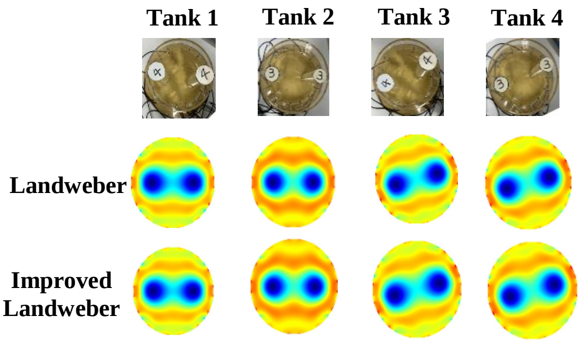


Fig. 3 Image reconstruction results of tank experiments

4. Conclusion

To solve the problem of the slow convergence speed of Landweber iterative method, an improved Landweber iterative reconstruction method is proposed. The method improves the convergence speed by adding regularization terms to the objective function of the inverse problem of EIT. From the results of the tank experiment, it can be seen that iterations of the Landweber method is more than double that of the improved Landweber method when the evaluation metrics of the reconstructed images are similar. In summary, the improved Landweber iterative method converges faster than the Landweber iterative method.

References

1. J. Solomon, D. Marinet, K. Roy Choudhury, et al., “Effect of radiation dose reduction and reconstruction method on image noise, contrast, resolution, and detectability of subtle hypoattenuating liver lesions at multidetector CT: filtered back projection versus a commercial model-based iterative reconstruction method,” Radiology, vol. 283, no. 3, pp. 777–787, 2017.
2. Y. Xu, Y. Pei, and F. Dong, “An adaptive tikhonov regularization parameter choice method for electrical resistance tomography,” Flow Measurement and Instrumentation, vol. 50, pp. 1–12, 2016.
3. L. Miao, Y. Ma and J. Wang, “ROI-Based Image Reconstruction of Electrical Impedance Tomography Used to Detect Regional Conductivity Variation,” in IEEE Transactions on Instrumentation and Measurement, vol. 63, no. 12, pp. 2903-2910, Dec. 2014, doi: 10.1109/TIM.2014.2326765.
4. S. Liu, L. Fu and W. Yang. “Optimization of an iterative image reconstruction method for electrical capacitance tomography[J],” Measurement Science and Technology, vol. 10, no. 7, pp.37-39, 1999.
5. W. Yang, Spink D M, York T A, et al., ”An image-reconstruction method based on Landweber′s iteration method for electrical-capacitance tomography [J],”Measurement Science and Technology, vol. 10, no. 11, pp.1065-1069, 1999.
6. D. C. Barber and B. H. Brown, “Applied potential tomography,” J. Physics E. Sci. instrum, Vo1.17, pp.723-733, 1984.
7. M. Cheney, D. Isaacson, E. J. Somersalo, et al., “Layer-stripping reconstruction method for impedance imaging,” 1992 14th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp.1694-1695, 1992, doi: 10.1109/IEMBS.1992.5761994.
8. P. V aukhonen, M. Vauhkonen, T. Savolainen, et al., “Three-dimensional electrical impedance tomography based on the complete electrode model,” IEEE Trans. Biomed. Eng, vol. 46, no. 9, pp. 1150–1160, Sep. 1999.
9. D. Liu, D. Gu, D. Smyl, et al., “B-Spline Level Set Method for Shape Reconstruction in Electrical Impedance Tomography,” in IEEE Transactions on Medical Imaging, vol. 39, no. 6, pp. 1917-1929, June 2020, doi: 10.1109/TMI.2019.2961938.
10. B. Grychtol and A. Adler, “FEM electrode refinement for electrical impedance tomography,” 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 6429-6432, 2013,doi: 10.1109/EMBC.2013.6611026.
11. R. Fu, Z. Wang, X. Zhang, D. Wang, X. Chen and H. Wang, "A Regularization-Guided Deep Imaging Method for Electrical Impedance Tomography," in IEEE Sensors Journal, vol. 22, no. 9, pp. 8760-8771, 1 May1, 2022, doi: 10.1109/JSEN.2022.3161025.

Authors Introduction

Qian Wang

She (1248265505@qq.com) graduated from Tangshan University with a bachelor's degree in 2018. She is currently pursuing a master's degree in School of Electronic Information and Automation, Tianjin University of Science and Technology.

Her research interests include electrical tomography (especially clinical applications of electrical impedance tomography), numerical solutions of inverse problems, and finite/boundary element methods.

Xiaoyan Chen

She, professor of Tianjin University of Science and Technology, graduated from Tianjin University with PH.D (2009), worked as a Post-doctor at Tianjin University (2009.5-2015.5). She had been in RPI, USA with Dr. Johnathon from Sep.2009 to Feb.2010 and in Kent, UK with Yong Yan from Sep-Dec.2012. She has researched electrical impedance tomography technology in monitoring lung ventilation for many years. She is in charge of the TUST-EIT lab and guides young researchers and graduate students to improve the electrical data acquisition hardware platform, to study the traditional and novel reconstruction algorithms with the prior structural information. Recently, her research team is focus on the novel methods through deep learning network models.

A Kinect-based Augmented Reality Game for Arm Exercise

Yoshimasa Tokuyama
Department of Digital Media and Product Design, Transworld University, No. 1221, Zhennan Rd., Douliu, Yunlin, Taiwan

R.P.C. Janaka Rajapakse
Graduate Institute of Animation and Film Art, Tainan National University of the Arts, No. 66, Daci, Guantian, Tainan, Taiwan

Junya Taguchi
Mimaki Engineering Co., Ltd., Nagano, Japan
E-mail: tokuyama.yoshimasa@gmail.com janakaraja@gmail.com invrbcilab@gmail.com
https://twu.edu.tw

Abstract

Augmented reality (A.R.) is the underlying technique where 3D virtual objects are integrated in real-time with a real environment. Augmented reality applications such as medical visualization, maintenance and repair, robot path planning, entertainment, military aircraft navigation, and targeting applications have been proposed. This paper introduces the development of an augmented reality game that allows the user to carry out arm exercises using a natural user interface based on Microsoft Kinect. The system has been designed as an augmented game in which the user's hands are in a world augmented with virtual objects generated by computer graphics. The player is sitting in a chair, just grasping the yellow stars that are displayed in the stage. It encourages the activities of a large number of arm muscles which will prevent decay. It is also suitable for rehabilitation.

Keywords: Augmented reality, Arm exercise, Rehabilitation, Healthcare

1. Introduction

The modern age has become an aged society, and the proportion of old people is increasing. The opportunity to move the body decreases when we get old, and the use of the brain also decreases. Just walking your legs will give you exercise, but if you don’t consciously move your arms, your muscles will gradually weaken without you even realizing it. Weakness of the arm muscles can lead to pain, such as a frozen shoulder. Besides, A typical rehabilitation program for stroke patients during hospitalization is paralyzed upper extremity functional training [1] (Figure 1). Also, when you stop using the brain, it may cause blur and dementia. Therefore, methods have been proposed to use the brain through games to prevent the hypofunction of the head and body of the elderly. For example, some traditional arcade games were used for rehabilitation purposes [2]. However, these arcade games are expensive and hard to relocate and use in a clinical space.

Augmented Reality (A.R.) systems have the advantage that information can be superimposed upon reality. The user can spatially relate virtual objects to the reality. Many researchers have proposed solutions for various applications, such as medical visualization, maintenance and repair, annotation, robot path planning, entertainment, military aircraft navigation, and targeting applications [3] [4]. Recently some Kinect-based AR systems were developed. Meng et al. [5] proposed a general method to interactively improve and correct the Kinect skeleton for A.R. Anatomy Learning. Anderson et al. [6] developed a system comprising a Kinect-based recording system and a corresponding training system with a large-scale augmented reality mirror.

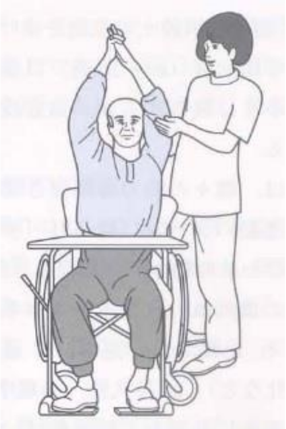


Fig.1 Elevation of both upper limbs

Casas et al. developed an Augmented Reality system for teaching key developmental abilities to individuals with A.S.D. The system has been designed as an augmented mirror where users can see themselves in a mirror world augmented with virtual objects [7]. Our previous work presented a framework for developing the Kinect-based augmented reality game for upper limb exercise and rehabilitation. We have got user opinions that were interesting and immersive [8].

In this paper, we present a Kinect-based augmented reality game for the exercise of the arms. The player is sitting in a chair and grasping the stars on the screen. It encourages the activities of a large number of the arm muscles. Using the Kinect device, the system can be operated with gestures. Another advantage of using Kinect is the simplified system improvement and the ability to modify the specifications at any time simply by rewriting the software source code. Therefore, although it is currently expressed only in Japanese, it is also possible to correspond to English or other languages.

2. System Overview

Figure 2 shows the experiment environment of playing our game for arm exercise. This is an augmented reality game in which gray stars and yellow stars are displayed on the stage, and users can naturally exercise their hands by grasping the yellow stars in turn (Figure 3, Figure 4). The purpose of the game in this study is to guide the participants to the movement of their hands naturally by making five reciprocations of holding the yellow star within the time limit. The stars are arranged in an arc along the Z-axis. As the level increases, the number of grab positions and the angle at which the arms are raised will increase. Figure 5 depicts the image of the Kinect device, and Figure 6 shows the flow chart of the game.

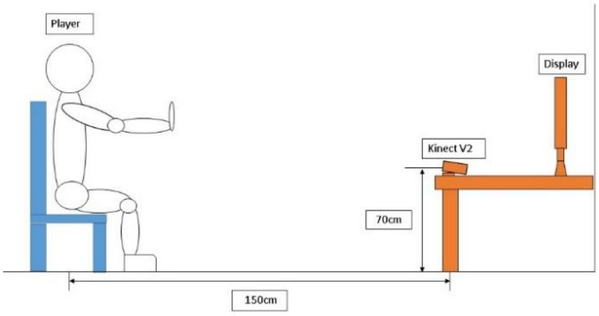

Player
Display
Kinect V2
70cm
150cm

Fig. 2 Experiment environment

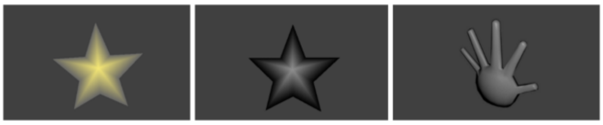


Fig. 3 Yellow star(left), Grey star(middle), left Hand model(right)

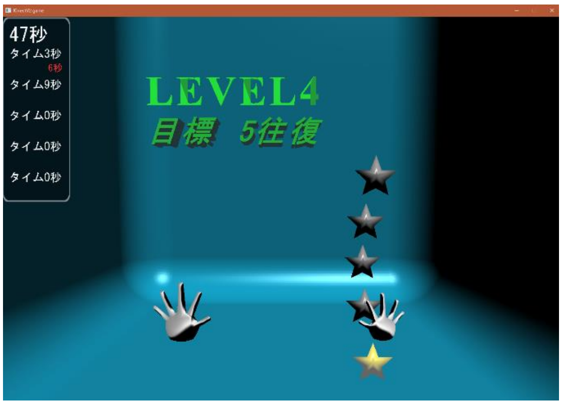

47秒
タイム3秒
6秒
タイム9秒
タイム0秒
タイム0秒
タイム0秒
LEVEL4
目標 5往復

Fig.4 The game screen

Fig. 5 Kinect device

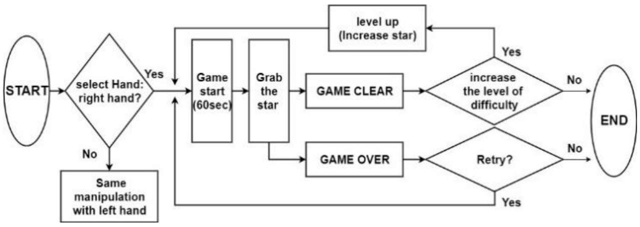

START
select Hand: right hand?
Yes
No
Same manipulation with left hand
Game start (60sec)
Grab the star
GAME CLEAR
level up (Increase star)
Yes
increase the level of difficulty
No
END
GAME OVER
Retry?
No
Yes

Fig. 6 Follow chart of the gameplay

2.1. *Arm coordinate position detection and placement of stars*

The game in this study is played with the arms outstretched. In order to reduce individual differences, such as difficulty in grasping stars due to differences in arm length depending on the body, we devised a method to detect the skeleton equipped with Kinect. We measure the coordinates of the neck, shoulders, elbows, and hands to determine the lengths of the upper arm, forearm, and shoulder width (Figure 7). Then, we calculate the arm length.

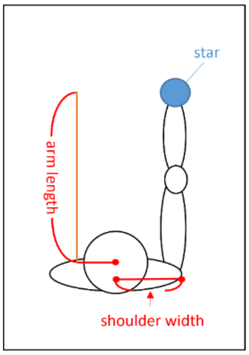


Fig. 7 Arm length calculation.

We position the stars along the arc marked by moving the stretched arm(Figure 8).

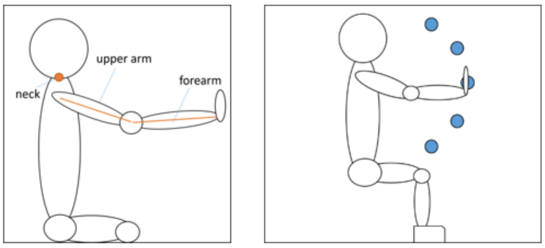


Fig. 8 The positions of stars

2.2. *Collision Detection*

The collision determination is defined when the distance between two centers is less than the sum of the radius of the two spheres which contain the object and palm Individually (Figure 9). Besides, we increased the judgment range to make it more responsive.

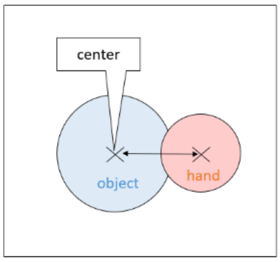


Fig. 9 Collision Detection

2.3. *Implementation*

This framework was implemented in Visual Studio 2010 development environment and used C++. And DirectX API is used for graphic rendering. To control the Kinect device, Kinect for Windows SDK toolkit is used in the implementations [9].

3. Results

Our framework was performed on a standard P.C. equipped with a 3.33 GHz Intel Xeon CPU and a NVIDIA Quadro FX 4600 graphics board with 768MB video memory. There are several characteristics that enhance usability, as stated in the following: Figure 10 shows the first screen that appears when the game is launched. At first, We measure the coordinates of the neck, shoulders, elbows, and hands to determine the lengths of the upper arm, forearm, and shoulder width (Figure 10).

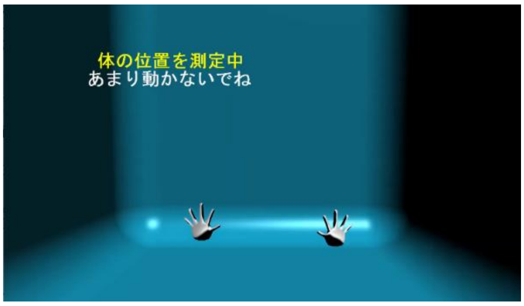


Fig. 10 Start Screen

Figure 11 shows the screen during the playing time. In the upper left position, the “time limit” and the time taken to grab each object is displayed. In the middle upper position, the current level and number of reciprocations for clear are displayed. In the middle lower position, two hands and the stars are displayed. If you clear the game for each level within the time limit of 60 secs, the “Goal Achievement” will be displayed, and you will be asked if you want to raise the level (Figure 12).

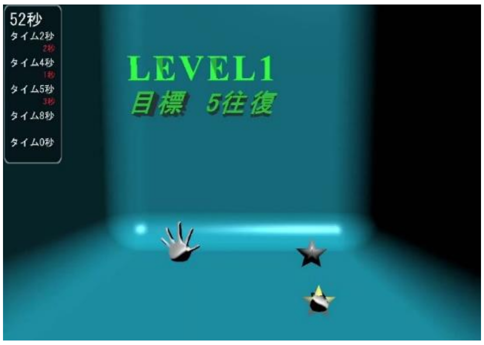


Fig. 11 Player Screen

The game ends when the user graps "no" at the right side of the screen. If you do not clear the game for each level within the time limit, you will be asked if you want to try again.

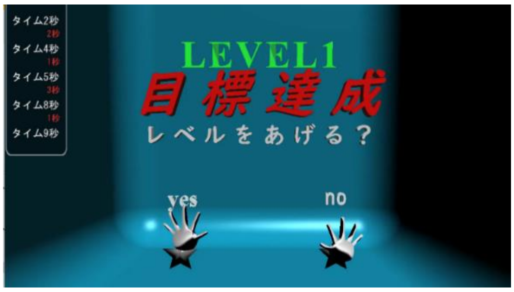


Fig. 12 Results Screen

In order to verify the usefulness of the proposed game, subjects (10 males and females of healthy subjects, all ages are in twenties) were asked to use the game and conducted a questionnaire about their feeling of use and concerns. Table 1 shows the results of that questionnaire.

Table 1. The result of the questionnaire

Good pinion	・Good for hand rehabilitation ・It can also be a hand gripper exercise ・Interesting to try
Point of concern	・Difficult to understand the depth ・Takes some time to get used to

4. Conclusion

In this paper, we have developed an augmented reality game for arm exercise through the use of a Kinect device. Hereafter, we will improve the concerns pointed out in the questionnaire. Since arm measurement is ambiguous due to camera blind spots, it is necessary to consider other measures, such as using two Kinect devices. It is necessary to make it easier for users to recognize the space in the game, such as changing the color of the stars when they enter the judgment range. Because the height at which the arm can be raised varies from user to user, it is necessary to match the placement of objects to the user. We will evaluate the affectivity of this game with the help of senior citizens in nursing homes and rehabilitation institutions. Future improvements and developments would be based on the requests of senior citizens or rehabilitation centers for specific exercise sessions.

References

1. Y. Sugawara. Occupational Therapy 1 – Disability. Kyodo Isho Publishing Co., Ltd., 2011.
2. H. Matsuguma, S. Fujioka, A. Nakajima, K. Kaneko, J. Kajiwara, K. Hayashida, F. Hattori, "Research and development of serious games to support stand-up rehabilitation exercise," Information Processing Society of Japan Journal, vol. 53, no. 3, pp. 1041–1049, Mar. 2012.
3. M. Billinghurst, A. Clark, and G. Lee, "A survey of augmented reality, Foundations and Trends in Human-Computer Interaction, vol. 8, no. 2-3, pp. 73-272, 2015.
4. D.W.F. van Krevelen and R. Poelman, "A Survey of Augmented Reality Technologies, Applications, and Limitations," The International Journal of Virtual Reality, vol. 9, no. 2, pp. 1-20, 2010.
5. M. Meng, P. Fallavollita, T. Blum, U. Eck, C. Sandor, S. Weidert, J. Waschke, N. Navab. "Kinect for Interactive AR Anatomy Learning," Proc. 2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR).
6. F. Anderson, T. Grossman, J. Matejka, G.Fitzmaurice. "YouMove: Enhancing Movement Training with an Augmented Reality Mirror," Proc. of the 26th annual ACM symposium on User interface software and technology, pp.311-320. 2013.
7. X. Casas, G. Herrera, I. Coma, M. Fernández. "A Kinect-based augmented reality system for individuals with autism spectrum disorders," Proc. GRAPP/IVAPP, pp. 440-446, 2012.
8. Y. Tokuyama, R. P. C. J. Rajapakse, S. Yamabe, K. Konno and Y. Hung, "A Kinect-based Augmented Reality Game for Lower Limb Exercise," 2019 International Conference on Cyberworlds, Oct. 2019.
9. Microsoft, Kinect – Windows app development. https://dev.windows.com/en-us/kinect.

Authors Introduction

Dr. Yoshimasa Tokuyama

He is a Professor at the Department of Digital Media and Product Design, Transworld University, Taiwan. He is also an emeritus Professor of the Department of Media and Image Technology at the Faculty of Engineering, Tokyo Polytechnic University, Japan. He received his Ph.D. in Computer Graphics from the University of Tokyo in 2000. His areas of research interest include computer graphics, haptics, virtual reality, shape modeling, game, and their applications. He is a member of the IPSJ, ITE, IIEEJ, and The Society for Art and Science.

Dr. R.P.C. Janaka Rajapakse

R.P.C. Janaka Rajapakse is an Associate Professor at the Graduate Institute of Animation and Film Art, Tainan National University of the Arts, Taiwan. And he is also a visiting scholar in the Media and Image Technology Department at the Faculty of Engineering, Tokyo Polytechnic University, Japan. He was a Postdoctoral Researcher at the Center for Hyper Media Research, Graduate School of Engineering, Tokyo Polytechnic University. He received his Ph.D. in Knowledge System Science from the Japan Advanced Institute of Science and Technology in 2008. His research interests include computer animation, motion capture, VR/AR/MR, haptic interfaces, AI, computer graphics, and Kansei Engineering. He is a member of the Motion Capture Society, The Society for Art and Science, ASIAGRAPH, and SIG-Design Creativity.

Mr. Junya Taguchi

He received his Bachelor's degree from the Department of Media and Image Technology at the Faculty of Engineering, Tokyo Polytechnic University, Japan. He is currently working as an Engineer at the Mimaki Engineering Co., Ltd., Nagano, Japan.

Development of EEG Based VR Application for Chakra Guided Meditation

Chien-Tung Lin, R.P.C Janaka Rajapakse
Graduate Institute of Animation and Film Art, Tainan National University of the Arts
No. 66, Daqi, Guantian Dist. Tainan City 72045, Taiwan
E-mail: feyin.tw@gmail.com, janakaraja@gmail.com
www.tnnua.edu.tw

Abstract

This paper studies how to properly meditate in a VR environment and increase the effect of chakra meditation by real-time EEG data. Meditation on the chakra is beneficial for the health of the body. For ordinary people, the chakra is invisible and unimaginable. Therefore, we developed the application in a VR environment. Users can sit inside a non-distracting environment with chakra visualization. The application can understand the condition of the user using the EEG data during the chakra meditation.

Keywords: Chakra, Meditation, EEG, VR, Mindfulness

1. Introduction

There are many meditation applications developed in the market that have a positive impact on people who are under tremendous pressure in modern times.

Virtual reality (VR) enables a person to be isolated from the environment and enter a separate space, thus reducing distractions. However, the majority of VR meditation applications use too many visual effects. It even places the users in a beautiful simulation world. Users find it difficult to focus on meditation. Limit stimuli are used only in a small number of applications but produce better results [1]. Therefore, this application is designed to utilize only the necessary visual guidance. Users can concentrate on meditation.

There are many methods of meditation. In this application, we introduce chakra meditation. Human life exists in two parallel dimensions, the ‘physical body’ and the ‘subtle body.’ In Figure 1, the subtle body consists of Nadi (energy channels), which are connected by the nodes of the energetic chakra (“cycles” or “wheels”) [2].

For the general public, it is difficult to imagine the chakras. We show the position of the chakra in a human model that sits on the opposite side. Guides users on the chakras and positions they must focus. Colors, patterns, and sound frequencies that correspond to the chakra are also used to enhance the effect of meditation.

Most meditation applications guide users by voice or visualization. However, it is a one-way guide. No

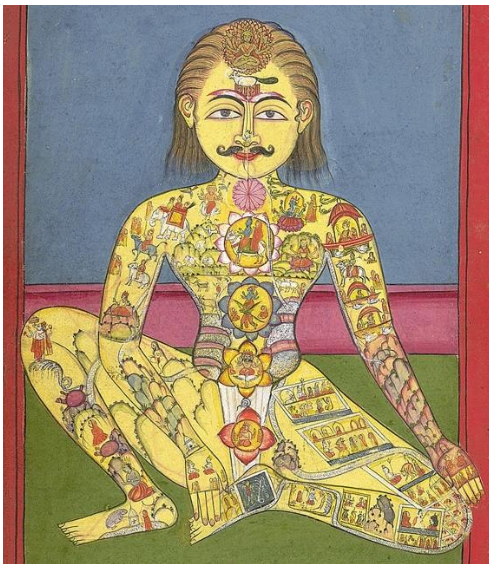


Fig. 1. Chakras from the 17th Century.

feedback method is available, and it is impossible to change according to the condition of the user. Therefore, the EEG is used in our application to determine whether the user is focused on meditation. It is used as a reference guide in the chakra meditation application.

2. Background

2.1. *Mindfulness meditation*

Positive meditation has been reported to be relaxing [3][4] and to improve physical health [5]. We often have a problem of thought that considers the past or the future; therefore, we need special training on it.

The primary purpose of mindfulness meditation is to focus on the present without judgment [6]. Our mind often thinks about the past or the future, so it needs special training and methods. Many forms of meditation have been developed since ancient times. Such focus is the breath, thought, or candlelight, contemplating a flower within, etc. The goal is essentially the same: to heal and relax the mind, body, and spirit. It is beneficial for maintaining physical and mental health.

Although chakra meditation has not become a method of mindfulness meditation, the goal is the same. In addition, chakra meditation may achieve more significant benefits. The user has a more precise focus: focus on the chakra facilitates the user to let go of mental judgments. The functions of the chakra are related to the energy of the body, and their balance contributes to physical and mental health. For the average person, the concept of chakra is challenging to imagine and feel. Therefore, we rely on VR and visualization as guides for chakra meditation.

2.2. *Virtual reality for meditation*

The use of VR as meditation allows people to disconnect from their daily environment and enter a different world conducive to meditation. Users are easy to immerse themselves in a virtual world. It could be a computer-generated or an omnidirectional video of the real-world [7] [8] [9]. Figure 2 shows the screenshot of an application. However, it is better to reduce the interaction and distraction in the virtual environment (VE) [10]. Most meditation methods require the user to close their eyes for better results. Therefore, our designs tend to be minimally visually guided. We use a mannequin sitting on the opposite side of the room to show the positions of every chakra and use visual effects and sound to guide the user on which chakra to focus at the moment.

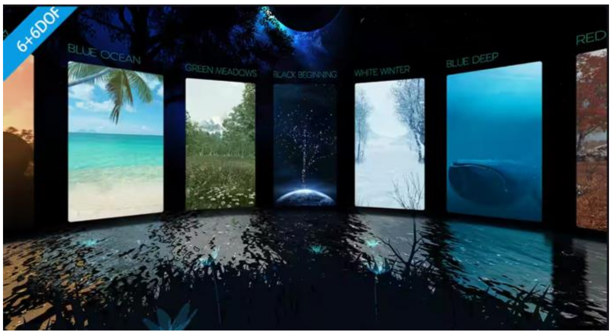


Fig. 2. A screenshot of Nature Treks VR [9].

2.3. *Chakra and frequency*

The theory of the chakra has been passed down from ancient times to the present. It is still present in modern society and is gaining more and more attention. The major chakras are arranged in the center of the body and correspond to the organs and secretory glands of each part of the body. They are also related to emotional aspects. Therefore, the balance of chakra energy helps to improve the state of mind and body.

The major chakras can correspond to different colors, sounds, and crystals. Optics and sound waves are different fields of frequencies. In quantum mechanics, the smallest fundamental particle is the string [11], which is vibration frequency. That is, the world is constituted by vibrational frequencies. Therefore, the use of corresponding frequencies in chakra meditation can be a real help. The frequency change of the sound also serves as a reminder, so that even with the eyes closed, the user can notice when it is time to switch to the next chakra.

2.4. *Electroencephalography (EEG)*

EEG sensors can detect brainwave signals, and it is possible to identify the user's attention [12]. Most applications are challenging to understand the users' status, have no feedback and are hard to quantify. The application can receive 500 EEG data per second and parse them through Looxid Link Core to provide concentration, relaxation, and balance of the left and right brains [13] [14]. It also provides alpha, beta, gamma, delta, theta, and other values updated every 100 ms. These data are transmitted to Unity programs via the unity API. We use the built-in EEG function to detect the user's concentration status in real-time and switch to the next pulse when a certain target is reached.

3. Design

3.1. *System architecture*

The development environment is an Intel I7 PC on a Windows 10 system. The graphic card is the NVIDIA GeForce RTX 2080TI. The programming language C# has been used to develop the application in the Unity SDK. The VR system is the VIVE Pro. We use the Looxid Link for the EEG device [13]. Its capabilities are easily integrated into the VIVE Pro. The system architecture is illustrated in Figure 3 and Figure 4.

3.2. *Chakras*

The location of the seven major chakras and the color and frequency of the reflected sound are related in Table 1.

The matching of the pulse to the color is certain. However, there are different theories for matching the sound's frequency. However, it is all based on the use of 432 Hz. To find the frequency Pn, the following definition may be used:

$$P_n = P_a \left(\sqrt[12]{2}\right)^{(n-a)} \qquad (1)$$

P_a : the frequency of a reference pitch (432 Hz).
n : numbers assigned to the desired pitch.
a : numbers assigned to the reference pitch.

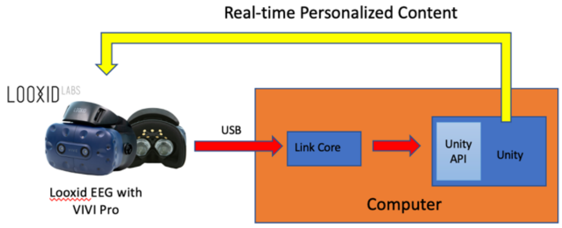


Fig. 3. Device and system structure.

The result of the frequency can be obtained in Table 2.

As shown in Figure 5, the octave note C as Root is used to derive the frequency corresponding to each chakra as the sound to be played during chakra meditation.

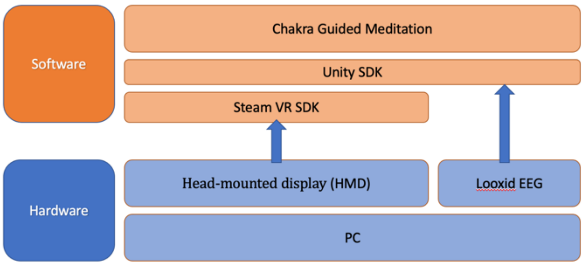


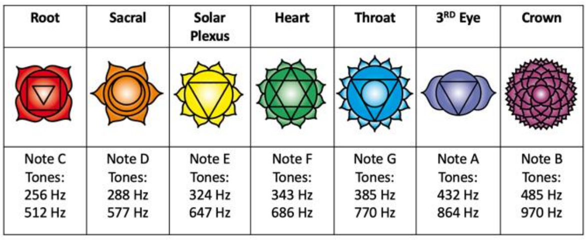
Fig. 4. System architecture.

Table 1. Chakra and attributes. (Source from M. Alcantara 2017 [15])

Chakra	Location	Color	Octave Note
Crown	Top of Head	Violet	B
3rd Eye	Brow Centre	Indigo	A
Throat	Base of Neck	Blue	G
Heart	Chest	Green	F
Solar Plexus	2 inches above navel	Yellow	E
Sacral	2 inches below navel	Orange	D
Root	Base of Spine	Red	C

Table 2. Pitches with a frequency of 432 Hz

Note	*Frequency (Hz)*
A3	216.00
A#3	228.84
B3	242.45
C4	256.87
C#4	272.14
D4	288.33
D#4	305.47
E4	323.63
F4	342.88
F#4	363.27
G4	384.87
G#4	407.75
A4	432.00



Root	Sacral	Solar Plexus	Heart	Throat	3RD Eye	Crown
Note C Tones: 256 Hz 512 Hz	Note D Tones: 288 Hz 577 Hz	Note E Tones: 324 Hz 647 Hz	Note F Tones: 343 Hz 686 Hz	Note G Tones: 385 Hz 770 Hz	Note A Tones: 432 Hz 864 Hz	Note B Tones: 485 Hz 970 Hz

Fig. 5. Chakra tone based on a 432 Hz grid.

How to guide and indicate the current chakra position of the user for meditation has been considered. Considering the need to keep the body upright during meditation, we place a seated figure in front of the user. In order not to let the user have extra imagination about this figure or ask who he is, we design his form as non-physical and unrecognizable. His body is marked with chakras. The user can follow the meditation in sequential order from the root chakra to the top chakra.

The application will display colors and totems of the chakras according to the current chakra to enhance the meditation effect, as shown in Figure 6. Totems will change size according to the user's EEG data and will change their size in real time. However, we have to be careful, as the visual effects may distract the user. In addition, to understand the user's meditation status, we use EEG to detect whether the user is concentrating on meditation and reaches the standard before moving on to the next chakra. The application is designed to achieve a concentration level of 70% and hold it for more than 1 minute. In addition to meditating from the root chakra, users are free to choose to meditate on a particular chakra. This prevents beginners from being stuck in the root chakra and unable to move forward.

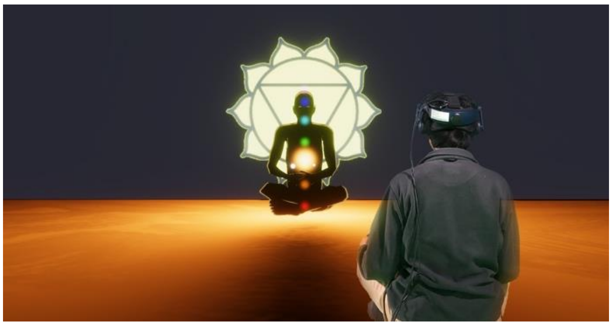


Fig. 6. User meditation in the application.

4. Result

Rick Ireton's study showed that some people are more sensitive to 432 Hz as the root chakra [16], so the application can provide options for users to choose the right frequency for themselves.

In addition, if only the sine wave frequency is used, it is not comfortable for the general public. Therefore, we use the sound of a Tibetan singing bowl as a sample [17] and then adjust the frequency according to the frequency table.

In the VR experience, how to create a safe and calm feeling in VR, but not too much visual impact. The first version was designed in a darker environment, which was scary for some people who thought they were going to play some scary game. Therefore, we can put firewood or candle flame in the scene, because the fire will make people feel warm and safe.

The EEG test needs more adjustment to make meditation effective for the user and not too difficult. We consider allowing the user to choose the difficulty level to achieve the right approach for the individual.

5. Conclusion

This application allows the user to perform chakra meditation in VR by guiding the user to perform a particular chakra meditation through video and sound frequencies. The EEG detects whether the user has reached the effective goal and proceeds to the next chakra. The images and sounds will not only guide but also help the chakra. We also try to use a minimum of visual effects as a guide to reduce distracting and flashy effects.

In the prototype version, we tested the effect of different sound frequencies on each chakra to find the right sound frequency. But we also need to maintain the flexibility to allow users to choose.

An important issue in VR is how to provide a comfortable and secure alternative without distracting the user. This sense of peace of mind varies from person to person, often depending on their own experiences. Therefore, it is necessary to provide different scenarios so that users can choose the right place for them. However, care must be taken not to become a world that the user wants to explore.

The last issue is about the use of EEG. Different levels should be added for users to choose and even the difficulty level can be adjusted flexibly with the users' use. If you are considering multiple users, you can set up to 3 users, record each user's usage status, and let the users adjust the difficulty level by themselves or automatically by the application.

References

1. C. Daudén Roquet and C. Sas, "Evaluating mindfulness meditation apps," in Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems, CHI EA '18, (New York, NY, USA), p. 1–6, Association for Computing Machinery, 2018.
2. D. G. White, Yoga in Practice. Princeton University Press, 2012.
3. C. Carissoli, D. Villani, and G. Riva, "Does a meditation protocol supported by a mobile application help people reduce stress? suggestions from a controlled pragmatic trial," Cyberpsychology, Behavior, and Social Networking, vol. 18, no. 1, pp. 46–53, 20
4. J. Carmody and R. A. Baer, "Relationships between mindfulness practice and levels of mindfulness, medical and psychological symptoms and well-being in a mindfulness-based stress reduction program.," Journal of behavioral medicine, vol. 31, no. 1, pp. 23–3
5. A. Howells, F. J. Eiroa Orosa, and I. Ivtzan, "Putting the 'app' in happiness: A randomised controlled trial of a smartphone-based mindfulness intervention to enhance wellbeing," Journal of Happiness Studies, vol. 17, 01 2015.
6. M. Speca, L. Carlson, E. Goodey, and M. Angen, "A randomized, waitlist controlled clinical trial: The effect of a mindfulness meditationbased stress reduction program on mood and symptoms of stress in cancer outpatients," Psychosomatic medicine, vol. 62,
7. BreatheVR, "Breathevr: A healthy mind has an easy breath.," 2021. https://www.discoverneon.com/breathevr". Accessed: 2022-12-30.
8. K. Shrikumar, "Playne: The meditation game.," https://www.playne.co. May 202. Accessed: 2022-12-30.
9. J. Carline, "Nature Treks VR" May 2019. https://www.oculus.com/experiences/quest/2616537008386430. Accessed: 2022-12-30.
10. R. M. Kelly, E. M. Seabrook, F. Foley, N. Thomas, M. Nedeljkovic, and G. Wadley, "Design considerations for supporting mindfulness in virtual reality," Frontiers in Virtual Reality, p. 155, 2022.
11. K. Becker, M. Becker, J. Schwarz, and P. Ramond, "String theory and m-theory: A modern introduction," Physics Today, vol. 61, p. 57, 05 2008.
12. N.-H. Liu, C.-Y. Chiang, and H.-C. Chu, "Recognizing the degree of human attention using eeg signals from mobile sensors,"sensors, vol. 13, no. 8, pp. 10273–10286, 2013.
13. "Looxid link." https://looxidlink.looxidlabs.com . Accessed: 2022-11-13.
14. G. Zwoliński, D. Kamińska, A. Laska-Leśniewicz, and Ł. Adamek, "Vibrating tilt platform enhancing immersive experience in vr," Electronics, vol. 11, no. 3, p. 462, 2022.
15. M. Alcantara, Chakra Healing: A Beginner's Guide to Self-healing Techniques that Balance the Chakras. Althea Press, 2017.
16. R. Ireton, "Chakra frequencies." https://chakrakeyacademy.com/chakrakey-frequencies. Accessed: 2021-12-18.
17. T. L. Goldsby, M. E. Goldsby, M. McWalters, and P. J. Mills, "Effects of singing bowl sound meditation on mood, tension, and well-being: an observational study," Journal of evidence-based complementary & alternative medicine, vol. 22, no. 3, pp. 401–406,

Authors Introduction

Mr. Cien-Tung Lin

He received his BS degree from the Department of Computer Science and Information Management, Providence University, Taiwan in 1999. He is a Graduate student at the Graduate Institute of Animation and Film Art, Tainan National University of the Arts, Taiwan. His research interests include virtual reality, interactive art, EEG, art therapy, computer animation and healing method.

Dr. R.P.C. Janaka Rajapakse

R.P.C. Janaka Rajapakse is an Associate Professor at the Graduate Institute of Animation and Film Art, Tainan National University of the Arts, Taiwan. And he is also a visiting scholar in the Media and Image Technology Department at the Faculty of Engineering, Tokyo Polytechnic University, Japan. He was a Postdoctoral Researcher at the Center for Hype Media Research, Graduate School of Engineering, Tokyo Polytechnic University. He received his Ph.D. in Knowledge System Science from the Japan Advanced Institute of Science and Technology in 2008. His research interests include computer animation, motion capture, VR/AR/MR, haptic interfaces, AI, computer graphics, and Kansei Engineering. He is a member of the Motion Capture Society, The Society for Art and Science, ASIAGRAPH, and SIG-Design Creativity.

A Study on Flower Patterns of Temple Cut-and-Paste Decorations based on L-system

Meng-Fan Huang
Tainan National University of the Art, No. 66, Daqi, Guantian District,
Tainan, Taiwan

Tzu-Hsien Yuan
Tainan National University of the Art, No. 66, Daqi, Guantian District,
Tainan, Taiwan

R.P.C. Janaka Rajapakse
Tainan National University of the Art, No. 66, Daqi, Guantian District,
Tainan, Taiwan
E-mail: owsla5520@gmail.com, chrisyuan.tw@gmail.com, janaka@mail.tnnua.edu.tw
https://www.tnnua.edu.tw

Abstract

"Cut-and-paste" is an architectural skill used in traditional Chinese temples to make various decorations by pasting tiles and glass. This traditional craft has more than 400 years of history, and each piece of work has noble artistry. This kind of craftsmanship is usually shown on roofs and walls of the temple, so it is also called "Art on the Roof." Due to the time-consuming or labor-intensive production and the aging of many craftsmen, this traditional handicraft is slowly disappearing. Although the paste decoration of the temple was registered and preserved by the Taiwan government many years ago, there are very few related studies, and many works are also in a state of being damaged and difficult to repair. Therefore, this study will discuss the decorative flower patterns in the paste decoration of the temple and analyze the growth of plants with the L-System. TouchDesigner visual development platform is used to create interactive works, hoping to transform traditional handicrafts into public art so that more people realize the importance of craft preservation.

Keywords: Cut-and-Paste decoration, Temple Art, Craft, L-System, Digital simulation.

1. Introduction

"Cut-and-paste" is a craft developed to decorate the roofs and walls of temples, in which "cut" involves cutting a bowl or glass into the desired size and shape, while "paste" involves using cement as a medium to glue the cut bowl or glass to the stucco sculpture (Figure 1). To present this craft requires a high level of skill, and it usually takes several years to complete the delicate decorative works, hence the name "Art on the Roof." These works are not just temple decorations but represent the iteration of Taiwan's traditional craftsmanship and

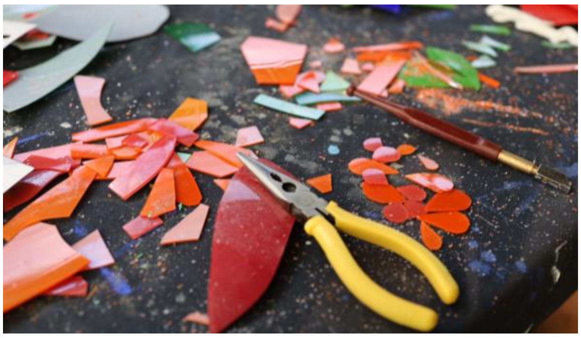


Fig. 1 Tools and materials for Cut-and-paste

are more like a dialogue with us through time and space. Even though the inextricable link between temples, faiths, and believers, the art of Cut-and paste went into decline with the high threshold of learning and the substitution of materials by mold making, coupled with the fact that the work is often decorated high up in temples, making it invisible to the public. The less the public cares, the harder it is to understand and pass on the importance of this culture. Although Cut-and-paste originated in China, the craft almost disappeared from the country after the Cultural Revolution[1], and Taiwan concerned them as an essential cultural asset that needs to be preserved. Currently, Cut-and-paste has been approved as a tangible cultural heritage in Taiwan but faces the dilemma of not having a well-established place for people to learn it and no related courses in the academy. In today's era of rapid technological development, many contents are being digitally transformed, but few traditional crafts are digitally reproduced. In this regard, our study will bring more attention to this cultural preservation issue through practical research and analysis in a cross-media approach, using technology to bring people closer to Temple roof craftsmanship.

2. Study Background

Cut-and-paste is not the temple's main structure, but the exquisite decoration is like putting on a unique style for the temple, bringing a different impression to the people. In the early agricultural society, education was not widespread among the people, temples as centers of faith were often decorated with themes from mythological stories and novels that conveyed the spirit of loyalty and filial piety, which not only had auspicious meanings but also served the purpose of educating people. Other works have a wide range of subjects, such as the common double dragons on the roof (Figure 2), the Seven Pagodas, and the Three Immortals of Fortunate Life. The walls are often decorated with flowers and birds, such as dragons and phoenixes, flowers and trees, birds and animals, and the Eight Immortals.

In the process of learning from existing craftsmen (Figure 3), we start with the basic decorative patterns, including the cutting of petal shapes, gluing techniques, flower shapes, and the composition of growing branches. As learned from the craftsmen, the "Chrysanthemum" is a traditional Chinese flower known as the Four Gentlemen of Flowers. Because of its allusive meaning of "longevity," it is used in many decorations. When observing Cut-and-paste chrysanthemum works, we can see that the flowers have a bright yellow color, and the petals were made by cutting and pasting with glass and have a slightly inclined angle. There are five to six layers,

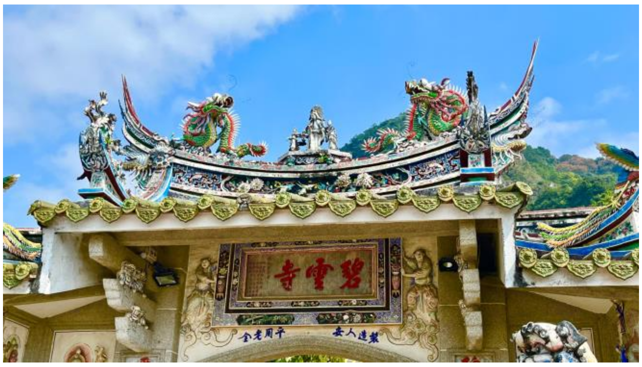


Fig. 2 Common Temple Roof Decoration

Fig. 3 The craftsman teaches cutting skill

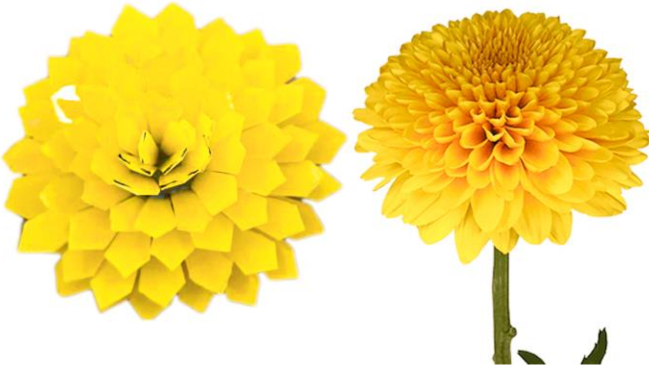


Fig. 4 Cut-and-paste chrysanthemum (Left), Real chrysanthemum (Left)

which slowly close together with the center. Although the shape of the petals of the cut-and-paste chrysanthemum has been altered to present them at a sharp angle, the reference form of this cut-and-paste work is a pom-pom type of chrysanthemum according to the above observation (Figure 4).

3. Digitalization

The preliminary phase of this research analyzed the most basic flower forms, including the expected selection of chrysanthemums and peonies (Figure 5) as the theme. The petals and leaves of the selected flower forms were simulated using mathematical procedural functions.

4. L-System

Lindenmayer system (L-System)[2] is a mathematical model of cellular interactions in growth and development. Because of the high similarity of plant models, it is easy to be defined, and such a rule can develop the growth structure of plants by increasing the number of layers of the L-System progression. For example, the plant-like structure generated by L-System (Figure 6) is defined by following Eq. (1).

In computer graphics and interactive content generation, L-System is widely used[3]. In the visualization software TouchDesigner[4], we tried to write different expressions of leaves and petals using the L-system, followed by fine-tuning the parameters to achieve more subtle changes.

$$
\begin{gathered}
n=5,\ \delta=22.5^{\circ} \\
X \\
X \rightarrow F-[[X]+X]+F[+FX]-X \\
F \rightarrow FF
\end{gathered}
\qquad (1)
$$

4.1. *Flower structures*

According to the peonies in the Cut-and-paste, many orange and red glasses are used to present the multi-layered petals. In the digitization process, the corresponding gradation of colors is used to reproduce the work, as shown in Figure 7.

4.2. *Plant structures*

Trees are often combined with floral patterns in temple decoration, and tree patterns can be simulated in L-System syntax. As Shown in Figure 8, the tree-like structures generated by L-System are defined by following Eq. (2).

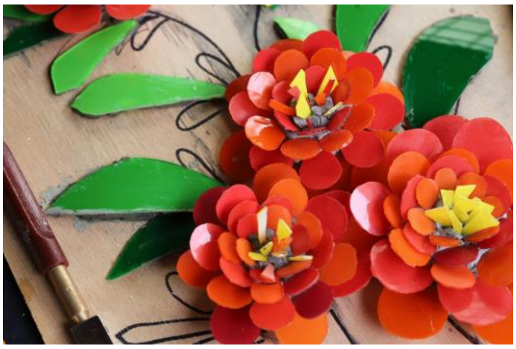


Fig. 5 Cut-and-paste peonies

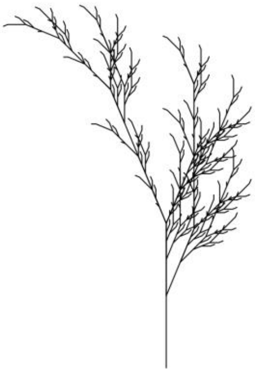


Fig.6 Plant-like structures generated by L-System [3]

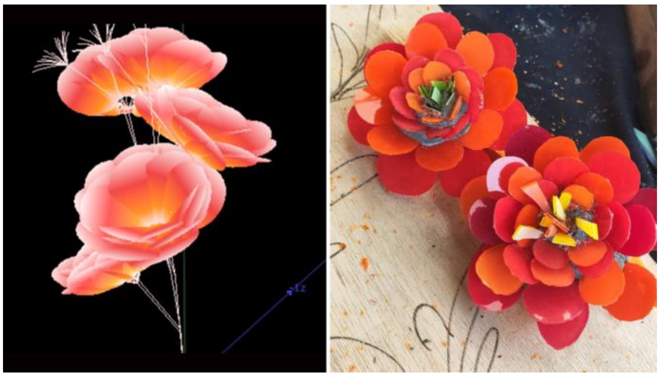


Fig. 7 Results of the peonies digitization

context_ignore: F+-
premise: A
A=";T~(4) F-[[A]+AJ]+F[-FA]+AJ (2)
F=~(3)!FF

4.3. *TouchDesigner*

Based on Cut-and-paste patterns and shapes of natural plants, the structures of flowers and plants were defined in the initial phase(Figure 9). Through the use of iteration, the L-system Surface Operator(SOP) of TouchDesigner allowed defining complex shapes of flower petals and tree structures. In the L-system SOP, an initial string of characters is evaluated repeatedly, and the generated geometry shapes were modified with Geometric nodes with instancing.

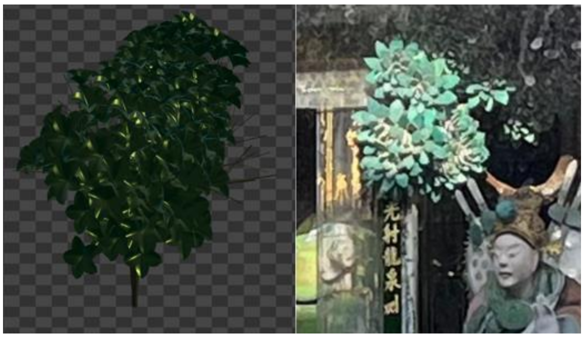


Fig. 8 Simulation of trees with L-system

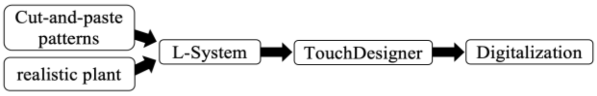


Fig.9 Digitization process through TouchDesigner

5. Conclusion

This study examined the decoration of Cut-and-Paste crafts in traditional Chinese temples. In conclusion, the observations indicated that natural landscapes and realistic plant growth patterns such as flower distribution, flower shapes, and branches were related to the mathematical representation of the cell interaction of visible growth and development. The current research is a preliminary analysis of the craft patterns to ensure they can be integrated with the L-system. In the future, the results of these findings will be transformed into public art to promote traditional handicrafts in a cross-media format.

References

1. Ding-Lin Chen, "A study of Cut-and-paste craft in Tainan. Tainan City: Tainan City Government," Bureau of Culture, 2016.
2. A. Lindenmayer, "Models for Plant Tissue Development with Cell division Orientation Regulated by Preprophase Bands of Microtubules," Differentiation, Volume 26, no. 1–3, 1984, pp. 1-10, https://doi.org/10.1111/j.1432-0436.1984.tb01366.x.
3. Przemyslaw Prusinkiewicz and Aristid Lindenmayer, "The Algorithmic Beauty of Plants", Springer-Verlag, Berlin, Heidelberg. ISBN:978-0-387-94676-4, 01 January 1996. 228 pages.
4. TouchDesigner, "A visual development platform" https://derivative.ca/

Authors Introduction

Ms. Meng-Fan Huang.

She received her B.Sc. degree from the Department of Journalism, Shih Hsin University, Taiwan, in 2021. She is a Graduate student at the Graduate Institute of Animation and Film Art, Tainan National University of the Arts, Taiwan.

Dr. Tzu-Hsien Yuan

He is an Assistant Professor at the Tainan National University of the Art in Taiwan. He received his Ph.D. in History from the University of Queensland, Australia, in 2017.

Dr. Janaka Rajapakse

R.P.C. Janaka Rajapakse is an Associate Professor at the Graduate Institute of Animation and Film Art, Tainan National University of the Arts, Taiwan. And he is also a visiting scholar in the Media and Image Technology Department at the Faculty of Engineering, Tokyo Polytechnic University, Japan. He received his Ph.D. in Knowledge System Science from the Japan Advanced Institute of Science and Technology in 2008. His research interests include computer animation, motion capture, VR/AR/MR, haptic interfaces, AI, computer graphics, and Kansei Engineering.

Pass-By: Development of Pedestrian Counts-based Art Installation for Passive Interaction

Jain-Lin Zhong
Graduate Institute of Animation and Film Art, Tainan National University of the Art, No. 66, Daqi,
Guantian District, Tainan, Taiwan

R.P.C. Janaka Rajapakse
Graduate Institute of Animation and Film Art, Tainan National University of the Art, No. 66, Daqi,
Guantian District, Tainan, Taiwan
E-mail: 1qazjk23@gmail.com, janakaraja@gmail.com
https://www.tnnua.edu.tw

Abstract

The trees in the streetscape grow as pedestrians pass by; we are always taking away and bringing in our surroundings without realizing it, perhaps some of the changes are very small, and we only notice them when they have grown to colossal size. This work presents a pedestrian counts-based art installation for multi-user passive interaction. This work focuses on the impact of people on their surroundings, which concerned a contemporary street scene projected in an open space and captured pedestrians' pass-by counts using an Arduino ultrasonic sensor. Based on the measured counts of pedestrians passing by, the developed application controls the color of the street scene and gradually decreases as the pedestrians cross. Based on the L-system mathematical model, the trees in the street scene progressively grow as the pedestrians cross.

Keywords: *L-system, Passive interaction, Interactive art, Touchdesigner*

1. Introduction

Most interactive artworks focus on tracking people's movements for instant reactions [1] [2] [3]. For instance, CutMod's Joy Displacement in Futura's Super Future event captures the human body movement and changes the image (Figure 1, left) [4]. Focus on real-time changes such as Daniel Rozin's Shiny Ball Mirror, two reflections being produced in front of the viewer using metal balls (Figure 1, right) [5]. This paper proposes an interactive art installation called "Pass-By," which does not create instant interactions and focuses on the long-term accumulated data for content generation.

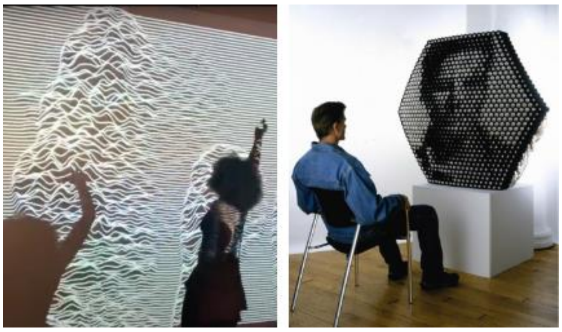


Fig. 1 Joy Displacement (Left) [4], Shiny Ball Mirror (Right) [5]

The aim of the Pass-By installation is to emphasize the less frequently overlooked elements of our lives, such as the intersections that we pass by on our way out of our homes or the trees we see on the streets. Therefore, the street in front of my house was captured as the background of the installation, and the count of pedestrians passing by was recorded using the computer as the passive interaction technique.

The Lindenmayer system (L-system [6]) was used to create a gradient between the background trees and the background color as more people passed. As a pedestrian passes, the changes in color and the growth of trees in the projected scene are much more minor and are not immediately visible. However, as time passes and pedestrians again walk through the installation, changes in background color and tree growth can be noticeable, as shown in Figure 2. Furthermore, since pedestrians are treated as separate entities in the street scene, and our installation was considered to include them in the street scene itself, the composition of the scene allowed the shadows of pedestrians to be cast in the background when they pass through this artwork. The camera captures the moment of a passing pedestrian recorded in a particular day and transforms into a passing pedestrian playing the next day's background.

2. Hardware structure

Figure 3 depicts the hardware structure of the installation, where the commodity webcam (Figure 4) was installed above the projector, recorded the pedestrians passing through information, and set up an ultrasonic sensor (Figure 4) below the projector to estimate the distance between the pedestrian and the projector. The installation maintained an appropriate distance between the projector and the projected screen, which allowed the audience's shadow to be cast on the screen and passively directed the audience to pass through.

3. Visual Composition

A webcam on the projector recorded the pedestrians' posture. TouchDesigner [7] visual development platform converted the captured poses into a grey silhouette and saved them as a second-day animation when the pedestrian passes through (Figure 5).

The distance r measured by the ultrasound sensor is transferred to the TouchDesigner using serial data. As shown in Figure 6, data r is categorized into three ranges in the TouchDesigner and splits the three ranges into three variables a, b, and c Eq. (1).

$$\begin{cases} a: r < 20 \\ b: 20 \leq r < 40 \\ c: 40 \leq r < 60 \end{cases} \tag{1}$$

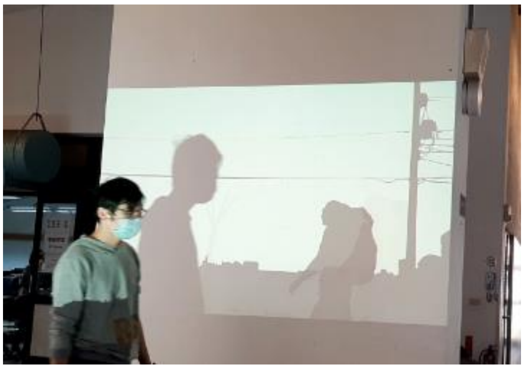


Fig. 2 Play the shadows of pedestrians as they pass by

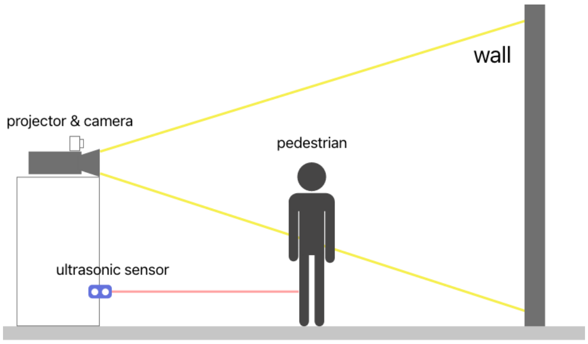


Fig. 3 Hardware structure

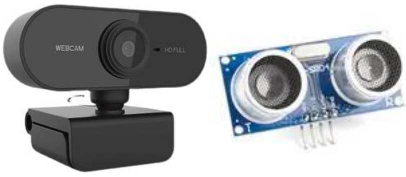


Fig. 4 Webcam and ultrasonic sensor

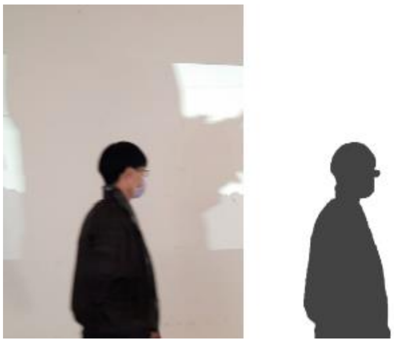


Fig. 5 Pedestrian transformed into silhouette

In addition, the number of people passing through three ranges, a, b, and c, is accumulated, and each person passing through the ranges increases the corresponding variable one by one. The three updated variables are exported to the projector view and generate changes.

3.1 *Color of background*

The first step is to input three variables a, b, and c into the RGB of the street scene, each with a negative value, the first variable a to red, the second variable b to blue, and the third variable c to green. The higher the value of the corresponding range, the more color loss occurs as each pedestrian passes through having a specific color, as shown in Figure 7.

3.1. *Grow tree by L-system*

L-system is an iterative function that uses mathematics. The central concept of L-systems is that of rewriting. Rewriting is a technique for defining complex objects by successfully replacing parts of a simple initial object using a set of rewriting rules or productions. Use this method to simulate the growth of plants and model. As seen from the Algorithm of Plant Beauty [6], the L system is complete in plant construction, such as creating a forest scene, the growth of a vineyard over a house, and a Lily of a valley.

In this work, the L-system is used to create a slow-growing tree in the scene, as shown in Figure 8. And design a, b, and c are three models of different branches, each of which corresponds to three variables of ultrasound output. The three branches grow to different degrees depending on the number of people passing through the respective regions.

4. Real-time self-localization method

The bottom layer of the projector is a street background (Figure 9). Only an empty street background and the trees above it are projected when no pedestrians pass by. When the ultrasonic sensor detects(i.e., at a distance r < 60) a pedestrian passing, an animation of the previous day's pedestrian passing is displayed on the upper layer of the background and trees. As shown in Figure 10, the top layer of the image is the shadow of the current pedestrian.

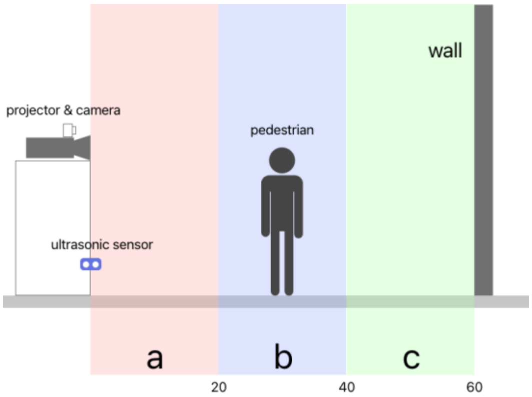


Fig. 6 Range of three intervals

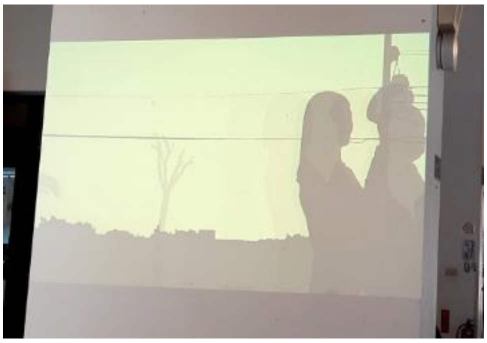


Fig. 7 Gradually yellowing of view

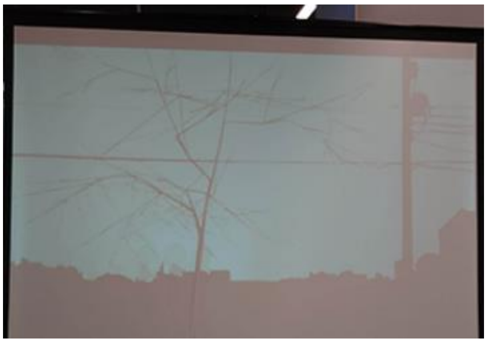


Fig. 8 The tree in the scene generated by L-system

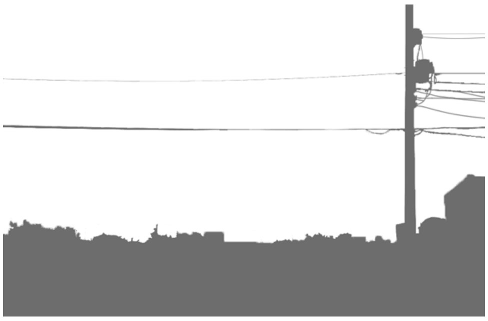


Fig. 9 Background

5. Conclusion

Pass-By can see the audience's own shadows projected onto the surface as part of the creation and contributes to the growth of the trees in the scene, with the appearance of the trees varying from one exhibition to the following (Figure 11). Some visitors stopped to inspect the trees when the Pass-By installation was exhibited at the testing exhibition. And few visitors were walking back and forth and trying to speed up the changes in the scene, especially children; they were waving their arms in front of the work to try to control the direction of the growth of the trees.

The future version would add the local weather conditions and affection of the sun's position to the scene's background, which would guide us to be more in tune with the time and place of the day and that each time and location will have its own characteristics.

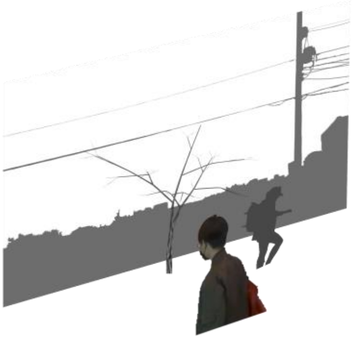


Fig. 10 Layers Sequence

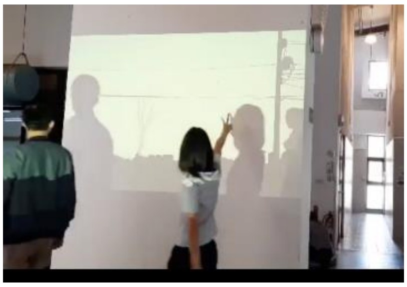


Fig. 11 Exhibition in the real world

References

1. David Z. Saltz, "The Art of Interaction: Interactivity, Performativity, and Computers," The Journal of Aesthetics and Art Criticism, vol. *55, no.* 2, pp. 117–127, 1997. https://doi.org/10.2307/431258
2. Dominic M. McIver Lopes. (2001). "The Ontology of Interactive Art." Journal of Aesthetic Education, vol. 35, no. 4, pp. 65–81. https://doi.org/10.2307/3333787
3. Xochihua, Alexis, "Looking to the Future of Education: A Social Art Practice Pedagogy" (2016). University Honors Theses. Paper 287, https://doi.org/10.15760/honors.302
4. Ralfonso Gschwend, "The Development of Public Art and Its Future Passive, Active and Interactive Past, Present, and Future," Journal of Arts, vol. 4, no. 3: 93-100, 2015. https://doi.org/10.3390/arts4030093
5. CutNod, Joy Displacement - Interactive Installation at Futra's Super Future event", https://reurl.cc/KXYpDq, Accessed 2022-12-31
6. Przemyslaw Prusinkiewicz and Aristid Lindenmayer, "The Algorithmic Beauty of Plants", Springer-Verlag, Berlin, Heidelberg. ISBN:978-0-387-94676-4, 01 January 1996. 228 pages.
7. TouchDesigner, "A visual development platform" https://derivative.ca/

Authors Introduction

Mr. Jain-Lin Zhong

He received his B.Sc. degree from the Department of Physics, National Taiwan Normal University, Taiwan, in 2021. He is a Graduate student at the Graduate Institute of Animation and Film Art, Tainan National University of the Arts, Taiwan. His research interests include interactive art, computer animation, projection mapping, human motion, and public art.

Dr. Janaka Rajapakse

R.P.C. Janaka Rajapakse is an Associate Professor at the Graduate Institute of Animation and Film Art, Tainan National University of the Arts, Taiwan. And he is also a visiting scholar in the Media and Image Technology Department at the Faculty of Engineering, Tokyo Polytechnic University, Japan. He was a Postdoctoral Researcher at the Center for Hyper Media Research, Graduate School of Engineering, Tokyo Polytechnic University. He received his Ph.D. in Knowledge System Science from the Japan Advanced Institute of Science and Technology in 2008. His research interests include computer animation, motion capture, VR/AR/MR, haptic interfaces, AI, computer graphics, and Kansei Engineering. He is a member of the Motion Capture Society, The Society for Art and Science, ASIAGRAPH, and SIG-Design Creativity.

Defect Solder Classification in Print Circuit Boards using Machine Learning

Watcharin Tangsuksant[1], **Jiraphan Inthiam**[2], **Hattapat Silangren**[3], **Phattharaporn Iamcharoen**[4], **Wisanu Jitviriya**[5]
Center of Innovation Robotics and Advanced Precision System: iRAPs,
Faculty of Engineering, King Mongkut's University of Technology North Bangkok,
1581 Pracharat 1 Road, Wongsawang, Bangsue
Bangkok, 10800, Thailand
E-mail: watcharin.t@eng.kmutnb.ac.th[1], jiraphan.i@eng.kmutnb.ac.th[2], hattaphat22642@gmail.com[3],
phattharaporniamcharoen@gmail.com[4] wisanu.j@eng.kmutnb.ac.th[5]

Abstract

This research proposes the solder inspection using the digital image processing technique and machine learning base with our machine vision prototype. There are five classes of classifying solder, including acceptable, short circuit, insufficient, blow hole and too much of solder. Automatic Optical Inspection (AOI) is used for the light source in the designed prototype and industrial camera which are installed on the mini-CNC. For the algorithm, this research applies the scanning line of binary image for detecting short circuit defection and the Random Forrest model for classifying other defects. According to the experiments, the system can classify the defect types for two classes (acceptable and unacceptable types) and five classes as 89% and 71% of accuracy, respectively.

Keywords: solder classification, solder defect, solder inspection, inspection machine

1. Introduction

Although, there have been developed automation system for the various tasks in the production line, soldering inspection on the Printed Circuit Boards (PCB) is still an important topic. There are many defected types occurring after wetting the solder, For example, solder bridging, insufficient solder, too much solder and blowhole. There defects result in damage to circuit on PCB. Moreover, the size of solders are very small that is quite hard for human vision and take a long time to inspect them, manually. The soldering inspection with the machine vision systems are proposed in the various solutions such as x-ray captured image [1], InSb IR camera [2]. Although, these are quite precise, the devices are expensive and complex. Another common solution is digital microscope capture that acquire the images from light source reflection. According to the light source types using for soldering inspection system, they can be categorized into two main types such as white and RGB light.

Based on white light reflection and digital microscope capture [3-8], there are different proposed method using digital image processing and some other machine learning techniques. For example, the YOLO is the real-time object detection that applies to detect the solder joint defects [3-4]. In 2020, Yu-Ting Li and *et al.* [3], YOLOv2 was used for detecting the foreground of the solders, then the 6 defect patterns was classified by using the ResNet-101. The accuracy was measured as 96.73% with 15 seconds per image. The updated version of YOLOv4, the algorithm processing time was faster than YOLOv2, was also applied for automatic inspection system for PCB solder joint by Ayhan Caliskan and *et al.* [4]. The 4,000 images approximately were train for increase accuracy as 97% with 4 seconds per image. Moreover, the Convolution neural network (CNN) is another option for object classification. According to the research of Shijia Gao et al. [5], the solder images were categorized into two classes as normal and abnormal. Then, the Support vector machine classifier (SVM) with the feature of Histogram of Oriented Gradient (HOG) was compared with the CNN solution. The result shown that the CNN had the higher accuracy as 85%. In addition, the traditional solution of image processing with OpenCV library such as image subtraction and blob detection are still usable, by Fa'Iq Raihan and Win Ce in 2017, [6] for the automated inspection system. Another interesting solution for the white light reflection, B.C. Jiang and *et al.* [7] proposed the machine vision and background remover-based that classified the solder defect such as the normal, short, open and no-solder. Analyzing the basic statistical values from the gray-level image for each defect types, using mean, median, range, standard deviation, and etc. The result was shown the 97.30% accuracy, but the feature selection and overlapped area of feature values problem were still difficult.

In order to add more information for light source reflection to the camera, the 3-color hemispherical LED array light source consisting of red, green and blue have been applied on many researches [9-12]. The various patterns of its reflection can be appeared on the image that depends on the curve and shape of solders. In 2020, Wenting Dai and *et al.* [9] applied the YOLO to detect the position of solder joint and semi-supervised learning for defect and non-defect classification of solders. The result is quite good for only 2-class by 1.5% error rate. Moreover, Yu Ting Li and Jiun In Guo [10] proposed the Faster RCNN model for PCP inspection using VGG16 as a pre-trained which result shown the mAP as 60% for the average of defect classification. In addition, the features can be extracted from each defect types. For example, place features, shape features such as color, occupancy ratio of area, center of gravity and continuous pixels [11]. Although, this proposed earned the good result, the case of solder joint fills with gas cavity within solder was still difficult and challenge. Ziyin Li and Qi Yang [12] designed the hardware prototype including illumination module, image acquisition module, motion control unit, PC, graphic display and operation unit. The prototype can detect many defects such as short and open circuit, wire gaps, voids, and scratch with around 95% of success rate.

This paper intent to present a system for detecting the solder defects on PCB that consists of a part of prototype design and classification algorithm. There are five types for classifying such as non-defect, short circuit, blowhole, insufficient solder, too much solder. The paper is organized as the follow. The proposed system including prototype design and classification is in the section 2. The experiments and results are in the section 3. Conclusion and discussion will be shown in section 4.

2. Proposed System for Solder Defects Inspection On PCB

The overview of system will be shown in Fig. 1, which have three main parts such as the vision, control, and classification in the personal computer (PC). All devices communicate to PC with the different protocol such as the GigE connecting the CMOS camera and USB for connecting the Automatic Optical Inspection (AOI) light source and controller.

The vision part consists of the AOI light source and GigE Industrial camera with 20 MP. Moreover, the AOI light source is an adjustable color-light mode between white and RGB color. The control part consists of the controller board, stepper motor driver and mini-CNC. This control part will be received the command as X, Y position from the PC for moving to the solder position on PCB. The process of classification, command and communication are done by PC.

2.1. *Prototype Design for Image Acquisition*

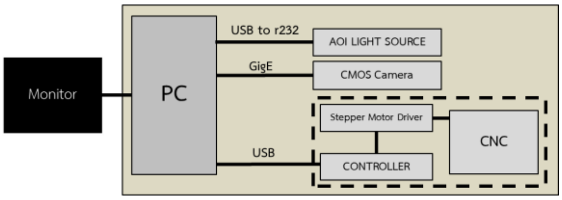


Fig. 1. Hardware system diagram

To acquire the solder images, this research designs the prototype as shown in Fig. 2. The position of each solder is saved on the PC initially. Then, the mini-CNC, that consists of three stepper motors for controlling moment along X, Y and Z axes, is commanded from PC as the position of X, Y position of solder on PCB via the Arduino UNO module. When the mini-CNC reaches the acquired position, the AOI light source and GigE Industrial camera, which are mounted on the Z-axis of the mini-CNC, will be worked as PC received command.

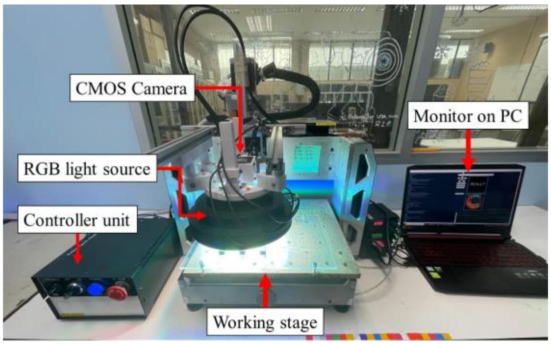
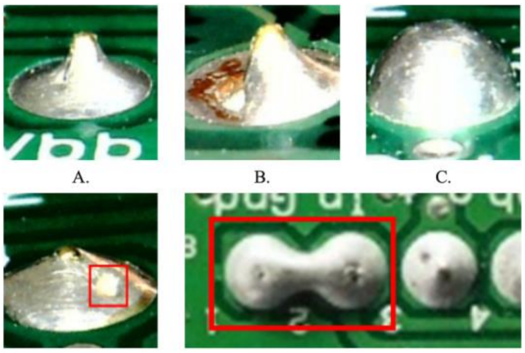


Fig. 2. Prototype design for Images acquisition

First is the camera will be captured the image with the white light for detect the short circuit defect. Then, the AOI is changed to RGB light mode and camera will capture the images for other defect type classification on PC.

2.2. *Solder Defects Classification*

This research categorizes the solder into five types, including the acceptable type. The solder types are acceptable, short circuit, insufficient, too much solder and blow hole as shown in Fig. 3. The proposed classified method is seperated into two solutions. Firstly, the basic techniuqe of image processing will be applied for detecting the short circuit defect. Second is to detect other defect types by using machine learning base.


A. B. C.

D. E.

Fig. 3. Four solder defect types and acceptable type (A. acceptable, B. Insufficient, C. Too much, D. Blow hole, E. Short circuit.)

In case of short circuit defect, the image is captured by white light mode. Then, the image thresholding and morphology techniuqe are used for pre-processing process. Then, the connected area of each solder will be checked by using contour techniuqe. Clearly, the short circuit area, in the image, is shown the huge connected area for comparing to the other solder pattern. Moreover, the position of short circuit will be detected by scanning the white pixel area of binary image between solder and its neighbor as shown in Fig. 4. If the summation of scaning line is greater than 200, that position will be defined for short circuit.

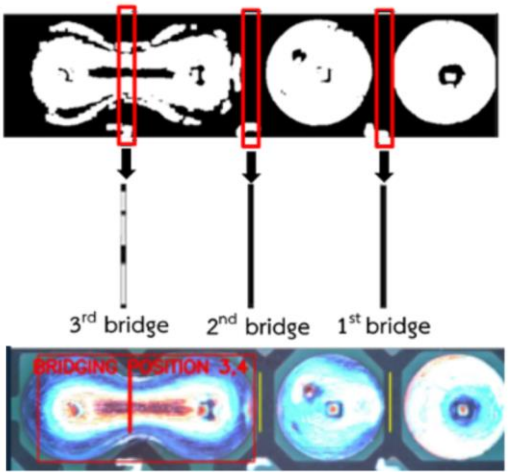


Fig. 4. Short circuit detection

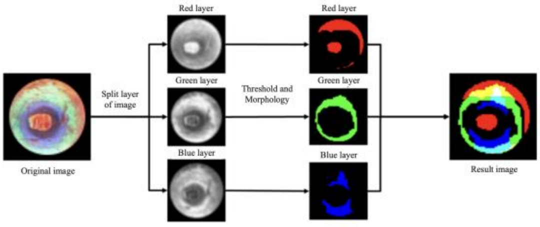


Fig. 5. Feature extraction of the solder image

Then, other defect types, including insufficient, too much and blow hole, are detected by machine learning techniuqe. This research apply the RGB light of AOI for more information whlie extracting features process. Fig.5 shows the feature extraction of solder image as the first step is seperation of image layer to red, green and blue. Then, thresholding value is set for segmenting each layer. Finally, these three layers are merged as the the feature for machine learning.

In order to extend the dataset of training process for machine learning. All merged images are augmented by rotating the angle from 20 to 340 degree, totally 408 new sample images are the dataset. Then, intensity values of the image size by 26×26 pixels. are preparing as the 676 vector feature and its lable.

3. Experimental Results

In the experiment, 7 print circuit boards are used as testing board, each board has 4 position solder pads. All circuit boards consist of acceptable type and defect type (Short circuit, Insufficient, Blow hole, and Too much).

3.1. *Verification I*

First verification of this proposed research is the comparison of classifier models using machine learning. The objective score is expressed in Eq. (1) which was used to evaluate the effective of classifier methods, where $\boldsymbol{\sigma}$ is the hyper-controlling factor ($\boldsymbol{\sigma = 0.04}$).

$$Objective\ score = \%\,Accuracy \times \left(\frac{Target\ time}{Process\ time}\right)^{\sigma} \quad (1)$$

Table 1. The comparison of classifier methods.

Classifier Methods	**Accuracy (%)**	**Target Time (s)**	**Processing Time (s)**	**Objective score (%)**
k-NN (k = 4)	87	0.3	0.39	86
k-NN (k = 1) with PCA	67	0.3	0.28	67
SVM	81	0.3	0.13	84
Decision Tree	73	0.3	0.75	70
Random Forest	**90**	**0.3**	**0.59**	**88**

From the experimental results in the Table 1, Random Forest had the maximum value of the objective score (88%). Therefore, It was selected as the classifier method in this research.

3.2. *Verification II*

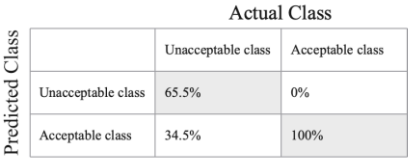
The confirmation of the suitable model was presented with the average accuracy of the system, including the confusion matrix. The experiment was repeated 10 times for all testing boards. In this research, the efficiency of solder defect classification was compared with 2-class classification (acceptable and unacceptable types) and 4-class classification (acceptable type and 3 defect types).

- **2-class classification**

The average accuracy of the system was 89%, F1-score accuracy was 0.79 and the confusion matrix is shown in Fig. 6.

- **4-class classification**

The average accuracy of the system was 71%, F1-score accuracy was 0.48 and Fig.7 presents the confusion matrix.

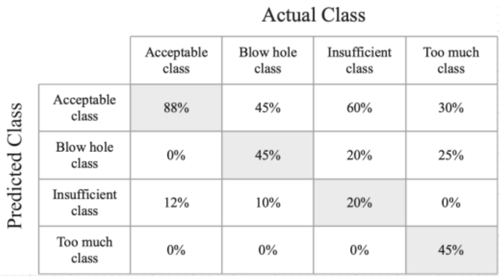


Actual Class

Predicted Class

	Unacceptable class	Acceptable class
Unacceptable class	65.5%	0%
Acceptable class	34.5%	100%

Fig. 6. Confusion matrix of 2-class classification



Actual Class

Predicted Class

	Acceptable class	Blow hole class	Insufficient class	Too much class
Acceptable class	88%	45%	60%	30%
Blow hole class	0%	45%	20%	25%
Insufficient class	12%	10%	20%	0%
Too much class	0%	0%	0%	45%

Fig. 7. Confusion matrix of 4-class classification

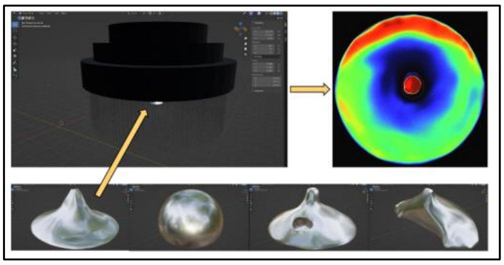


Fig. 8. The sample synthetic dataset

4. Conclusion and Discussion

From the experimental results, when Random Forrest was used to test the defect solder classification, it was found that the 2-class solder classification had higher percentage accuracy and F1-score than the 4-class classification. The non-defect type (or acceptable type) was the most accurate for the prediction, and the insufficient solder was the least accurate. The part of short circuit detection was used the morphology scanning pixels methods that is highly accurate because these types of defects are clear and distinction. For the future work, we aim to raise the accuracy of the inspection system using the synthetic dataset. In Fig. 8 shows the sample synthetic dataset that is created from Blender3D software.

Acknowledgements

This work was funded by Faculty of Engineering, King Mongkut’s University of Technology North Bangkok (Contract no. ENG-NEW-65-36). Received cooperation from TREND Electronics (Thailand) Co., Ltd.

References

1. Li, Yu-Ting, Paul Kuo, and Jiun-In Guo. "Automatic industry PCB board DIP process defect detection system based on deep ensemble self-adaption method." *IEEE Transactions on Components, Packaging and Manufacturing Technology* 11.2 (2020): 312-323.
2. Caliskan, Ayhan, and Guray Gurkan. "Design and Realization of an Automatic Optical Inspection System for PCB Solder Joints." *2021 International Conference on INnovations in Intelligent SysTems and Applications (INISTA)*. IEEE, 2021.
3. Jiang, B. C., C. C. Wang, and Y. N. Hsu. "Machine vision and background remover-based approach for PCB solder joints inspection." *International Journal of Production Research* 45.2 (2007): 451-464.
4. Ce, Win. "PCB defect detection using OPENCV with image subtraction method." *2017 International Conference on Information Management and Technology (ICIMTech)*. pp. 204-09, IEEE, 2017.
5. Raghuvanshi, Vediya, et al. "PCB solder pad inspection mechanism using gerber file." *2016 International Conference on Communication and Signal Processing (ICCSP)*. pp.1321-1325, IEEE, 2016.
6. Gao, Shijia, Hui Zhang, and Hanguang Mi. "Solder Joint Defect Detection Based on Image Segmentation and Deep Learning." *2019 IEEE International Conference on Signal, Information and Data Processing (ICSIDP)*. pp. 1-6, IEEE, 2019.
7. Dai, Wenting, et al. "Soldering defect detection in automatic optical inspection." *Advanced Engineering Informatics* 43 (2020): 101004.
8. Li, Yu Ting, and Jiun In Guo. "A VGG-16 based faster RCNN model for PCB error inspection in industrial AOI applications." *2018 IEEE international conference on consumer electronics-Taiwan (ICCE-TW)*, pp. 1-2, IEEE, 2018.
9. Wu, Fupei, et al. "An AOI algorithm for PCB based on feature extraction." *2008 7th World Congress on Intelligent Control and Automation*, pp. 240-247, IEEE, 2008.
10. Fan, Fangliang, et al. "Efficient Faster R-CNN: Used in PCB Solder Joint Defects and Components Detection." *2021 IEEE 4th International Conference on Computer and Communication Engineering Technology (CCET)*, pp. 1-5, IEEE, 2021.
11. Wu, Fupei, and Xianmin Zhang. "Feature-extraction-based inspection algorithm for IC solder joints." *IEEE Transactions on components, packaging and manufacturing technology* 1.5, pp. 689-694, 2011.
12. Li, Ziyin, and Qi Yang. "System design for PCB defects detection based on AOI technology." *2011 4th International Congress on Image and Signal Processing*. Vol. 4, pp. 1988-1991, IEEE, 2011.

Authors Introduction

Dr. Watcharin Tangsuksant

He is a lecture of Department of Production and Robotics Engineering at King Mongkut's University of Technology North Bangkok, in Thailand. He received his D. Eng. degree in Graduate School in 2019. His research interest is Asistive Devices/Application with Artificial Intelligence and Machine Learning.

Dr. Jiraphan Inthiam

He is a lecture of Department of Instumentation and Electronics Engineering at King Mongkut's University of Technology North Bangkok, in Thailand. He received his D. Eng. degree in Graduate School in 2019. His research interest is Robotics fields and Artificial Intelligence and Machine Learning.

Mr. Hattapat Silangren

He received his Bachelor of Engineering Program in Robotics engineering and Automation System from the Faculty of Engineering, King Mongkut's university of technology North Bangkok, Thailand, in 2022. He is currently a master's degree student at Kyushu Institute of Technology Iizuka Campus, Japan.

Ms. Phattharaporn Iamcharoen

She received her Bachelor of Engineering Program in Robotics engineering and Automation System from the Faculty of Engineering, King Mongkut's university of technology North Bangkok, Thailand, in 2022.

Asst. Prof. Dr. Wisanu Jitviriya

He is a lecturer in Department of Production and Robotics Engineering, King Mongkut's University of Technology North Bangkok, Thailand. He received his D. Eng. degree in Graduate School in 2016. His research interest that covers the design and analysis of robotic system combined with an artificial neural network

A dynamic nurse scheduling using reinforcement learning:
Dealing with various sudden absences of a nurse

Masato Nagayoshi
Niigata College of Nursing, 240 shinnan-cho
Joetsu, Niigata 943-0147, Japan
E-mail: nagayosi@niigata-cn.ac.jp, elderton@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

In nurse scheduling, whereby work schedules for nurses are created, it is very difficult to create a work schedule that satisfies all the various requirements. Hence, various studies have been conducted on the nurse scheduling problem. However, for practical use, adjustments including various constraints and evaluation values are required, as the created shift schedule is often not practical as it is. Therefore, we have proposed a work revision method using reinforcement learning on a constructive nurse scheduling system. In this paper, we extend the proposed method to dynamic nurse scheduling, in which the work schedule is revised or rescheduled when an absence occurs. Furthermore, we confirm whether or not the extended method can be used to create a work schedule that is feasible in various sudden absences of a nurse.

Keywords: dynamic nurse scheduling, reinforcement learning, constructive search, sudden absence

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 extend the proposed method to dynamic nurse scheduling, in which the work schedule is revised or rescheduled when an absence occurs. Furthermore, we confirm whether or not the extended method can be used to create a work schedule that is feasible in various sudden absences of a nurse.

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 take into account detailed conditions.
3. It does not take into account the evaluation value for the entire shift schedule for a month.
4. Based on the priority of job *j* for nurse *n*, the problem is transformed into a N × N assignment problem and can be solved efficiently using the Hungarian method. Note that jobs are prepared for

the number of nurses by dividing work shift *w*, including holiday, by the group *g* to which the nurses belong.

2.2. *Work Revisions*

The constructive scheduling system considers only the basic constraints that would be required in a hospital with a large number of 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. In this case, it is important that the constraints are satisfied by the exchange. If an exchange is made that does not satisfy the constraints, a warning message is displayed.

(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. Of course, the next solution displayed is not necessarily a satisfactory solution, but the above procedure is repeated in a timely manner until a satisfactory solution is obtained.

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:

$$min \sum_{n} \sum_{v} \mathrm{V_{nw}} \qquad (1)$$

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 considered to be 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: $\mathrm{V_{nw}}$ (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, $\sum_{n} \sum_{v} \mathrm{V_{nw}} = 0$ (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 before a sudden absence occurs.

4. Dynamic Nurse Scheduling Using Reinforcement Learning

When an assignment of work shift has to be changed suddenly due to an absence, etc., an alternative nurse is first secured according to the following procedure. In the case that a change of day shift is unavoidable, a change to the night shift (evening or late-night shift) may be considered, but only holiday is considered in this paper.

(1) Designate the work shift w_0 of the absent nurse n_0 on the day of the absence d_A as the source of the exchange, and designate the destination work shift w_1 as holiday.
(2) Deduce the group $g(j_0)$ in which the absent nurse n_0 is in charge of a job j_0, which is assigned as the shift w_0 on the day d_A.
(3) Determine a nurse n_1 who belongs to group $g(j_0)$ and whose shift on the day d_A 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_A. there is no applicable nurse, determine the nurse n_1 by changing the source work shift w_0 to day shift, late-night shift, and evening shift, in that order.
(4) The nurses n_0 and n_1 are exchanged their shifts on the day d_A. Then, change the work shift of the nurse n_0 to a designated work shift so that the work shift of the nurse n_0 will not be exchanged any more.

Next, the exchange procedure is obtained by reinforcement learning in the same way as the procedure described in Section 3.1 after the occurrence of an absence. However, the day with the highest priority is determined after the absence occurrence date d_A.

5. Computational Experiments

5.1. *Nurse Scheduling Problem*

The extended method is applied to a nurse scheduling problem similar to 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.

Table 1. Evaluation of shift 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

Next, Table 1 shows the evaluation of shift patterns for 2 days with M = 2.

5.2. *RL Agent*

In the state space of the RL agent, the total number of violations is assumed to be [0, 2] and can take 3 states.

Table 2. Parameters for experiments

Parameter	Value
α_Q	0.1
γ	0.9
τ	0.1

The computational experiments have been done with arameters as shown in Table 2. In addition, all initial Q-values are set at 5.0 as the optimistic initial values.

5.3. *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 the occurrence of a sudden absence on day, evening and late-night shift were observed during learning over 50 simulations, as described in Fig. 1 and Fig. 2, respectively. Here, a day of the absence and an absent nurse were determined by random numbers.

It can be seen from Fig. 1 and Fig. 2 that (1) in day and late-night shift absences, the target state was reached in a few steps, (2) required steps for accomplishing the task was 100 only once in evening shift absence, because the target state could not be reached, (3) the number of violations has not increased much from the initial number of 2.

In evening shift, where the target state could not be reached, no alternative nurse could be found for the 17th nurse absence on 28th day. This situation could not be adjusted without increasing the number of violations excluding over-holiday violations in the current month only, since this absence occurred late in the month. In practice, it is thought that an alternative nurse is secured in consideration of the work shift schedule for the following month.

Thus, we confirmed that the extended method can deal with absences other than designated work shift except at the end of the month in this task.

6. Conclusion

In this paper, we extend the proposed method to dynamic nurse scheduling, in which the work schedule is revised or rescheduled when an absence occurs. Through computational experiments, we confirmed that the extended method can deal with absences other than designated work shift except at the end of the month.

Our future projects include to respond to a sudden absence of a designated work shift and to clarify the rules for creating shift schedules, etc.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number JP19K04906.

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, 1998.
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.

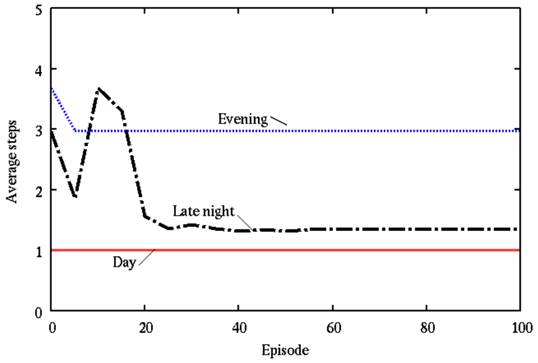


Fig. 1. Required steps for accomplishing the task.

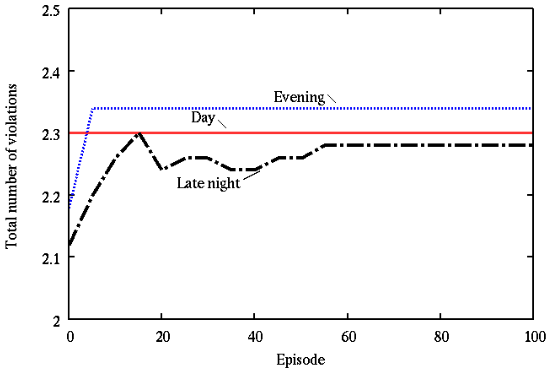


Fig. 2. Total number of violations when accomplishing the task.

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.

Classification of Time Series Data Obtained by the Satellite by Using Rule-Based and Machine-Learning Methods

Satoko Saita
National Institute of Technology, Kitakyushu College, 5-20-1 Shii, Kokuraminamiku,
Kitakyushu, Fukuoka, 802-0985 Japan

Mariko Teramoto
Faculty of Engineering, Kyushu Institute of Technology, 1-1 Sensui-Cho, Tobata,
Kitakyushu, Fukuoka, 804-8550

Kentarou. Kitamura
Faculty of Engineering, Kyushu Institute of Technology, 1-1 Sensui-Cho, Tobata,
Kitakyushu, Fukuoka, 804-8550

E-mail: saita@kct.ac.jp
researchmap.jp/saita

Abstract

Because of its communication volume constraints, scientific observations by nanosatellites need to reduce downlink data by onboard data preprocessing. Therefore, we tried to classify time series data of the geomagnetic field obtained by the SWARM satellite to determine the most appropriate method for onboard classification of a phenomenon in the geomagnetic field. The classifications have been executed using rule-based, K-means, and combined CNN methods. The experimental results demonstrated the effectiveness of the machine-learning model with LSTM networks.

Keywords: Deep-Learning, Nanosatellite, Machine-Learning, Time-series Classification

1. Introduction

It is known that electromagnetic waves occurring in geospace are affected by the near-earth plasma environment. Among those, the electromagnetic ion cyclotron (EMIC) waves are generated by the cyclotron-resonant instability of the anisotropic energetic (10 – a few 100 keV) protons [1], [2]. In the Earth's magnetosphere, the EMIC waves can be dealt with in the ULF (Ultra-low Frequency) range (0.1 - 5 Hz).

The electrons can resonate with the EMIC waves, and the energy of the electrons may go up to 50 MeV. The electrons in the MeV energy range can affect neural dynamics [3], [4].

Therefore, studies of EMIC waves are essential for a better understanding of their overall role in the interaction of the solar wind with the Earth's magnetosphere-ionosphere-atmosphere coupled system.

In general, magnetic measurements to monitor the populations of EMIC waves along their orbits have been conducted by conventional satellites with a weight of more than 100 kg. On the other hand, nanosatellites with a weight of less than 50 kg are exponentially increasing since 2013. Therefore, they are expected to be a potential candidate for affordable tool in-orbit observation for the sciences (Esper et al., 2000) [15]. The Laboratory of Lean Satellite Enterprises and In-orbit Experiments (LaSEINE), established by the Kyushu Institute of Technology, is developing these nanosatellites. One of the most expected outcomes of the development is an enhancement of geomagnetic pulsation (including EMIC waves) observations. However, the observation by nanosatellites is required to reduce the downlink data because the miniaturization of the satellites constrains their communication capabilities.

To reduce the downlink data, preliminary data analysis should be done in orbit, and the minimum required data already analyzed can be downlinked to the ground station. In this paper, we developed deep-learning and rule-based methods to detect geomagnetic pulsations. This study will provide a system to evaluate the EMIC-wave-appearance period. Furthermore, this study will contribute to creating a system that allows more downlink data of a particular period, such as specific geomagnetic pulsations.

2. Materials and Methodology

2.1. *Data Preparation*

Geomagnetic field data obtained from satellites are used for this study. Swarm, the European Space Agency's first constellation mission for Earth Observation, consists of three satellites for the Earth's magnetic field observation. The identical satellites Alpha, Bravo, and Charlie (A, B, and C) were launched on 22 November 2013 into a near-polar orbit with orbital inclinations of 87.4° for Swarm-A and -C and 88.0° for Swarm-B. In this study, we utilized high-resolution vector magnetic field data sampled at 50 Hz with an accuracy of 0.01 nT from the satellite “A” during the magnetic storm of 22 June 2015.

The fluctuating field was then expressed in the field-aligned coordinate system, in which the ***z***-axis is parallel to the background magnetic field lines (***B***), the ***y***-axis is defined as (***B*** x ***R***), where ***R*** is the radial vector from the Earth's center to the satellite location, and ***x***-axis corresponds to ***y*** × ***z***. By and Bx points toward the geomagnetic east and upward, respectively.

To determine the background magnetic field, we applied a Savitzky-Golay smoothing filter with a window

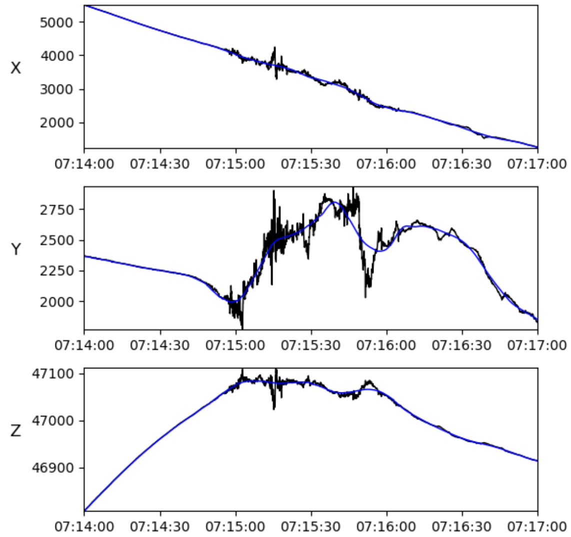


Fig. 1. Original data (black) and smoothing data (blue) of the 50-Hz magnetic field

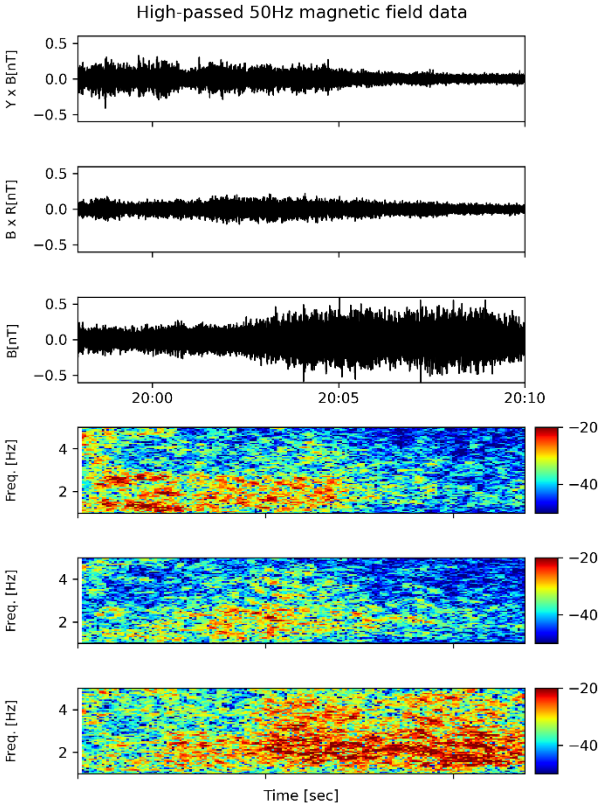


Fig. 2. High-passed 50 Hz magnetic field component in the mean-field-aligned coordinates.

of 30 x 50 data points and a second-order polynomial [5]. Above stated method was also used by Kim et al. (2018) and Wang et al. (2019) [6], [7]. Fig. 1 shows the original 50-Hz data (black) and the smoothing data (blue). An example wave spectrogram is shown in Fig. 2. The top three panels show the temporal variations of the fluctuating magnetic field in the field-aligned coordinate system as observed by Swarm A from 19:58 to 20:10 universal time (UT) on 22 June 2015. The cutoff frequency for a high-pass filter is 0.1 Hz. Examples of detected EMIC waves observed by Swarm A on 22 June 2015. The bottom three panels of Fig. 2 show the results of the spectrogram analysis. The fast Fourier transform (FFT) was performed with a window size of 800 data points (~16 s) and an overlap of 600 points (~12 s).

In previous studies, we need visual checks of FFT spectra to identify EMIC wave events because other electromagnetic activities might contaminate the fluctuating magnetic field data, including the signal of EMC waves. We can determine that EMIC wave events can be detected based on the following criteria: the signals satisfy at least 0.15 Hz bandwidth and at least 1 min duration. Then we can also exclude the other type of pulsations or background noises classified with amplitudes of bottom frequency [8].

The magnetic field data obtained from satellite “A” had been normalized, standardized, and reshaped to construct machine-learning datasets. First, these time series data are split into 6000 data points and labeled with “EMIC wave appearing (1)” or “no EMIC wave appearing (0)”. Then we split fragmented data for training, validation, and testing datasets.

2.2. *Machine Learning*

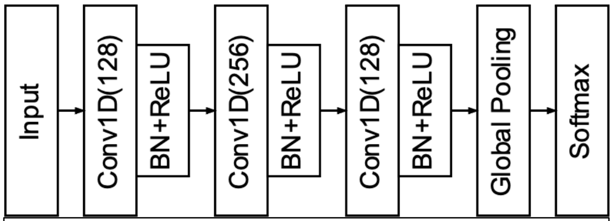


Fig. 3. The basic design of the fully convolutional network model.

This study used the fully convolutional network model to determine EMIC wave events [9], [11]. Fig. 3 shows the architecture of the fully convolutional network model. This model uses CNNs to extract image features, then transforms the number of channels into the number of classes via a convolutional layer, and finally converts the height and width of the feature map to those of the input image via the transposed convolution. As a result, the model output has the same size as the input data.

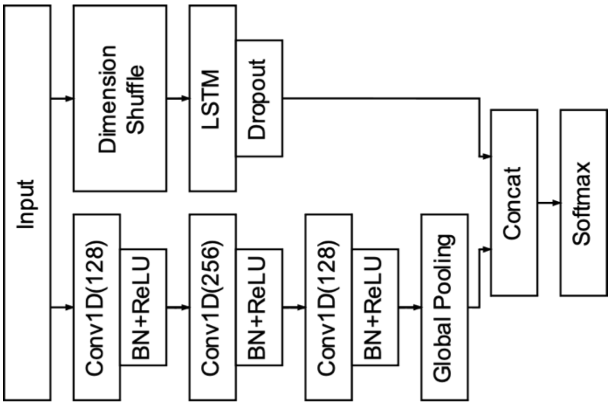


Fig. 4. Architecture of the 1d-CNN + LSTM model. In this architecture, the shortcut connections are used to pass information from the downsampling layers to the upsampling layers.

A model consisting of convolutional neural networks (CNNs) and long short-term memory networks (LSTM) was also applied to the datasets. In the model, the fully convolutional block is augmented by a Long Short Time Memory block [12]. The fully convolutional block consists of three stacked temporal convolutional blocks. Each convolutional block is identical to the convolution block in the CNN architecture. Each block consists of a temporal convolutional layer, accompanied by batch normalization followed by a ReLU (Rectified Linear Unit) activation function. Finally, global average pooling is applied following the final convolution block. Simultaneously, the time series input is conveyed into a dimension shuffle layer. The transformed time series from the dimension shuffle is passed into the LSTM block. The LSTM block comprises either a general LSTM layer or an Attention LSTM layer. Fig. 4 shows the architecture of the 1d-CNN + LSTM model. This method is a new approach for extracting EMIC waves. In this case, we proposed a method for acquiring spectrograms. We provided an example of pre-processing time series data for a group of training datasets needed in this case.

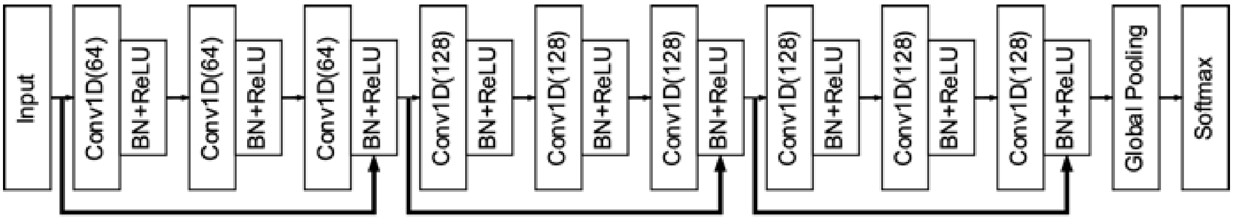


Fig. 5. Architecture of the Residual Network

We also applied the Residual Network model to determine EMIC wave events. ResNet was developed by He et al. 2016 [13]. Fig. 5 shows the architecture of the Residual Network. The advantage of the ResNets model compared to other architectural models is that the performance of this model does not decrease even though the architecture is getting deeper. The ResNet model is implemented by skipping connections on two to three layers containing ReLU and batch normalization among the architectures. He et al. showed that the ResNet model performs better in image classification than other models, indicating that the image features were extracted well by ResNet.

2.3. *Data Visualization*

We used Class Activation Maps (CAMs) to visualize where the model paid attention. CAMs enable us to see not only the class the network predicts but also the part of the image the network is especially interested in (Zhou et. al., 2015). The visualization provides a powerful way to communicate data-driven findings, motivate analyses and detect flaws. Fig. 6 shows an example of applying CAMs to the magnetic field data. The left figure shows the result of the Class Activation Mapping. We can see that the CNN network pays attention to about a frequency range, which is consistence with the eigenfrequencies of geomagnetic pulsations. The data analysis and machine learning were performed in accordance with Nielsen (2019) [14].

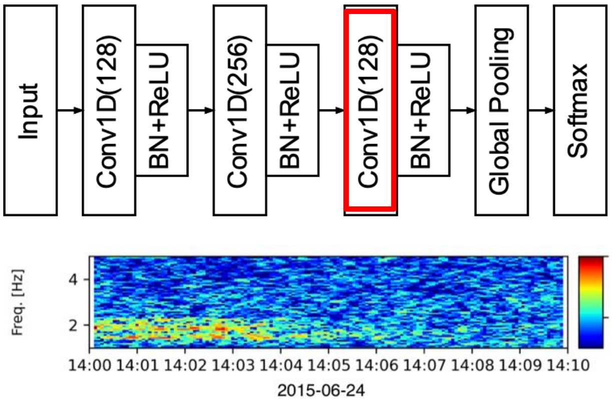


Fig. 6. The example of the results of Class Activation Maps (CAMs).

3. Results

In this study, we applied machine learning to the time series data obtained from SWARM satellites. This study aimed to develop a satellite system to discriminate specific natural phenomena in a future mission. The observation data of the SWARM satellites were assumed as magnetic field data observed by a CubeSat, and the model was trained. The SWARM satellites are scientific satellites for measuring the earth's magnetic field and were launched in 2013 by ESA [10]. The magnetic field data observed by SWARM are available to the public and can be acquired via the Internet.

Table 1. Comparison of accuracy in several testing datasets for machine learning models.

Kind of data	1DCNN-LSTM	FCN	ResNet
line plot	72.6	72.3	73.4
spectrogram	79.6	79.6	80.2

We use 2 kinds of input data and applied 3 machine learning models to identify EMIC wave events. ResNet applied to both kind of data shows the highest accuracy. The results using spectrograms show good performance on all of the machine learning models (Table 1).

4. Conclusions

We got the results by using the data separated every 600 seconds. However, these results also include possibilities for improvement. the correct answer rate is expected to be improved when the data of 300 seconds was prepared in the 60-second unit and the learning data was increased by 10 times. And this study addresses only a few EMIC wave events. We must continue to accumulate EMIC wave events and build a rich data set. In the forthcoming works, we will propose incorporating added architectures with their variants using the corresponding datasets.

Acknowledgements

This work was done in the framework of joint studies between the Kyushu Institute of Technology and KOSEN.

References

1. Cornwall JM (1965) Cyclotron instabilities and electromagnetic emission in the ultra low frequency and very low frequency ranges. J Geophys Res 70(1):61–69
2. Kennel CF, Petschek H (1966) Limit on stably trapped particle fluxes. J Geophys Res 71(1):1–28
3. Meredith NP, Thorne RM, Horne RB, Summers D, Fraser BJ, Anderson RR (2003), Statistical analysis of relativistic electron energies for cyclotron resonance with emic waves observed on crres. J Geophys Res 108:A6
4. Tsurutani B, Hajra R, Tanimori T, Takada A, Remya B, Mannucci A, Lakhina G, Kozyra, J, Shiokawa K, Lee L et al (2016) Heliospheric plasma sheet (hps) impingement onto the magnetosphere as a cause of relativistic electron dropouts (reds) via coherent emic
5. Savitzky, A. and Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. Analytical Chemistry, 36(8), 1627–1639. https://doi.org/10.1021/ac60214a047
6. Wang, H., He, Y. F., Lühr, H., Kistler, L., Saikin, A., Lund, E., & Ma, S. (2019). Storm time EMIC waves observed by Swarm and Van Allen Probe satellites. Journal of Geophysical Research: Space Physics, 124, 293– 312.
7. Kim, H., Hwang, J., Park, J., Bortnik, J., & Lee, J. (2018). Global characteristics of electromagnetic ion cyclotron waves deduced from Swarm satellites. Journal of Geophysical Research: Space Physics, 123, 1325– 1336.
8. Erlandson, R. E., & Anderson, B. J. (1996). Pc 1 waves in the ionosphere: A statistical study. Journal of Geophysical Research, 101(A4), 7843–7857.
9. Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., and Torralba, A., “Learning Deep Features for Discriminative Localization”, *arXiv e-prints*, 2015.
10. Friis-Christensen, E., H. Lühr, D. Knudsen, R. Haagmans (2007), Swarm – An Earth Observation Mission investigating Geospace, Advances in Space Research, Volume 41, Issue 1, 2008, Pages 210-216, ISSN 0273-1177.
11. Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3431–3440).
12. Karim MA, et al. (2018). Distinct features of multivesicular body-lysosome fusion revealed by a new cell-free content-mixing assay. Traffic 19(2):138-149.
13. He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770-778.
14. Nielsen, A.(2019), Practical Time Series Analysis: Prediction with Statistics and Machine Learning, O'Reilly Media, ISBN: 9781492041627.
15. Esper J, P. V. Panetta, M. Ryschkewitsch, W. Wiscombe, and S. Neech, NASA-GSFC nano-satellite technology for Earth Science Missions, Acta Astronautica, 46, 2-6, 287-296, 2000

Authors Introduction

Dr. Satoko Saita

She is an Associate Professor of the Department of Creative Engineering at the National Institute of Technology, Kitakyushu College in Japan. She graduated from the Department of Earth and Planetary Sciences, Kyushu University, in 2000. She received his Ph.D. degree in Science from Kyushu University in 2005. Her research interest is Space and Upper Atmospheric Sciences.

Dr. Mariko Teramoto

She is an Associate Professor at the Faculty of Engineering at Kyushu Institute of Technology in Japan.

She received the B.S., M.S., and Ph.D. degrees from the Department of Geophysics, Kyoto University, Kyoto, Japan, in 2005, 2007, and 2010 respectively. Her research interest is geomagnetic variations and dynamics of plasmas in geospace.

Dr. Kentaro Kitamura

He is a Professor at the Faculty of Engineering at Kyushu Institute of Technology in Japan. He graduated from the Department of Earth and Planetary Sciences, Kyushu University, in 1997. He received his Ph.D. degree in Science from Kyushu University in 2001. His research interest is Space Weather and nano-satellite.

Microalgae Detection by Digital Image Processing and Artificial Intelligence

Watcharin Tangsuksant[1], **Pornthep Sarakon**[2]
Department of Production and Robotics Engineering,
Faculty of Engineering, King Mongkut's University of Technology North Bangkok,
1581 Pracharat 1 Road, Wongsawang, Bangsue, Bangkok, 10800, THAILAND
watcharin.t@eng.kmutnb.ac.th[1], pornthep.s@eng.kmutnb.ac.th[2]

Abstract

This article presents a technical approach to the video computer analysis, to automatically identifying the two most frequently identified microalgae in water supplies. To handle some difficulties encountered in image segmentation problem such as unclear algae boundary and noisy background, we proposed a deep learning-based method for classifiers or localizers to perform microalgae detection and counting process. The system achieves approximately 91% accuracy on Melosira and Oscillatoria detection, which around 4.82 seconds per grid. (Intel Xeon(R) CPU E5-2667 12 CPU at 2.66GHz and 32.0GB RAM, NVIDIA Quadro K5200 with 2304 CUDA cores). The system can significantly reduce 33.33 - 55.56% of the counting time when compared with the visual inspection of manual methods, and eliminate the error due to the human fatigue.

Keywords: Microalgae detection, Melosira, Oscillatoria, YOLO

1. Introduction

Nowadays, the quality of water from natural sources is de-creasing by human activities, agriculture, and industrial consumption. Piped water requires to be clean because people use this water for their consumption. Therefore, the Metropolitan Waterworks Authority monitors the water quality to detect possible pollutants and eutrophication phenomena which affect the production and quality control of piped water. Water supply production needs to screen all the algae with the conventional water filtration system. The filtration system is reducing its efficiency by approximately 30 percent [1], due to the algae clogging the sand filter. For this reason, the workers require to manage many algae and clean the filter more often.

The Metropolitan Waterworks Authority, in Thailand, has several departments responsible for inspecting and monitoring water sources and water supply canals. Furthermore, they collect water samples for quality analysis regularly over the year. Moreover, they also request cooperation from other organizations, such as the Department of Health, the Pollution Control Department, the Harbor Department, the Department of Agriculture Provincial, and the Public Health Office. To ensure, raw water used to produce the water supply has high quality and appropriate use for produce water supply with ISO 9002 quality standard of the Metropolitan Water-works Authority [2].

The microalgae problem in the water supply system is the main problem that is focused on this research. Microalgae are unicellular organisms that have various shapes, sizes, and structures. Classifying these microalgae manually may require experts and different experts have a different standard. To classify the microalgae, it has several steps for pre-paring the test slide and needs to use a counter-press to count a multiple of them according to the type of microalgae. The human visual system is fast and accurate, allowing us to perform complex tasks with little conscious thought. As being human, eye fatigue can make a mistake, but the computer does not. Therefore, it has various difficulties in the operation. It takes approximately four hours each day to finish the classification and counting process that means the Metropolitan Waterworks Authority spent a lot of time and cost on this work. Therefore, it will be better if they have an application that can detect and counting the microalgae directly.

There have been many researches proposing the algae detection and classification. For example, research proposed the technique which extracted features and segmentation in the algae images such as distance of contour points from centroid, number of edges, and the number of square pixels inside the edge [3]. Another work was presented by Sansoen Promdaen and et.al [4], 12 microalgae that most commonly found in water source of Thailand was classified by using blurry texture object with Sequential Minimal Optimization (SMO). Their result is pretty high with 97% of accuracy. Moreover, some researches aimed to classify the plankton image that casually share some characteristics with microalgae. However, the planktons were extracted the larger features [5]–[7] for satisfied result for classification. Currently, the machine learning become the useful tool for the object classification and recognition in the image processing. Many works applied the Support Vector Machine (SVM) for their plankton or microalgae classification [8]–[9]. For instance, Lili Xu and et.al. used the supervised learning with SVM classifier [9]. This approach obtains the information through a density method that could be sensitive to the microalgae size. Thus, the result of this proposed was still unfeasible. In addition, Paulo Drews-Jr and et.al [10]. pro-posed the microalgae classification using the semi-supervised and active learning based on Gaussian mixture models. Especially, deep learning is the popular and powerful for image processing which can apply for microalgae classification. Lago Correa and et.al [11]–[12] solved the problem using deep convolution neural network (CNN) for microalgae classification with their collected datasets. Although, these previous researches proposed the various techniques for algae or plankton classification, they could not apply counting microalgae practically. There was a re-search of microscopic algae detection by segmentation with Mask-RCNN, especially of diatoms [13]. This research had a drawback is that the performance of the detection step limits the performance of the segmentation.

With the above limitations, it is a drawback to keep up with work in conditions that algae grow well in the nature sources. In addition, the number of samples and analysis frequency could not be increased. This research aims to solve this problem with a deep learning technique called YOLO (You Only Look Once). It is a single neural network that predicts bounding boxes and class probabilities directly from full images in one evaluation [14]. We also created a classified dataset for training and making data weights to apply with the Window Form Application that we write for automated image processing the type of microalgae and evaluate the program accuracy in identifying algae type against the expert analysis.

2. Materials and Methods

2.1. *Microalgae*

Algae is a low-class organism with a simple structure that important in balancing nature. It is commonly found in nature, mostly in humid areas. Algae can be growth using water source from municipal, some industrial runoff and rainwater. According to Bold (1985) [15] dividing algae into 9 divisions. It can be categorized by characteristics of the cell. It could be single-celled algae (Unicellular) or multicellular algae (Multicellular). Algae is a simple plant that can range from the microalgae to macroalgae. Most microalgae growth requires light, carbon dioxide, water and a few nutrients [16].

Microalgae are unicellular organisms of various shapes, sizes, and structures. Only the most common and problematic microalgae were selected for data collection in this research. The expert at the Metropolitan Waterworks Authority provided information and labelled data on the two most common algae while we collected data, which are Melosira and Oscillatoria as shown in Fig. 1.

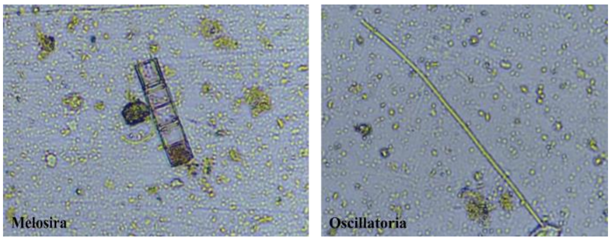


Fig. 1 Sample of Melosira and Oscillatoria.

Melosira belongs to the Division Chrysophyta, the genus Bacillariophyceae (Diatom), characterized by a round, oval, or cylindrical cell. The melosira arranged in a straight line like a chain. There are unequal dots lined up around the edge of the cell. Melosira is distinguished from Aulacoseira and other fresh-water diatoms with similar colonial growth habits by uniformly structured walls, without costae or septae, and lack of spines visible under the light microscope. The most common freshwater species that can be found in both sea and fresh water. It is generally benthic in growth habit, but it is also commonly entrained into the plankton. Some authorities consider it as indicator of organic pollution [17].

Oscillatoria is a genus of blue-green algae common in freshwater environments, including hot springs. These unbranched filamentous algae, occurring singly or in tangled mats, derives its name from its slow, rhythmic oscillating motion, which is thought to result from a secretion of mucilage that pushes the filament away from the direction of excretion [18]. Blue-green algae poisoning occurs when the algae form a scum on top of ponds or other stagnant waters. That is why this problem should be a serious concern [19].

2.2. *Methods*

This research proposes the microalgae detection on the test slide. Two main types of microalgae consist of Melosira and Oscillatoria which are very difficult to distinguish without expert skill. Classification and localization of these microalgae are extremely challenge because of complex step of preparation for microalgae sample and unclear objects on the various background in the images. This article overcomes the mentioned challenge by deep learning-based object detection technique. Therefore, this section explains the detection technique with YOLOv3, preparation of sample and system setup.

YOLO (You Only Look Once) is a real-time object detection algorithm based on the convolutional neural networks which is originally developed by Joseph Redmon et al., 2016 [14]. In additional, YOLO can detect multiple objects on a single image which the class prediction and location identification. This article applies third version of YOLO, which is faster and better than original version.

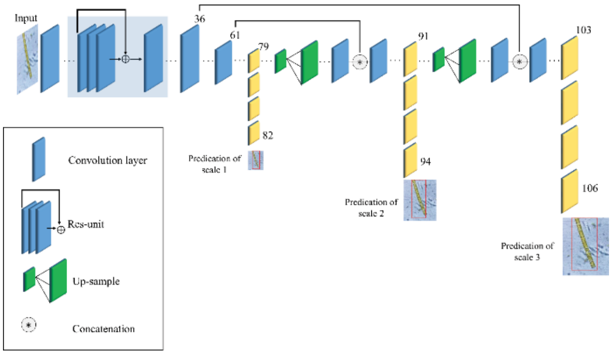


Figure 2 YOLOv3 network architecture.

The YOLOv3 network architecture is shown in the Fig. 2. This version is design with a deeper architecture of feature extraction namely Darknet-53 [14]. According to its name, there are 53 convolutional layers for this network which can detect objects on multi-scale feature maps as 79, 91 and 103 layers. Moreover, each layer followed by a batch normalization layer and Leaky ReLU activation function.

Table 1 Configurations for training model using YOLOv3.

Parameters	Configuration
Size of input image	2048 px. × 1152 px.
Number of training images	794 images for Melosira, 1159 images for Oscillatoria
Learning rate	0.001
Decay of Learning rate	0.0005
Number of epochs	4000
Batch size	64

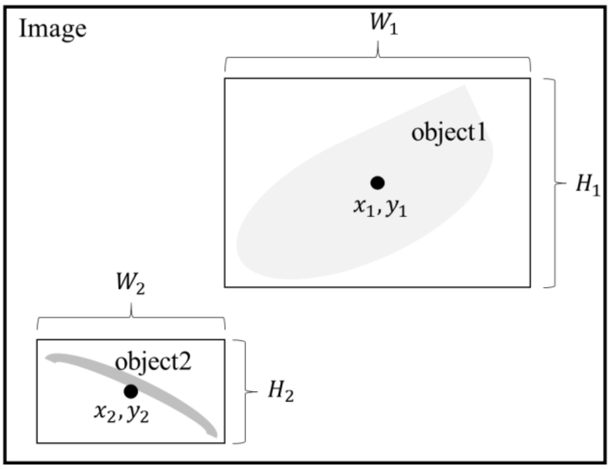


Figure 3 Bounding box for training the YOLO.

For the training process of YOLOv3, the bounding boxes of the whole object targets with their labels are required. Fig. 3 shows the bounding box and its essential parameters for training the YOLOv3. There are totally five parameters for each object including x-coordinate of center in a bounding box (x), y-coordinate of the center in a bounding box (y), width of the bounding box(W), height of the bounding box (H) and an object label, which can create the training matrix of these parameter as Eq. 1. Furthermore, training configurations are described in Table 1. This article focuses on two types of microalgae, Melosira and Oscillatoria. In order to obtain the correct microalgae images, these training images are labelled by the expert from Metropolitan Waterworks Authority.

$$Y = \begin{bmatrix} object1 & x_1 & y_1 & W_1 & H_1 \\ object2 & x_2 & y_2 & W_2 & H_2 \\ & & \vdots & & \\ objectn & x_n & y_n & W_n & H_n \end{bmatrix} \qquad (1)$$

The output of YOLO will provide the five values involving the object location in the image as same as the values of training parameters. Moreover, the confidence values will return the probability (0–1) of detected object for each bounding box which zero means the lowest confidence of object prediction and one are the highest confidence of prediction.

2.3. *Preparation of Microalgae Samples*

Samples of microalgae were obtained from raw water supplies. Raw water with 50 mL., from the raw water source, were poured to the test tube. To separate the components of liquid in the test tube, a centrifuge was used for this process. The centrifuge was set as 2500 round per minute, with 20 minute of process time. The obtained liquid in the test tube would be separated into two parts, which are liquid and pel-let (microalgae) for the top and bottom parts respectively. Then, upper part of liquid was drawn off 40 mL. using pipette. The remained liquid with 10 mL., in the test tube, was brought to the Vortex mixer machine for mixing the sample of 10 second. To prepare the samples of microalgae, finally, dropping the sample from the test tube to the microscope slide. The Fig. 4 is shown the overview process of these preparation.

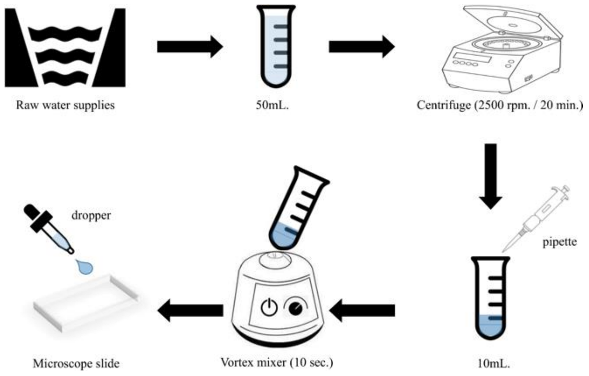


Figure 4 Steps of Microalgae samples preparation.

2.4. *Apparatus and system setup*

In order to collect the digital images of microalgae, apparatus selection and system setup are the crucial process. There are three main components consisting of light micro-scope, digital microscope and computer.

The proposed system uses the microscope which is set the 10× magnification for objective lens, because it is fit enough to observe the microalgae. Next component is the digital microscope, which is the important part for collecting the digital images from microscope. This article uses Eakins Digital Microscope which specification of this device is shown in Table 2.

Table 2 Specification of Eakins Digital Microscope.

Specifications	Configuration
Model	21MP 30MP Microscope camera
Picture pixel	38MP / 30MP / 21MP
Video size	30fps for 4K / 60fps for 1080P
CMOS Sensor	1/2.33 inch
Lens Interface	Industrial 200X, 500X, 100X

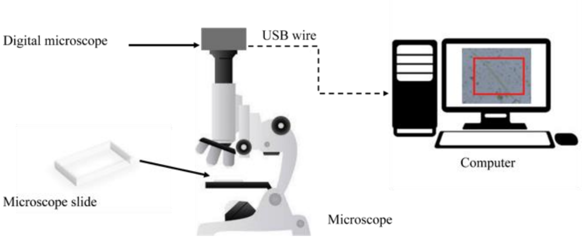
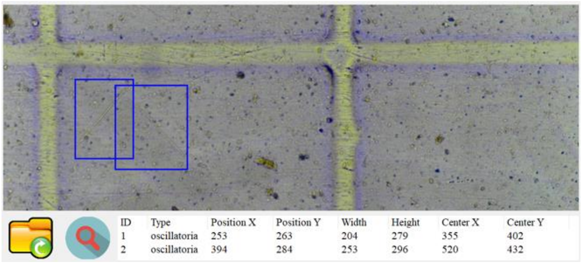


Figure 5 Image acquisition system.

The digital microscope was attached to the C-mount lens where was on the top of the microscope as shown in the Fig 5. The images with the size of 8,320 px. × 4,680 px. were acquired by saving on the micro SD card, which was built-in on the Eakins Digital Microscope. Because a large number of images were trained by using YOLOv3, the GPU on a computer was required. The specification of the computer was Intel Xeon(R) CPU E5-2667 12 CPU at 2.66GHz and 32.0GB RAM, NVIDIA Quadro K5200 with 2304 CUDA cores.

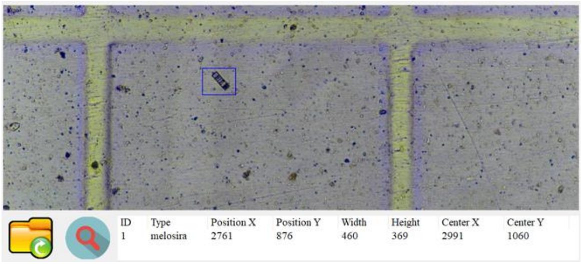
3. Experiments and Results

After training the model of YOLOv3 as explaining in previous section, the performance of detection will be measured. This article provides and develops the user-interface on Visual Studio C# that can read the images and detect the micro-algae using YOLOv3 as shown in Fig. 6–7. The tested images show the various sizes of microalgae and noisy background. The experiment tests the 23 unknown images which contains the ground truth of Melosira and Oscillatoria as 9 and 14 images, respectively. Table 3. shows the confusion matrix of microalgae detection, which can calculate necessary performance as shown in equation 2 to 8. The proposed method shows the high accuracy of detection by 0.91, and high performance of F1 score both of Melosira and Oscillatoria.



ID	Type	Position X	Position Y	Width	Height	Center X	Center Y
1	oscillatoria	253	263	204	279	355	402
2	oscillatoria	394	284	253	296	520	432

Figure 6 Detection of Oscillatoria.



ID	Type	Position X	Position Y	Width	Height	Center X	Center Y
1	melosira	2761	876	460	369	2991	1060

Figure 7 Detection of Melosira.

Table 3 Confusion matrix of detection.

Actual Values Predicted Values	Melosira	Oscillatoria
Melosira	9	0
Oscillatoria	2	12

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} = 0.91 \tag{2}$$

$$Precision(Melosira) = \frac{TP}{TP + FP} = 1.00 \tag{3}$$

$$Precision(Oscillatoria) = \frac{TN}{FN + TN} = 0.85 \tag{4}$$

$$Recall(Melosira) = \frac{TP}{TP + FN} = 0.82 \tag{5}$$

$$Recall(Oscillatoria) = \frac{TN}{FP + TN} = 1.00 \tag{6}$$

$$F1(Melosira) = \frac{2 \times Precision \times Recall}{Precision + Recall} = 0.90 \tag{7}$$

$$F1(Oscillatoria) = \frac{2 \times Precision \times Recall}{Precision + Recall} = 0.92 \tag{8}$$

Generally, a test slide has a thousand of grid. The expert usually takes 2-3 hours per slide for counting the microalgae. For this proposed technique, the average processing time of detection is 4.82 seconds per image (or per grid). Therefore, it will take around 1 hour and 20 minutes for counting microalgae per a test slide, which is 1.50-2.25% faster than visual inspection of an expert. The proposed system can significantly reduce 33.33-55.56% of the counting time, and eliminate the error due to the human fatigue.

4. Conclusion

This article has shown the feasibility of microalgae detection with the noisy background and various size of microalgae. Especially, Melosira and Oscillatoria are the most frequently found microalgae in water supplies that impact to the quality assessment of the water source. According to the experiment, the system can detect and count the Melosira and Oscillatoria by 91.00% accuracy and 4.82 second per grid using the YOLOv3. The system can reduce 33.33-55.56% of the counting time of microalgae compared to human counting. However, the limitations and problems for some images are still faced such as the missed detection in case of scratch on the microscope slide and lacerated microalgae.

Future work is to develop the automatic microalgae counting system with moving the stage of the microscope. Moreover, other microalgae types will be detected for this automatic counting system.

Acknowledgements

This work is supported by Metropolitan Waterworks Authority – MWA and Associate Professor Doctor Ramil Kesvarakul from King Mongkut's University of Technology North Bangkok.

References

1. Provincial Waterworks Authority with the problem of algae and drought [Online]. 2014, Website: https://www.pwa.co.th/news/view/30398.
2. Quality of raw water for provincial waterworks authority in Thailand. 2005, Website: https://www.mwa.co.th/ewt_news.php?nid=428&filename=Event.
3. P. Coltelli and et al., "Water monitoring: automated and real time identification and classification of algae using digital microscopy", Environmental Science: Processes & Impacts vol. 16, no. 11 ,pp. 2656-2665, 2014.
4. S. Promdaen, P. Wattuya and N. Sanevas, "Automated microalgae image classification", Procedia Computer Science, vol. 29, pp. 1981-1992, 2014.
5. M.A. Mosleh and et al., "A preliminary study on automated freshwater algae recognition and classification

system”, BMC bioinformatics, vol. 13. no. 17, BioMed Central, 2012.

6. M.C. Benfield and et al., “RAPID: research on automated plankton identification”, Oceanography, vol. 20, no.2, pp. 172-187, 2007
7. A. Verikas and et al., “An integrated approach to analysis of phytoplankton images” IEEE Journal of Oceanic Engineering, vol. 40, no.2, pp. 315-326, 2014.
8. Q. Hu and . Davis, “Accurate automatic quantification of taxa-specific plankton abundance using dual classification with correction”, Marine Ecology Progress Series, vol. 306, pp. 51-61, 2006.
9. L. Xu, and et al., “Red tide algae classification using svm-snp and semi-supervised fcm.”, 2010 2nd International Conference on Education Technology and Computer, vol. 1. IEEE, 2010.
10. R.G. Colares snf et al., “Microalgae classification using semi-supervised and active learning based on Gaussian mixture models”, Journal of the Brazilian Computer Society, vol. 19, no.4, pp 411-422, 2013.
11. I. Correa and et al., “Deep learning for microalgae classification”, 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA). IEEE, pp. 20-25, 2017.
12. S. Lakshmi, and R. Sivakumar, “Chlorella algae image analysis using artificial neural network and deep learning”, Biologically Rationalized Computing Techniques For Image Processing Applications, Springer, Cham, 2018, pp. 215-248.
13. J. Ruiz Santaquiteria, and et al., “Semantic versus instance seg-mentation in microscopic algae detection”, Engineering Applications of Artificial Intelligence, vol. 87, pp. 103271, 2020.
14. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection”, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2016-Decem, pp. 779–788, 2016, doi: 10.1109/CVPR.2016.91.
15. H. C. Bold and M. J. Wynne, Introduction to the algae: structure and reproduction, 1978.
16. S. Zullaikah, A. T. Utomo, M. Yasmin, L. K. Ong, and Y. H. Ju, “Eco-fuel conversion technology of inedible lipid feedstocks to renew-able fuel”, Adv. Eco-Fuels a Sustain. Environ., pp. 237–276, 2019, doi: 10.1016/b978-0-08-102728-8.00009-7.
17. Kociolek, J. Patrick, et al. "Centric and araphid diatoms." Freshwater Algae of North America. Academic Press, 653-708, 2015.
18. T. K. Bhutia, K. Rogers, and G. Shukla, Oscillatoria, Britannica, T. Editors of Encyclopedia. 2016, [Online]. Available: https://www.britannica.com/science/Oscillatoria.
19. W. K. Rumbeiha and D. B. Snider, “Veterinary Toxicology, in Encyclopedia of Toxicology: Third Edition”, Elsevier, 2014, pp. 915–928.

Authors Introduction

Dr. Watcharin Tangsuksant

He received B.Eng. degree in Biomedical Engineering from Srinakharinwirot University, Bangkok, Thailand in 2013. His M.Eng. degree in Biomedical Engineering from King Mongkut’s institute of technology Ladkrabang, Bangkok, Thailand in 2015. His D.Eng. degree in Department of life science and system engineering from Kyushu institute of technology, Wakamatsu campus, Japan, in 2019. He is currently a lecturer in Department of production and robotics engineering at King Mongkut’s University of Technology North Bangkok. His research interest is imageprocessing and machine learning.

Dr. Pornthep Sarakon

He is a lecturer in the Department of Production and Robotics Engineering, King Mongkut's University of Technology North Bangkok, Thailand. He received a B.Eng. (Hons.) degree in Biomedical Engineering from Srinakharinwirot University, Nakhon Nayok, Thailand in 2015 and an M.Eng. degree in Information and Communication Technology for Embedded Systems from the Sirindhorn Inter-national Institute of Technology, Thammasat University, Thailand in 2017. He received a Ph.D. degree in Electrical and Electronic Engineering from Kyushu Institute of Technology (Kyutech), in 2021. His research interests include artificial intelligence, model compression, computer vision, and 3D body data.

Smartcropplanting: IOT-Based Mobile Application for Hydroponic System

Sung Jun Kyu
Institute of Computer Science & Digital Innovation, UCSI University,
UCSI Heights, 1 Jalan Puncak Menara Gading, Kuala Lumpur, WP Kuala Lumpur 56000, Malaysia

Chit Su Mon
School of Mathematical and Computer Sciences, Heriot-Watt University, Putrajaya 62200, 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
E-mail: jkyu.sung99@gmail.com, c.mon@hw.ac.uk, kasthurisuba@ucsiuniversity.edu.my
www.ucsiuniversity.edu.my, www.hw.ac.uk/malaysia/
,

Abstract

Hydroponic crop production is challenging for farmers and gardeners; they must continuously monitor and control the crop environment to achieve or maintain the best plant growth. Climate changes and disease/pests might occur anytime at anywhere which will consequence in crop-damaging. Studies show that slight rises in temperature from 1°C to 4°C can resulting in a decrease in the production of 10 to 41%. Another study shows that depending on the severity of leaf spot disease it will damage the crop by 10 to 50%. Therefore, this study proposed to improve the current hydroponic system by implementing the Internet of Things and Image Processing technique to optimize and reduce the infeasible tasks with mobile applications. This study has conducted a mixed-mode method with home gardeners, farmers, and related experts. The results show that 80% of respondents agree that SmartCropPlanting will improve the productivity and efficiency of crop planting with IoT and Image Processing techniques.

Keywords: Agriculture, Hydroponic, IoT, Automation, Machine learning, Image Processing.

1. Introduction

Plants are not always easy to grow, farmers/gardeners might lose their crop due to climate changes, plant disease/pest, improper watering, etc. In Malaysia, climate changes have highly impacted negatively on agriculture production. One study shows that the temperature rises from 1°C to 4°C, the production of oil palm can decrease up to 10 to 41% [1]. Thus, this scenario not only effecting oil palm production it also affects many other plant productions. Hence, reduce in-plant production will be affecting the price in the market which can address food security issues in Malaysia [1]. Therefore, improving the agriculture system is an extremely important challenge for food security in Malaysia. Currently, there is a new farming technique where farmers/gardeners can grow their plants indoors and without the presence of soil [2]. This modern farming method growing plants with only nutrient-rich water instead of getting water and nutrients from the soil, Hydroponics planting system has arisen. With the ability to plant crops indoors, they will not be affected by any climate changes outdoor [3]. On the other hand, measuring and controlling hydroponic crops is quite challenging and an infeasible task for farmers/gardeners. Therefore, with the help of modern Internet of Things (IoT) technology, the challenges faced by farmers/gardeners can be solved with low-cost sensors and a few other low-cost electrical components [4][5][6]. With the help of IoT technology, it allows farmers/gardeners to monitor and control their crop environment parameters remotely by providing a mobile application with a user-friendly graphical user interface (GUI) [3] without the need for laboratory instruments [7][8][9].

Besides climate changes, plant disease and pests could affect crop production. Plant disease and pests are resulting in endanger to agriculture around the world [10]. One study shows that depending on the severity of the leaf spot disease it will damage the sugar yield by 10% to 50% [11]. Thus, this study shows that the effect on plant disease and pests will cause serious damage to the crop. The disease can cause by 2 factors, first are biotic which are fungus, bacteria, virus, and nematodes. The second is abiotic which is caused by environmental parameters like temperature, humidity, nutrient deficiency [12]. Disease on leaf spot initially emerges in small spots, then slowly evolves, and eventually spread to the whole leaf [11] and causes permanent damage to the leaf. Therefore, plant disease/pest should detect instantly to prevent further spread or growth of the disease/pest. To reduce the damage of crops, applying image processing techniques will be extremely helpful for the early detection of diseases and pests. Currently, farmers/gardeners check their crop disease with their naked eyes which might cause human error and time consuming [11]. To solve this issue, it is believed that the image processing technique will reduce infeasibility, reducing human error, and save time to identify the disease [11].

2. Literature Review

Hydroponic planting system became one of the modern planting systems without the presence of soil [2]. Farmers/gardeners need to measure the hydroponic environment parameters by using an additional instrument [3]. Therefore, researchers and developers try to carry out a solution to control and monitor Hydroponic planting systems with IoT technology without the need for additional instruments. There are other similar IoT based Hydroponic systems such as "IoT based hydroponics system using Deep Neural Network (IHDNN)" [2], "IoT Based automated Indoor Vertical Hydroponics Farming Test-Bed (IAIVHF)" [3], and "Automated smart hydroponics system using IoT (ASHSI)" [13]. In this review, it will provide a features comparison table of 3 different IoT-based hydroponic system on Table 1.

Table 1. Related work comparison

Features	**IHDNN**	**IAIV HF**	**ASHSI**
Web application	/	/	X
Mobile application	/	X	/
Notification	/	X	X
Monitor Crop Parameter	X	X	X
Control crop Parameter	X	X	/
Automatic Control Parameter Action	X	/	X
Plant disease/pest detection	/	/	/

Table 1 shows 3 different types of IoT-based hydroponic systems. As shown above, all 3 types of IoT-based hydroponic systems did not provide plant disease/pest detection. However, it does provide monitoring and controlling features. Thus, farmers/gardeners still need to physically check their plant's health.

Plant diseases/pests have been a significant concern in farmer's organizations since they will cause the reduced quality and production of the crop. Thus, early detection is important for preventing further crop damage. To detect plant diseases/pests, image processing technology can be applied for plant disease detection and classification [14]. Kusumo, B. S., et al. has tested different type of various image processing features such as [15]

1. Red Green Blue (RGB)
2. Scale-invariant feature transform (SIFT)
3. Speeded up robust features (SURF)
4. Oriented FAST and rotated BRIEF (ORB)
5. Histogram of oriented gradients (HOG)

Different type of machine learning algorithm has been tested such as Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF), and Naïve Bayes (NB) [12]. To elaborate the comparison result, the statistic of comparison is shown below.

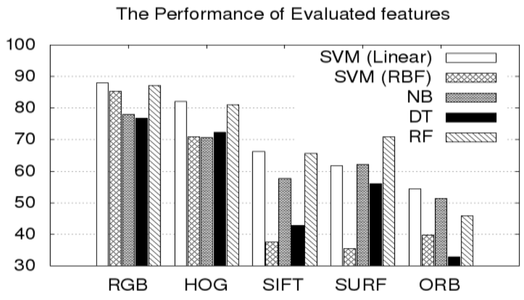


Fig. 1. Statistic of Comparison Result [6]

Fig. 1 shows that the best classifier type is RGB which has the best accuracy compared to other classifiers. Therefore, by using a free online plant disease dataset with the plant name and disease type [16] train it with RGB classifier with SVM machine learning algorithm, believed that the accuracy result will be accurate to detect plant disease/pests.

2.1. *Problem Statement, Question & Objective*

This study aims to develop an IoT- based mobile application with low-cost sensors and hardware that could help farmers/gardeners to control and monitor their crops without additional equipment. Furthermore, it also provides an image processing technique to detect plant disease/pests. This system could give more convenience to farmers/gardeners and maintain a healthy crop.

RO1: To develop a system with low-cost sensors and hardware with IoT-based Mobile application Hydroponic System. This system only requires a low-cost sensor and hardware to monitor and control hydroponic crops.

RO2: To cut the additional equipment for testing the parameter of their crops with IoT-based Mobile application Hydroponic System. With the help of this system, farmers/gardeners can monitor environment parameters without additional equipment.

RO3: To develop a fully automated system that allows the farmer to control environmental parameters and monitor their plant health remotely with an IoT-Based Mobile application Hydroponic System. The system allows farmers/gardeners to control and monitor their crops remotely and send an instant alert message to Farmers/gardeners via a mobile application if the system detected plant disease/pests.

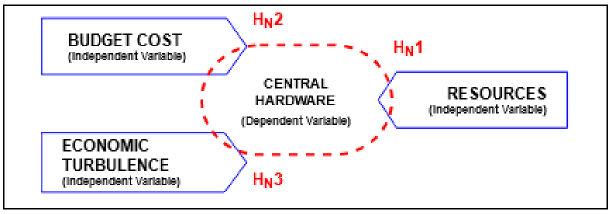


Fig. 2. Hypothesis Research Model

From the objectives derived, these are the summarized hypothesis shown in Fig. 2.

H1: Farmers/gardeners do not require expensive hardware or additional maintenance with IoT-based Mobile application Hydroponic systems.

Constant monitoring of the environment is an extremely important task to maintain a healthy crop. Depends on a different type of plant, crop environment parameter is important to grow a healthy plant. Farmers/gardeners have to physically measure parameters with overpricing and not reusable equipment. Therefore, providing an automation system with IoT-based technology will prevent farmers/gardeners to spend additional costs on equipment.

H2: Farmers/gardeners allow to control and monitor real-time data of their hydroponic crop parameter and their plant health over the internet with IoT-based Mobile application Hydroponic System.

Physically monitoring and controlling crops is a time-consuming and inefficient task. By providing an IoT-based technology, farmers/gardeners allow to control and monitor their crops remotely. Furthermore, this system also provides an image processing technique to detect plant disease/pest and send an alert message to a mobile application which is a very efficient way for detecting plant disease/pest.

H3: Farmer's organization will contribute to an increase in food demand harvest and reduce the cause of plant disease with IoT-based Mobile application Hydroponic System.

By providing this system to farmers' organizations, the food harvest demand will increase which resulting in an improvement in food security. It also provides instant alert notification to the farmer if there is any plant disease/pest detection. Thus, it will improve the early action for the farmer to stop further spreading of plant disease/pests and protect their crop.

3. Methodology

The research methodology for this research is using mixed mode.

The quantitative research method is research that asks narrow questions and collects quantifiable data from existing or potential customers by sending out an online survey or questionnaires to ask for their opinion. After collecting this information, it will be used by the researcher to deploy the mathematical frameworks and theories. We are conducting this method by using an online questionnaire. We share the survey with the related field people such as IT-field, gardener, farmers, and researchers related in this field so that we can maintain the accuracy of the result.

The quantitative research method is the most widely used when conducting market research. The qualitative research method is more focused on collecting data through conversational communication. This method not only asks about what they think but also asking why they think so. It is more like psychology, sociology, and anthropology. Therefore, it is a method that allows getting a further explanation from respondents. By understanding how respondents make the decision, it will gain advantages while concluding the research. We are conducting one-to-one interviews through social media.

4. Results and Findings

According to the research have been found, the awareness of hydroponic crop farming system will reduce the cost of labor for hiring workers. The Fig. 3 below summarize the responses.

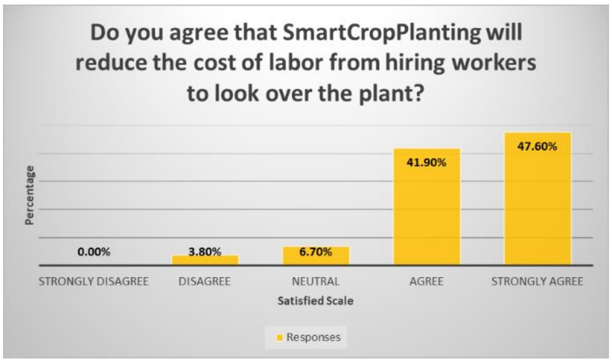


Fig. 3. Survey Question (Cost)

From Fig. 3, we found out that 89.5% of respondents agreed that SmartCropPlanting will get to reduce the cost of labor to take care of the plants and 3.8% of respondents disagree with that. In this case, our expectation for this survey is accurate and the hypothesis for this research is appropriate. In the plantation industry, it requires a lot of manpower to fulfill the tasks of operations. This sector not only far from the city but it also dirty, dangerous, and difficult. So, with the combination of technology, the industry will be able to reduce the cost of labor at the same time it solves the problem for farmers who lack manpower. With around 90% of respondents agree on automation will reduce the cost of labor, it means the research objectives for cost-related is valid. It helps in solving the major issue of the plantation industry which is the budget. By reducing cost farmers get to increase their profit margin and able to expand their business larger than before.

In the interview we asked about "If there is a way for farmer's organization to detect and monitor their plant's health instantly, should it be implemented? Why?" and most of the answer is yes. It means the image processing which is one of our system features is worth implementing and matches the research objective. With this feature, it will help farmers to detect the disease and check the condition of the plant anytime. Many responses were answered to prevent further spread and increase the efficiency of the overall tasks. The research objective seems to be valid due to the interview responses. With the answer of increased efficiency, stopping disease for further spread and quickly rehabilitate of plant these show that by monitoring the plant's health instantly is required by the farmer or the organization to have the daily routine easier and take lesser time to complete all of it.

5. Conclusion

In conclusion, this application provides the user full control and monitors their hydroponic crop remotely. Users can monitor crop parameters such as pH of the water, the temperature of the water, crop condition, artificial light condition, and water cooler condition. With this feature on the system, it not only safe money on lab equipment it also safe time and reduce infeasible task for the user. Therefore, users can achieve the best plant crop health and harvest plants more effectively. Although this application is used for hydroponic system, but it might be harder to implement for a larger crop. Especially, the image processing and live streaming of hydroponic crops, and the larger the crop the higher the budget. Therefore, this project is only built for home hydroponic systems. For future work, this project can include controlling water pH function, water ppm, and having a backup battery which will bring a huge benefit for hydroponic crop system.

References

[1] Sarkar, M. S. K., Begum, R. A., Pereira, J. J., 2020, "Impacts of climate change on oil palm production in Malaysia", *Environmental Science and Pollution Research*, 27(January), 9760-9770.

[2] Mehra, M., Saxena, S., Sankaranarayanan, S., Rijo, J. T., Veramanikandan, M., 2018, "IoT based hydroponics system using Deep Neural Network", *Computers and Electronics in Agriculture*, 155(December), 473-486.

[3] Muhammad, E. H. C., Amith, K., Saba, A., Fatima, A.-K., Jalaa, H., Fahmida, H., Mamun, B. I. R., Ahmed, A. S., Nasser, A.-E., 2020, "Design, Construction and Testing of IoT Based Automated Indoor Vertical Hydroponics Farming Test-Bed in Qatar", *sensors*, 20(19), 1-24.

[4] Thiab, A.S., Yusoh, Z.I.M., Bin Shibghatullah, A.S.,2018, "Internet of things-security and trust in e-Business", *Journal of Engineering and Applied Sciences*, 13 (13), 4939-4948.

[5] Al_Barazanchi, I., Abdulshaheed, H.R., Shibghatullah, A.,2019, "The communication technologies in wban", *International Journal of Advanced Science and Technology*, 28 (8), 543-549.

[6] Thiab, A.S., Bin Shibghatullah, A.S., Yusoh, Z.I.M., 2018, "Internet of things-proactive security approach", *Journal of Engineering and Applied Sciences*, 13 (9), 2668-2671.

[7] Thiab, A.S., Bin Shibghatullah, A.S., Mohd. Yusoh, Z.I.,2018, "Internet of Things (IoT): Architectural framework for secure

payment mode", *Journal of Engineering and Applied Sciences*, 13 (2), 415-421.

[8] Thiab, A.S., Bin Shibghatullah, A.S., Yusoh, Z.I.M., 2018, "The role of access control and device authentication in the internet of things", *Journal of Engineering and Applied Sciences*, 13 (9), 2680-2684.

[9] Subaramaniam, K., Shibghatullah, A.S., Lean, F.L, Zainal Abidin, Z., 2019, "Intruder System with Raspberry Pi in Rural Areas of Malaysia", *The 6th Mechanical Engineering Research Day 2019*, pp. 167-168.`

[10] Zhang, J., Huang, Y., Pu, R., Pablo, G. M., Yuan, L., Wu, K., Huang, W., 2019, "Monitoring plant diseases and pests through remote sensing technology: A review" *Computers and Electronics in Agriculture*, 165(October), 1-14.

[11] Ozguven, M. M. and Adem, K., 2019, "Automatic detection and classification of leaf spot disease in sugar beet using deep learning algorithms", *Physical A*, 535(December), 1-8.

[12] Kusumo, B. S., Heryana, A., Mahendra, O., Pardede, H. F., 2018, "Machine Learning-based for Automatic Detection of Corn-Plant Diseases Using Image Processing," 2018 International Conference on Computer, Control, Informatics and its Applications (IC3INA), Tangerang, Indonesia, 93-97.

[13] Ravi, L., Mohamed, D., Sathish, K. S., Raed, A., 2020, "Automated smart hydroponics system using internet of things", *International Journal of Electrical and Computer Engineering (IJECE)*, 10(6), 6389-6398.

[14] M. Arsenovic, et al. 2019, "Solving Current Limitations of Deep Learning Based Approaches for Plant Disease Detection", *Symmetry*, 11(7), 1-21.

[15] Barbedo, J. G. A., 2018, "Factors influencing the use of deep learning for plant disease recognition," *Biosystems Engineering*, 172(August), 84-91.

[16] Ferentionos, K. P., 2018, "Deep learning models for plant disease detection and diagnosis," Computer and Electronics in Agriculture, 145(February), 311-318.

Authors Introduction

Sung Jun Kyu

He received the Bachelor of Science (Hons) in Computing from Institute of Computer Science and Innovation (ICSDI) in UCSI University, Malaysia.

Assistant Prof Ts Dr. Chit Su Mon

She received the BSc (Hons) IT: Computer Science from Lancaster University, UK as well as BSc (Hons) Information Technology from Sunway University, Malaysia in 2009, the MSc Networked and Mobile Systems from Lancaster University, UK, MSc Mobile Systems from Sunway University, Malaysia in 2012 and Doctor of Philosophy (Computing) from Sunway University, Malaysia in 2023. She is currently assistant professor at School of Mathematical and Computer Sciences, Heriot-Watt University, Malaysia. Her current research interests include Haptic, Virtual Reality, Augmented Reality, Visually Impaired, Human Computer Interaction.

Asst. Prof. Ts. 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. She obtained her doctoral degree from Malaysia University of Science & Technology. She has supervised many undergraduate students as main supervisors and co-supervisors. She has publications in Scopus-Indexed Journals and Web of Science. She was also a co-researcher in Pioneer Science Incentive Fund (PSIF) in the area of augmented reality. Her research interests include human-computer interaction, human personality types, augmented reality, e-learning, mobile commerce and e-commerce.

Detection of Eye Misalignment Using an HMD
with an Eye-tracking Capability

Yoki Nagatomo, Noriyuki Uchida, Takuya Ikeda, Kayoko Takatsuka, Masayuki Mukunoki, Naonobu Okazaki

Faculty of Engineering, University of Miyazaki
1-1 Gakuen Kibanadai-nishi, Miyazaki-City, Miyazaki, japan
E-mail: takatuka@cs.miyazaki-u.ac.jp

Abstract

In this study, we implemented the Cover Test, a test method for diagnosing eye misalignment using a head-mounted display with an eye-tracking capability. Specifically, we created a virtual examination environment in a VR space. The eye-tracking technique collected eye movements immediately after the covering or uncovering of the eyes. Thus, we calculated the amount of eye deviation and developed a system to determine the presence and magnitude of strabismus and heterophoria.We assessed the system in the verification experiment by examining the consistency between the judgment results provided by this system and the clinical evaluation approach with the Maddox rod. The result was that we could verify the horizontal eye movements more accurately.

Keywords： eye position examination, cover test, strabismus/heterophoria

1. Introduction

Eye misalignment, as represented by strabismus and phoria, means abnormalities in the way we see. When one eye is out of alignment when fixating on an object, it is called strabismus, which prevents stereoscopic vision. Although stereopsis is possible, it is a condition of misaligned eyes for unknown causes. In most cases, abnormalities in the ocular muscle cause strabismus and phoria, and the abnormal location defines the treatment plan. Even if the ocular muscles are healthy, the strabismus-like symptoms may appear as a sign of a fatal disease such as cerebral infarction caused by abnormal cranial nerves. Regardless, early detection is essential, though qualified personnel are in low supply, and the practice is time-consuming[1]. Since the patient had to look at the indicator from a distance of five meters, a substantial amount of space was necessary. Therefore, in our previous study, we developed a simple and space-saving eye position testing method[2]. Specifically, we automated the conventional eye position testing method Cover-testby controlling the LCD shutter of the 3D glass with a microcontroller board Arduino. We also constructed a system that calculates the amount of eye deviation from the amount of misalignment by analyzing moving images of the eye captured by a web camera installed in the 3D glasses[2][3][4]. However, this kind of method could not completely resolve the problems of eyelid and eyelash movements as well as the reflection of fluorescent light on the surface of the eyeball, which affects the accuracy of detection. We had to ask the examinee not to blink as much as possible throughout the examination to eliminate the influence of eyelids and eyelashes. Eliminating the lighting effects was required as necessary for each examination room.

2. Purpose of Research

Our study aims to create a novel eye misalignment detection system that is unaffected by eyelid movements, eyelashes, or exam room lighting conditions. Using a head-mounted display (HMD) with an eye-tracking capability, we developed a technique to determine the existence of abnormalities. In a virtual environment, we set up an examination environment and implemented the Cover-Test to detect the amount of eye deviation from the data of eye movements retrieved from the eye tracking function during the eye covering and uncovering.

In the verification experiment, we assessed the system by comparing the results of an examination using the Maddox rod (hereafter abbreviated as the Maddox test) with the detected amount of eye deviation .

3. Current method

3.1 Cover-Test

Based on the movements of the eyes when covered, the CoverTest distinguishes strabismus from orthophoria. Strabismus occurs when one eye moves when the other eye is covered, and phoria occurs when the occluder is removed. There are two types of cover tests: the Alternating Cover Test (ACT) and the Cover-Uncover Test (CUT). We perform ACT by always covering one eye to detect the total deviations, manifest and latent. To distinguish between strabismus/heterophoria and orthophoria, we perform CUT by covering one eye and observing its motions, including the activities in the other eye when exposed. These are two procedures we used in this study.

3.2 Maddox Rod Test

The Maddox test is a procedure to measure subjective total deviation in cases of binocular vision, such as strabismus and micro-angle strabismus with peripheral fusion[5]. As shown in Fig. 3, the examinee fixates one eye at a point lighted by a pen torch by a five meter distance, and the other covered with a Maddox rod. (Fig.1-(2)) If there is no ocular deviation causing binocular disparity, the point light source of the uncovered eye coincides with the red light ray perceived by the other eye. (Fig. 1-(2)). When there is an ocular deviation, such as strabismus or heterophoria, the point light source and the red light ray do not coincide. (Fig.1-(3)). If there is no match, we place a prism bar in front of the eye covered with the rod and slide it from the weakest to the highest prism degree until the point source matches the red light ray. The prismatic power at the point of coincidence is the actual ocular deflection of the examinee.

4. Method Proposal

4.1 Proposed System

Fig.2 describes the hardware configuration of the system. Fig.2-A characterizes the HMD with an eye-tracking function. The camera and infrared sensor of the HMD record eye movements for data[6][7]. The examination environment comprises a test room with an optotype, a red dot displayed in he middle, and occluders. When the program functions, the HMD displays a screen shown in Fig.3 in a binocular stereoscopic form. While the examinee fixates on a target on the screen, we performed the covering and uncovering based on ACT and CUT. We used an occluder, which turns transparent and non-transparent, making the cover uncover possible.
During this procedure, we continued to collect the data to test strabismus from the degree of the eye movements: ACT calculates the total deviation of both eyes. CUT determines strabismus and heterophoria by calculating the deviation of strabismus (strabismus amount) and heterophoria (heterophoria amount).

4.2 Functional Structure

Fig.4-(1) presents a block diagram of the system's functional structure, whereas Fig.4-(2) and Fig.4-(3) present the function composition of the "HMD" and "controller," respectively. Fig.4-(4) in the flowchart describes the overall flow of the ACT examination (Chapter 3.1), and (5) - (7) specifies each step of the procedure (4). The flow of (4), (5), and (7) for the CUT examination are the same (Chapter 3.1); The difference in procedure between ACT and CUT is the processing procedure for acquiring eye data in (6). This is because the varieties of eye movements we should observe for abnormality determination are different. Fig.4-(6) shows a flowchart for occluding and exposing the eye, effective at manifesting abnormal judgment through eye movements.

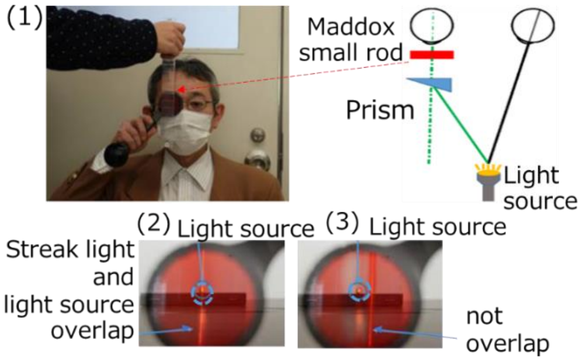


Fig.1 Inspection with the Maddox rod

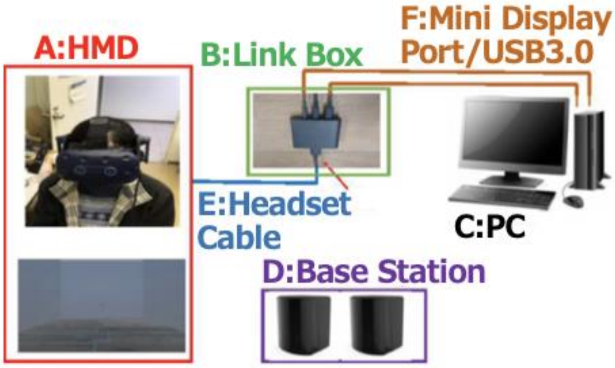


Fig.2 Configuration of the system

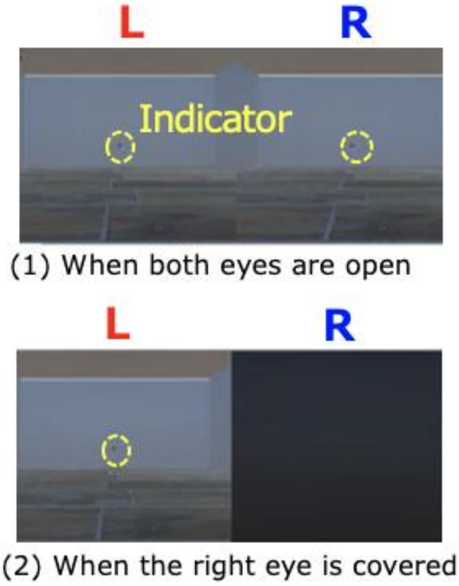

L
R
Indicator
(1) When both eyes are open
L
R
(2) When the right eye is covered

Fig.3 Image as it appears to the examinee during the Cover-Test

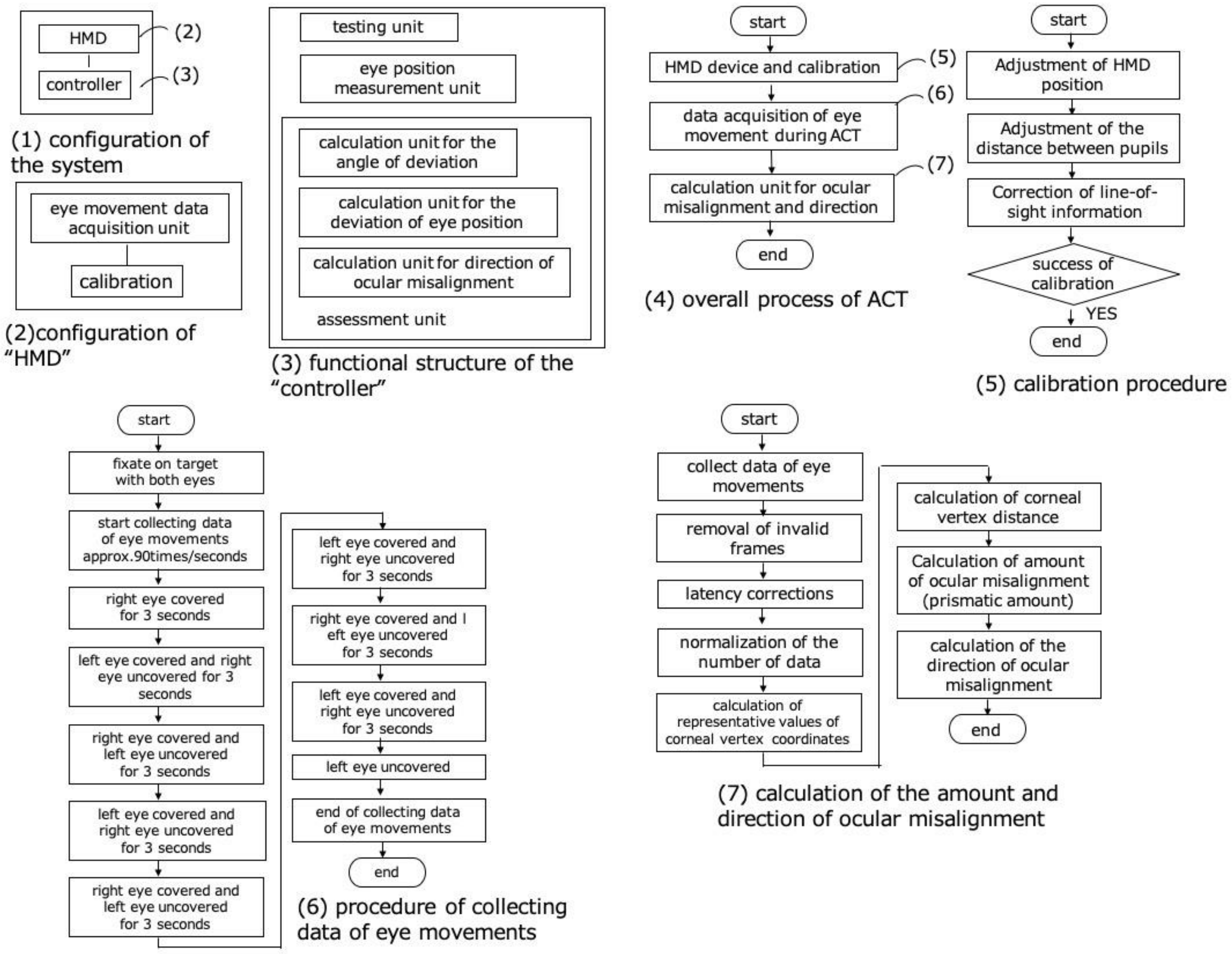

HMD
controller
(2)
(3)
(1) configuration of the system
eye movement data acquisition unit
calibration
(2)configuration of "HMD"
testing unit
eye position measurement unit
calculation unit for the angle of deviation
calculation unit for the deviation of eye position
calculation unit for direction of ocular misalignment
assessment unit
(3) functional structure of the "controller"
start
HMD device and calibration
data acquisition of eye movement during ACT
calculation unit for ocular misalignment and direction
end
(5)
(6)
(7)
(4) overall process of ACT
start
Adjustment of HMD position
Adjustment of the distance between pupils
Correction of line-of-sight information
success of calibration
YES
end
(5) calibration procedure
start
fixate on target with both eyes
start collecting data of eye movements approx.90times/seconds
right eye covered for 3 seconds
left eye covered and right eye uncovered for 3 seconds
right eye covered and left eye uncovered for 3 seconds
left eye covered and right eye uncovered for 3 seconds
right eye covered and left eye uncovered for 3 seconds
left eye covered and right eye uncovered for 3 seconds
right eye covered and l eft eye uncovered for 3 seconds
left eye covered and right eye uncovered for 3 seconds
left eye uncovered
end of collecting data of eye movements
end
(6) procedure of collecting data of eye movements
start
collect data of eye movements
removal of invalid frames
latency corrections
normalization of the number of data
calculation of representative values of corneal vertex coordinates
calculation of corneal vertex distance
Calculation of amount of ocular misalignment (prismatic amount)
calculation of the direction of ocular misalignment
end
(7) calculation of the amount and direction of ocular misalignment

Fig.4 Functional Configuration and Operation

5. Experiment

5.1 Content of Experiment

We used the Maddox test and the system to compare the quantified results of the ACT and CUT tests in the experiment. Table.1 and Table.2 show the quantitative results. XT/XP and ET/EP represent exotropia and esotropia, respectively. Exotropia occurs when one eye moves outward when fixating on an object, while esotropia occurs when the opposite eye turns inside. Blue in the table shows the results of cases considered exotropia of 1△ or more and those in red considered esotropia of 1 △ or more.

5.2 Assessment/Discussion

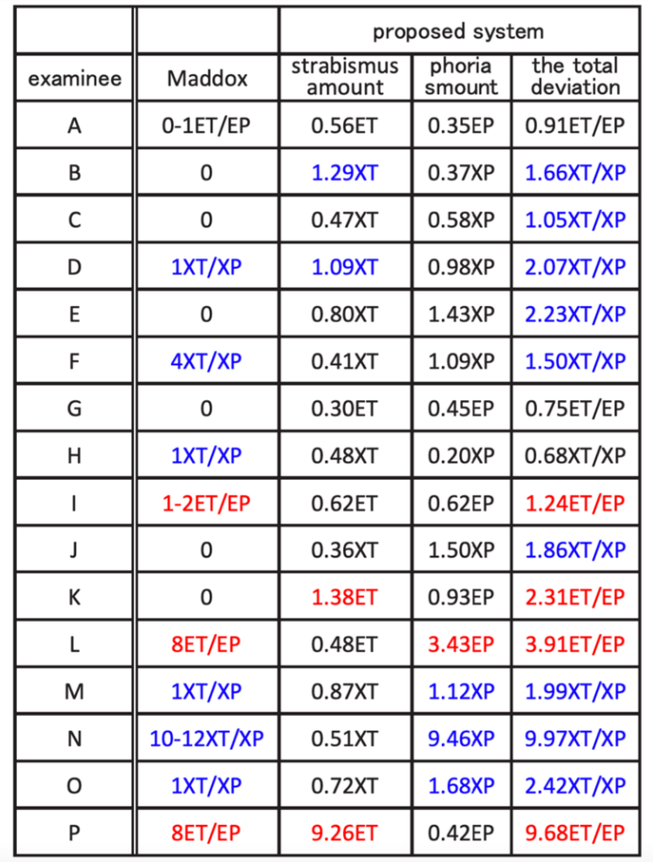
Table.1 points out the total horizontal deviations for the Maddox test and ACT. The mean error for both tests was 0.61 △, less than 1 △, which is close to the correct answer.

Table.2 gives the amount of horizontal strabismus and obliqueness in the Maddox test and CUT, together with the total deviation generated by adding them up. The results show that the average error of the CUT and the Maddox test is 1.50 △, which is insufficiently accurate.

Table.1 measured value of a total amount of deviation

examinee	Maddox	proposed system
A	0-1ET/EP	0.88ET/EP
B	0	0.52XT/XP
C	0	0.36ET/EP
D	1XT/XP	1.44XT/XP
E	0	0.52XT/XP
F	4XT/XP	3.22XT/XP
G	0	1.06ET/EP
H	1XT/XP	0.88XT/XP
I	1-2ET/EP	2.27ET/EP
J	0	0.50XT/XP
K	0	0.64ET/EP
L	8ET/EP	9.51ET/EP
M	1XT/XP	0.57XT/XP
N	10-12XT/XP	12.10XT/XP
O	1XT/XP	1.49XT/XP
P	8ET/EP	7.91ET/EP

Table.2 measurements of strabismus and obliquity



		proposed system		
examinee	Maddox	strabismus amount	phoria smount	the total deviation
A	0-1ET/EP	0.56ET	0.35EP	0.91ET/EP
B	0	1.29XT	0.37XP	1.66XT/XP
C	0	0.47XT	0.58XP	1.05XT/XP
D	1XT/XP	1.09XT	0.98XP	2.07XT/XP
E	0	0.80XT	1.43XP	2.23XT/XP
F	4XT/XP	0.41XT	1.09XP	1.50XT/XP
G	0	0.30ET	0.45EP	0.75ET/EP
H	1XT/XP	0.48XT	0.20XP	0.68XT/XP
I	1-2ET/EP	0.62ET	0.62EP	1.24ET/EP
J	0	0.36XT	1.50XP	1.86XT/XP
K	0	1.38ET	0.93EP	2.31ET/EP
L	8ET/EP	0.48ET	3.43EP	3.91ET/EP
M	1XT/XP	0.87XT	1.12XP	1.99XT/XP
N	10-12XT/XP	0.51XT	9.46XP	9.97XT/XP
O	1XT/XP	0.72XT	1.68XP	2.42XT/XP
P	8ET/EP	9.26ET	0.42EP	9.68ET/EP

6. Conclusion

In this study, we automated the Cover-Test, an eye misalignment test, and used it to create a detection system. Using VR, we created a virtual testing environment. In this setting, we automated several steps, including the presentation of an optotype to the acquisition of the eye position, analysis, and test result acquisition during the cover-uncover operation. It is now possible to perform examinations in a space-saving setting. This resolves the issues from earlier studies where eyelid and eyelash movement interfered with the test's accuracy and the exam room lighting conditions. The HMD is more comfortable to wear and less burdensome than the earlier 3D glasses, making it easier to use. We assessed the system by comparing the results from ACT and the Maddox test, an actual clinical test. As a result, it provided correct outcomes in ACT but could not identify microscopic abnormalities in CUT. ACT showed a mean error of 0.61 △ (prism dioptre) and a correlation of 0.956, demonstrating excellent correlation and practicality. However, the CUT has an average error of 1.50△ and a correlation of 0.810, a significant inaccuracy. We believe that further precision is required. The following three issues are our future challenges.

The first is to improve the accuracy of CUT. The second is to make it possible to determine whether the CUT is accurate or slightly abnormal.

Despite no abnormalities, the examinee's eyes display slight tremors (fixation micro-tremors). Therefore, it is extremely difficult to tell if it is orthophoria. The goal of this system is to enable differentiation between the two. The need to differentiate orthopia, minor strabismus, and phoria is not clinically high. However, it is a critical issue for the system's development. We may use the findings gained through the development process in other systems using VR in the future.

The third is to detect vertical misalignment. The current system can monitor horizontal deflection with excellent accuracy despite the vertical measurement's low accuracy. We need a considerable amount of data to clarify how to improve the accuracy. Data on subjects with severe vertical abnormalities are rarely available, and we view data collection as a short-term issue.

References

1. Fukuda Toshimasa; Japan Ophthalmologists Association Board of Public Health, Three-year-old child ophthalmological health checkup survey report, Ophthalmology in Japan, 85-3, pp.296-300 (2014)
2. Noriyuki Uchida, Kayoko Takatuka, Hisaaki, Yamaba, Masayuki Mukunoki, Naonobu Okazaki; Method for detecting eye misalignment based on movement near the center of the pupil, International Conference on Artificial Life and Robotics (ICAROB 2021), OS-15-3
3. N Uchida, K Takatuka, A Nakazawa, M Mukunoki, N Okazaki；HMD-based Cover Test System for the Diagnosis of Ocular Misalignment, Journal of Artificial Life and Robotics, Springer, pp.1-6 (2019),
4. Noriyuki Uchida, Kayoko Takatuka, Hisaaki Yamaba, Masayuki Mukunoki, Naonobu Okazaki；Development of a System to Detect Eye Position Abnormality based on Eye-Tracking, Journal of Robotics, Networking and Artificial Life, Vol. 8(3), pp.205–210 (2021)
5. Manual for Orthoptist, Quantitative ocular position examination Self-perceived strabismus angle assessment (using red glass or [Maddox rod]) (accessed 2021-1-20).
6. VIVE Pro eye overview (accessed 2020-12-27).
7. SRanipalSDK (accessed 2019- 9-28).

Authors Introduction

Yoki Nagatomo

He was received the B. E degrees in computer science and system engineering from Miyazaki University, Miyazaki Japan, in 2022. He is now a master course student of Miyazaki University. His research interest virtual realit

Noriyuki Uchida

He is an orthoptist. He is currently a Doctoral course student in Miyazaki University, Japan. His research interests include visual optics, strabismus and amblyopia.

Takuya Ikeda

He was received the M. E degrees in computer science and system engineering from Graduate school of Miyazaki University, Miyazaki Japan, in 2021. His research interests virtual reality.

Kayoko Takatsuka

She is an Expert technical staff with the Faculty of Engineering, University of Miyazaki, Japan. She graduated from Kumamoto University, Japan, in 1992, and the Ph D. degree in interdisciplinary graduate school of agriculture and engineering from University of Miyazaki, Japan, in 2014. Her research interest is process systems engineering

Masayuki Mukunoki

He is now a Professor of Faculty of Engineering at University of Miyazaki in Japan. He received the bachelor, master and doctoral degrees in Information Science from Kyoto University. His research interests include computer vision, pattern recognition and video media processing.

. 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.

Using OpenCV for real-time image recognition through augmented reality devices

Gabdullina Dinara
Kazan Federal University, Kazan, Russian Federation, 420008, Rep. Tatarstan, Kazan,
St. Kremlin, 35, Russia

Zykov Evgeniy
Kazan Federal University, Kazan, Russian Federation, 420008, Rep. Tatarstan, Kazan,
St. Kremlin, 35, Russia

Kugurakova Vlada
Kazan Federal University, Kazan, Russian Federation, 420008, Rep. Tatarstan, Kazan,
St. Kremlin, 35, Russia
E-mail: dinaragabdullina086@gmail.com, Evgeniy.Zykov@kpfu.ru, vlada.kugurakova@gmail.com

Abstract:

This article describes the peculiarities of writing libraries and the rapid development of augmented reality applications on the Unity platform and the OpenCV open-source library. The application of OpenCV functions to work with augmented reality objects is discussed in detail. On their basis, algorithms and methods were developed and an open library was created that implements the stable recognition of objects in various conditions by means of binding to the marker and their visualization on the augmented reality devices. On the basis of the developed library, a complete application was created to illustrate the possibility of applying virtual reality technologies for optical control of the assembly of radio electronic boards and for personnel training.

Keywords: Augmented reality, Pattern recognition, Computer vision, Optical control, Industrial automation.

1. Introduction

Augmented reality is a good tool for implementation in various industries, but the prevalence of the use of augmented reality applications in industry is still low. The one article says [1], that the reasons for the non-acceptance of new technologies are described: risk factors, which include time risks, financial and security risks, and the factor of psychological acceptance of the technology.

The list of successful examples of the use of augmented reality in industrial processes could go on and on. But so far, every augmented reality application is a "one-off product", so there is a need to streamline the development of augmented reality applications, easily using different glasses and relying on proven approaches to sustainably recognize dynamically changing 3D objects in different lighting conditions, expanding the range of applications in industrial processes – such as conveyor assembly.

However, in order to solve all these problems a considerable amount of work needs to be done, so it was decided to start by narrowing down the scope of the tasks to be solved, with possible scaling up by adding more functions and algorithms.

So, the task was to create an open library of sustainable object recognition in different light conditions, visualizing them and other related information on augmented reality devices based on freely distributed software.

2. SDK analysis for AR

ARKit [2] is an SDK for creating augmented reality applications for the IOS operating system. Available features include environmental depth estimation using the Depth API and LiDAR camera, capturing human motion with the camera and identifying human bones, and creating topological space maps.

ARCore [2] allows the development of Android and iOS applications. The depth of the environment can be measured with an RGB camera by creating a depth map. There is a light estimation function, marker recognition.

The Vuforia augmented reality SDK [2], developed by Qualcomm, provides capabilities for creating Android and iOS applications. With Vuforia, you can recognize objects from a 3D model, scan objects in space, and recognize groups.

Wikitude [2] is another SDK, object recognition from 3D models, image recognition and tracking, point-to-point mapping using location, video overlay, smartglass integration, integration with external plugins.

3. Problem statement. Description of the basic functionality for an augmented reality library on the programmer and user side

Let's describe the basic functionality with an example: you need to recognize an object in space. It was started by defining the user's actions:

1. The user opens the application.
2. The user points the camera of the glasses at the workspace.
3. A frame appears around the object of interest, indicating that recognition is correct. You will be prompted for further action.
4. Repeat steps 2-3 until the work is completed. Close the application when the work is completed.

Let's describe the basic functionality on the developer side:

1. Login to the app and initialize the augmented reality device.
2. The algorithm recognizes key points in the image and in the frame.
3. Finding correspondences between the key points.
4. Drawing a frame around the object.
5. Exit the app.

3.1 Image recognition with OpenCV

OpenCV is an Open-Source Computer Vision Library. The library is quite popular, currently having over 5 million downloads, and is the baseline for the Khronos computer vision standard. The library is adapted to different platforms. Independence from closed code of various commercial frameworks and possibility to implement new algorithmic approaches are the main advantages of using OpenCV.

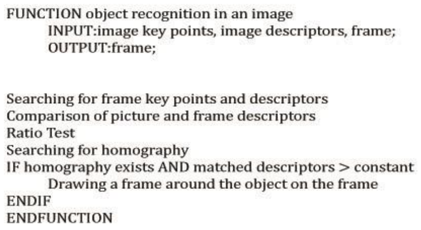
The OpenCVSharp plus Unity resource kit [3] and Unity 2021.2.10f1 development environment [4] were used to implement this functionality.

ORB is an algorithm for identifying key points in an image. According to [5], this algorithm is the best algorithm for low performance devices. FAST algorithm [6] is used to select the key points: a pixel with 16 pixels is compared, if there is more than half of the pixels darker or brighter than this one, it is selected as the key point. A binary feature vector is used to describe the key points using the BRIEF algorithm [7]. To begin with, smoothing is performed, then a pair of pixels is selected around the key point and the intensities are compared. Based on the result, a value of 0 or 1 is selected.

The Flann Matcher algorithm [8] was used to match the points from the image and the camera frame. Unlike the Brute-Force Matcher algorithm [9], this algorithm is faster on large data sets because it is an implementation of the k-d-tree algorithm [10]. The complexity of the algorithm is at best $O(h)$, where h is the height of the tree.

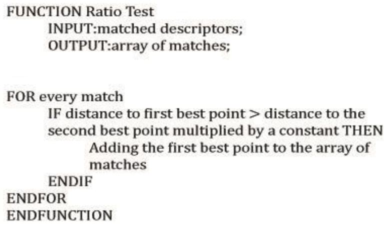
The Lowe's Ratio Test algorithm [11] was used to remove key points that are noise. In the next step, a frame was mapped around the recognized image. In order to map the corresponding points from the image to the frame, a homography, which is a 3×3 matrix, was used.

The algorithm pseudocode shown in Fig. 1 and Fig. 2, shows the steps described earlier.



```
FUNCTION object recognition in an image
        INPUT:image key points, image descriptors, frame;
        OUTPUT:frame;


Searching for frame key points and descriptors
Comparison of picture and frame descriptors
Ratio Test
Searching for homography
IF homography exists AND matched descriptors > constant
        Drawing a frame around the object on the frame
ENDIF
ENDFUNCTION
```

Fig. 1. Pseudocode to recognize key points on a frame and display the frame



```
FUNCTION Ratio Test
        INPUT:matched descriptors;
        OUTPUT:array of matches;


FOR every match
        IF distance to first best point > distance to the
        second best point multiplied by a constant THEN
                Adding the first best point to the array of
                matches
        ENDIF
ENDFOR
ENDFUNCTION
```

Fig. 2. Pseudocode for deleting keypoints

Recognition of key points in the image is performed only once, at the beginning of work, and recognition of key points in the frame is performed on each frame from the camera. It should be said that the correctness of the frame display depends on environmental factors, so the frame is drawn only if more than 20 points are found when Lowe's Ratio Test is applied.

The app was run on the Vuzix M400 Starter Kit augmented reality glasses.

3.2 Improving the algorithm

After running several tests, the weaknesses of the developed algorithm were revealed. Firstly, the frame was not always displayed correctly on the scene. Secondly, it required a certain amount of time and good lighting to recognize a sufficient number of key points. In addition, jitter, or camera shake led to the loss of an object followed by the initiation of a new search, resulting in a chaotic flicker on the screen.

Thus, the following code improvements have been suggested:

1. Stabilization of the frame display;
2. Image pre-processing:
   2.1. bringing it into greyscale format,
   2.2. overlay filters to get rid of noise and sharpen the image;
3. Experiment with different levels of light in the workplace and choose the most appropriate filters.

In order to stabilize the frame display, it was decided to draw a single object, Cv2.Rectangle [12].

A conditional operator has also been added: if after Lowe's ratio test the matched points are less than the set threshold value, in this case 30, and at the same time greater than 20, the frame coordinates are not recalculated, but the previous ones are taken. In the other case, it is updated according to the code. If the homography is not counted, then the average value of the last 10 frames is searched using AccumulateWeighted [13].

3.3 Implementing the functionality in the application

Following this research, the functionality has been implemented in an application to assist the technician in the lead assembly of radio components on a printed circuit board (Fig. 3) [14].

The advantages of manual assembly are:

- High flexibility when changing production sites.
- The possibility of a permanent visual check, allowing defects in boards or components to be detected in good time.

The disadvantages are: low productivity, high labour intensity of the technological process, use of sufficiently qualified operating personnel (3rd-4th categories) with the necessary practical and theoretical knowledge [15].

The use of automated optical inspection systems allows higher inspection confidence and minimizes defects at high assembly speeds.

There are few companies in the Russian electronics market that produce products in large batches. This makes the development of low-cost optical inspection systems an urgent task.

The basic requirements for budget optical inspection systems are [16]:

- identifying the main types of defects – missing, misaligned, misaligned components, jumpering and lack of contact in solder joints,
- a visual presentation of information on the locations of suspected defects,
- low cost.

Based on these requirements, it is possible to formulate a number of tasks that need to be solved when creating budget optical inspection systems [12]:

- obtain an image of the reference object and obtain an image of the object under test,
- perform an initial processing of the acquired images,
- identify the presence of defects on the object under test and display a message on the type and location of the suspected defect.

On the basis of the above, the application contains the following functionality: the recognition of the board and the radio components, the display of the name, the quantity, the position on the board and the polarity. Also, if you look at the example instructions [17], the resistance or capacitance of the radio component is also indicated. This is the basic information required for a correct assembly.

Other article [18] says that the assembly procedure is covered. The main points that can be highlighted are checking the polarity of the part, fitting the part to the board, trimming the pins back and starting soldering, once all the parts have been fitted.

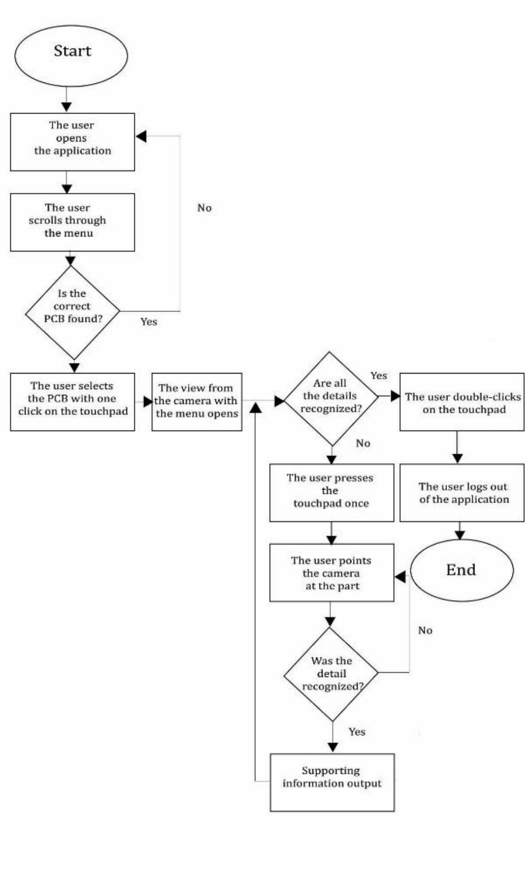

Start
The user opens the application
The user scrolls through the menu
Is the correct PCB found?
No
Yes
The user selects the PCB with one click on the touchpad
The view from the camera with the menu opens
Are all the details recognized?
Yes
No
The user double-clicks on the touchpad
The user logs out of the application
End
The user presses the touchpad once
The user points the camera at the part
Was the detail recognized?
No
Yes
Supporting information output

Fig. 3. Application flowchart

A mask in computer vision is used to highlight pixels with specific values in an image [19]. The use of a mask for colour recognition of radio components is a solution to the problem as these radio components do not have a lot of inscriptions or drawings on the surface. Let's take the red diode as an example.

The HSV colour extremes were first set to (155, 25, 111) and (180, 255, 255). Each frame was then converted to HSV format. To filter specific colours, the InRange function [20] was used to feed the boundary colour values to form a mask. Function Dilate [21] was used to expand mask – increase white area of the mask and get rid of possible noise. The mask was overlaid on the frame using BitwiseAnd [22].

The Cv2.FindContours method [23] was used in OpenCV to select object contours on the frame. The method takes an image on which you want to find contours. To save memory, we used the method CHAIN_APPROX_SIMPLE, which returns only the outermost points at the outline. These reference points were used to draw a rectangle in the outline area.

Different radio components have their own polarity characteristics. If the polarity is not correct, the radio element may fail. To avoid this, the user is prompted in the top panel to check the polarity [24]. There can be several identical radio components, so the number is also displayed.

As all code was implemented in C# in Unity, the interface was also created in the same environment. Due to the fact that the screen size for a single eye and its resolution is quite small, the interface should be as simple as possible, with simplified UI elements and no complex hierarchy. At the same time, only the information that the user needs at a given step should be displayed [25]. The interface is controlled by the touchpad using the Key Codes function [26].

4. Image pre-processing

The introduction of filters and lighting experiments was decided to be carried out in the last phase of the implementation, in order to obtain the most representative experimental results.

Using different filters for image processing allows you to extract additional information needed to correctly identify key points. Consider the following image processing methods: grayscale conversion, Gaussian blur, and histogram alignment.

The image in OpenCV is converted to greyscale format [27] using a formula:

$$RGB\ to\ Gray: \quad Y \leftarrow 0{,}299 * R + 0{,}587 * G + 0{,}114 * B, \tag{1}$$

where Y – is the result of the conversion, *R is* the red channel, *G* – green channel, *B* – blue channel.

A conversion is made from a three-channel image to a single-channel image with values in the range 0 to 255.

Smoothing by mean value is one of the simplest smoothing methods. This method takes an area of pixels of a certain size and finds the average value for the pixel value you are looking for [21].

To determine what size the counting area of the mean will be, a kernel is used, which is a matrix with dimensions $M \times N$, where M and N - are odd numbers [28]. Also, the larger the dimensionality of the kernel, the stronger the smoothing. Example of kernel (Eq. 2):

$$K = \frac{1}{K_{height} * K_{width}} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \tag{2}$$

where K – 3×3 core.

Gaussian antialiasing is used to smooth out sharp corners and get rid of noise. This smoothing is based on the Gaussian function formula in two-dimensional space. It is applied to each pixel, after which a convolution matrix is formed. The pixel value is updated using a weighted average formula based on the kernel dimension.

Impulse noise, or randomly occurring black and white pixels, is a type of noise in images. For this case, a median filter is used and a kernel is also used, but the median in a given vicinity rather than the mean value is considered [21].

Histogram alignment is a method of increasing the intensity range of an image. This operation is performed by overriding the pixel value using a cumulative distribution function [29]. If, however, there are heavily overexposed parts of the image, conventional histogram equalization will lose information in these areas due to the relatively large range of variation. Adaptive histogram equalization helps to solve this problem by applying the histogram equalization operation to parts of the image separately [30].

5. Introducing image preprocessing

The recognized image and the camera image were processed in the following order: conversion to grayscale, Gaussian smoothing and adaptive histogram equalization. There are corresponding methods in OpenCV: Cv2.CvtColor(), Cv2.GaussianBlur(), Cv2.Blur(), Cv2.MedianBlur, CLAHE.Apply().

A comparison was made between the recognition times before and after filter processing. This experiment was due to the fact that the developed application runs in real time. Experiments were conducted with different degrees of light, the time of correct recognition was considered the moment when the dialog prompt appeared in the upper panel. The time was measured using the StopWatch class functions: Start() and Stop.

The degree of illumination was measured using a Testo 540 luxmeter. To determine the mean, medians were chosen as the most resistant to experimental emissions values [31].

The graph (Fig. 4) shows that the recognition is much faster when image preprocessing is used, and the lower the illumination, the higher the effect of preprocessing on the object recognition time.

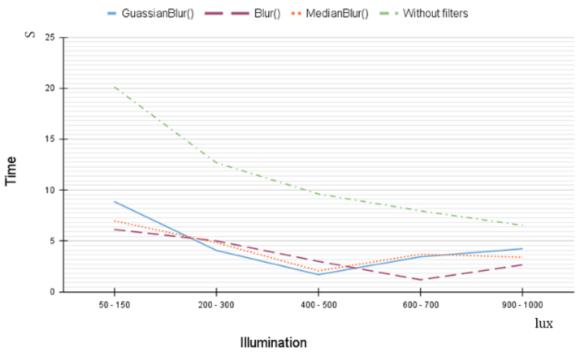


Fig. 4: Graphs without processing and processing by low-pass filters with a kernel

6. Conclusion

The algorithms of the OpenCV library allow the creation of basic functionality for augmented reality applications. Various combinations of image preprocessing, keypoint recognition and matching methods produce results that can be used for a variety of purposes. The algorithms considered combine to provide image recognition functionality in space.

The application also revealed weaknesses such as the dependence on lighting quality and on the speed of the software to display the frame on each frame. Applying pre-filtering and remembering previous values has greatly improved the display. Thus, proper use of OpenCV provides the opportunity to design a variety of functionality.

Acknowledgements

The work was carried out at the expense of the Strategic Academic Leadership Program of Kazan (Volga Region) Federal University ("PRIORITY 2030").

References

1. K.E. Schein, P.A. Rauschnabel, "Augmented Reality in Manufacturing: Exploring Workers' Perceptions of Barriers", IEEE Transactions on Engineering Management, pp. 1-14, 2021.
2. Y. Chen, Q. Wang, "An H.C. Overview of augmented reality technology", Journal of Physics: Conference Series, 2019.
3. OpenCV plus Unity (2019). (Accessed: 2 February 2022).
4. Unity 2021.2.10 (2021). (Accessed: 5 February 2022).

5. A. Muryan, Y.Verdi, “Application of oriented fast and rotated brief (orb) and bruteforce hamming in library opencv for classification of plants”, JISAMAR (Journal of Information System, Applied, Management, Accounting and Research), pp. 51-59, 2020.
6. P. Chhabra, N.K. Garg, M. Kumar, “Content-based image retrieval system using ORB and SIFT features”, Neural Computing and Applications, pp. 2725-2733, 2018.
7. F. Siddiqui, S. Zafar, S. Khan, N. Iftekhar, “Computer Vision Analysis of BRIEF and ORB Feature Detection Algorithms”, Applied Computational Technologies, pp. 425-433, 2022.
8. V. Vijayan, P. Kp, “FLANN Based Matching with SIFT Descriptors for Drowsy Features Extraction”, Fifth International Conference on Image Information Processing (ICIIP), pp. 600-605, 2019.
9. N. Antony, B.R. Devassy, “Implementation of Image/Video Copy- Move Forgery Detection Using Brute-Force Matching”, International Conference on Trends in Electronics and Informatics (ICOEI), pp. 1085-1090, 2018.
10. Y. Chen, L. Zhou, Y. Tang, “Fast neighbor search by using revised k-d tree”, Informatics and Computer Science Intelligent Systems Applications, pp. 142-162, 2019.
11. D.G. Lowe, “Distinctive Image Features from Scale-Invariant Keypoints”, International journal of computer vision, Vol. 60. pp. 91-110, 2004.
12. Drawing functions in OpenCV (2022). (Accessed: 20 March 2022).
13. E. Niloy, J. Meghna, M. Shahriar “Hand Gesture-Based Character Recognition Using OpenCV and Deep Learning”, International Conference on Automation, Control and Mechatronics for Industry 4.0 (ACMI), pp. 1-5, 2021.
14. PCB Mounting Output Process (2022). Accessed: 25 March 2022).
15. V.L. Lanin, “Assembly technology, installation and control in the production of electronic equipment”, Bel. State University of Informatics and Radioelectronics. – Minsk: Inpredo. p. 64, 1997. (In Russian).
16. A. Ovchinnikov, “The concept of building budget systems for optical inspection of the quality of printed circuit board assembly”, Technologies in the electronic industry, Vol. 8, No. 36, pp. 41-44, 2009. (In Russian).
17. Soldering kit (2022). (Accessed: 25 March 2022).
18. A.I. Tsukanov, O.V Kuchevasov, “Technologies for mounting and dismantling components and elements of radio-electronic and radio-television equipment: a teaching aid”, St. Petersburg State Educational Institution ", College of Electronics and Instrument Engineering" St. Petersburg, p. 105, 2017.
19. OpenCV – Invert mask (2021). (Accessed: 22nd April 2022).
20. A. Qashlim, B. Basri, H. Haeruddin, A. Ardan, “Smartphone Technology Applications for Milkfish Image Segmentation Using OpenCV Library”, International Journal of Interactive Mobile Technologies (iJIM), pp. 150-163, 2020.
21. H. Singh, “Advanced Image Processing Using OpenCV”, Practical Machine Learning and Image Processing, pp. 63-88, 2019.
22. M.A. Mir, F. Qazi, M. Naseem “Invisibility Cloak using Color Extraction and Image Segmentation with OpenCV”, Global Conference on Wireless and Optical Technologies (GCWOT), pp. 1-6, 2022.
23. V. Harini, V. Prahelika, I. Sneka, “Hand Gesture Recognition Using OpenCv and Python”, New Trends in Computational Vision and Bio-inspired Computing, pp. 1711-1719, 2018.
24. Capacitor Polarity: Understanding Polarity (2022). (Accessed: 12 March 2022).
25. UI Best Design Practices (2022). (Accessed: 23 March 2022).
26. How to map Android keycode to Unity keycode (2021). (Accessed: 23 March 2022).
27. W. Hou, D. Xia, H. Jung “Video road vehicle detection and tracking based on OpenCV”, International Conference on Information Science and Education (ICISE-IE), pp. 315-318, 2020.
28. Convolutions with OpenCV and Python (2016). (Accessed: 24 March 2022).
29. Y. Wei, G. Xu, H. Liu “OpenCV-based 3D printing physical surface defect detection”, IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)”, pp. 1924-1927, 2020.
30. F. Alkhalid, A.M. Hasan, A. Alhamady, “Improving radiographic image contrast using multi layers of histogram equalization technique”, IAES International Journal of Artificial Intelligence (IJ-AI), pp. 151-156, 2021.

31. Median in statistics (2022). (Accessed: 1 May 2022).

Authors Introduction

Ms. Gabdullina Dinara

Laboratory assistant, Research Laboratory of Augmented Reality Technologies in Industrial Processes, Kazan Federal University. Research area: augmented reality, virtual reality, computer vision

Mr. Zykov Evgeniy

Candidate of Science in Physics and Mathematics, Associate Professor, Department of Radio Astronomy, Kazan Federal University. Research area: image processing, artificial intelligence, neural networks, robotics, augmented reality, bionics.

Ms. Kugurakova Vlada

Candidate of Technical Sciences, Associate Professor of the Department of Software Engineering, Kazan Federal University. Research area: virtual reality, augmented reality, 3D models, game engines, artificial intelligence.

A Structure Pattern Extraction by Using Morphological Component Analysis in the Aerial Image Edge Detection

Wataru Oshiumi[1]
[1]Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-Ku, Kitakyushu, 808-0196, Japan

Ankur Dixit[1]
[1]Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-Ku, Kitakyushu, 808-0196, Japan

Hiroaki Wagatsuma[1,2]
[1]Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-Ku, Kitakyushu, 808-0196, Japan
[2]RIKEN CBS, Japan

E-mail: oshiumi.wataru745@mail.kyutech.jp, ankur.dixit003@gmail.com, waga@brain.kyutech.ac.jp

Abstract

Automated extraction of areas of interest from aerial images is a central issue for map makers. The road skeleton is one of the most frequent targets, and then the prefiltering of road structures is highly important; however, the isolation of man-made structures from a natural landscape is technically tricky because image parts to represent natural geographical features are not uniform patterns. In the present study, we focused on the function of the Morphological Component Analysis (MCA) method to extract structural patterns when dictionaries were appropriately given and demonstrated the effectiveness of edge detection if the prefiltering was done by using MCA. MCA decomposes the image into two patterns in our computer experiment with Curvelet transform and Local Discrete Cosine Transform (LDCT) dictionaries. This approach will explore extensive possibilities of structural data extraction from complex images as an automated method.

Keywords: Morphological Component Analysis, Edge detection, Curvelet transform, Local discrete transform

1. Introduction

The detection of man-made structures from aerial photographs, such as roads on the land surface, can be performed with a certain level of accuracy using deep learning [1][2][3]. On the other hand, detections of non-uniform structures, especially with natural patterns, including forests, are still challenging issues with respect to schemes using machine learning and deep learning.

As an alternative option, it is worthwhile to apply morphological component analysis (MCA) to tackle those issues[4], which has the advantage of extracting a target texture pattern by implementing an appropriate dictionary as a signal basis. The MCA method also considered the possibility of image denoising and decomposition when extracting specific patterns in the map image data [5]. In the sense of the decomposition of components that conform image or signal, target data can be separated from the original data. In the case of aerial photography, the same strategy can be applied to obtain the target structure. In the present study, we focused on the structural decomposition method of MCA and demonstrated the effectiveness of edge detection from the decomposed data.

Therefore, the decomposition performance will be a measure to evaluate the accuracy of MCA for the present purpose the effectiveness of edge detection can be discussed. In MCA, a dictionary selection is also an issue, while as a preliminary study, we focused on the function of local discrete cosine transform (LDCT) and curvelet transform (CURVE) [6]. This approach will explore extensive possibilities of structural data extraction from complex images as an automated method.

2. Methods

2.1. *Image decomposition in sparse modeling*

In the present paper, it is assumed that an image consists of two structures. If the original image is X and the two

structures are x_0 and x_1, we can express them from Eq. (1) and Eq. (2). Furthermore, the decomposition was performed with x_0 as the background and x_1 as the structure information of the aerial photograph.
Next, consider the original image X. After decomposing from Eq. (3), structures x_0 and x_1 can represent the original image X, and the coefficients corresponding to the dictionary α_k are denoted by T_k to represent x_k from Eq. (4).

$$\mathrm{X} = x_0 + x_1 \tag{1}$$

$$\mathrm{X} = \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} \tag{2}$$

$$\mathrm{X} = \sum_{k=0}^{1} T_k \alpha_k \tag{3}$$

$$T_k \alpha_k = x_k \tag{4}$$

The basis α_k and the coefficients T_k are X using Eq. (3) and Eq. (4). Find the basis α_k and the corresponding coefficients T_k.

Since the original image X is represented sparsely using a dictionary, it can be defined as an optimization problem for image decomposition from Eq. (5).

$$\{\alpha_0, \alpha_1\} = \arg min_{\{\alpha_0, \alpha_1\}} \|\alpha_0\|_0 + \|\alpha_1\|_0$$
$$\text{Subject to}: \mathrm{X} = \mathrm{T}_0 \alpha_0 + \mathrm{T}_1 \alpha_1 \tag{5}$$

By relaxing the l_0 norm to the l_1 norm and using the Lagrange undetermined multiplier method, an approximation can be made to Eq. (6).

$$\{\alpha_0, \alpha_1\} = \arg min_{\{\alpha_0, \alpha_1\}} \|\alpha_0\|_1 + \|\alpha_1\|_1$$
$$+\lambda \|X - \mathrm{T}_0 \alpha_0 - \mathrm{T}_1 \alpha_1\|_2^2 \tag{6}$$

This paper used LDCT and curvelet transform dictionaries to decompose the images. In the LDCT transform, the window interval size and the truncation percentage of the low-frequency component represented by each interval were changed. The number of layers was considered and investigated in the curvelet transform.

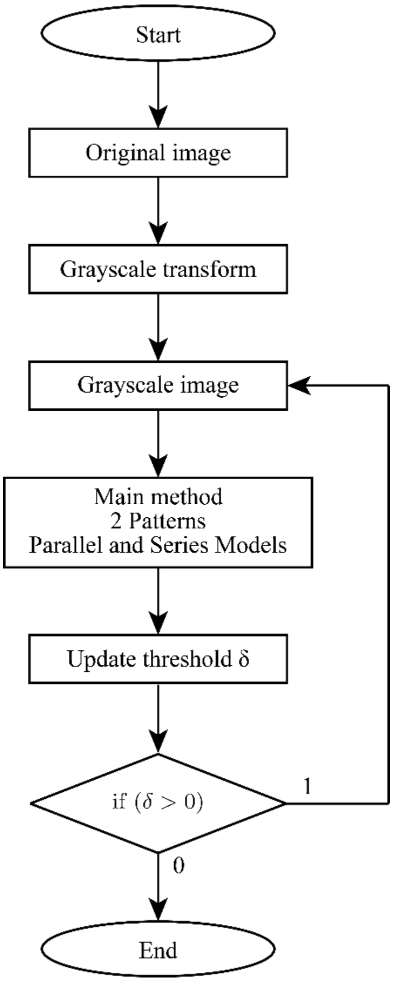
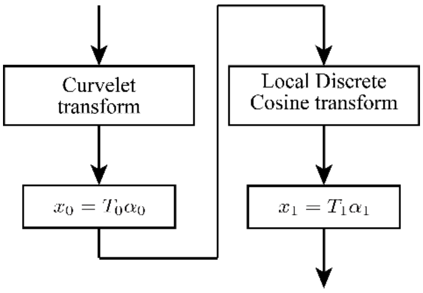
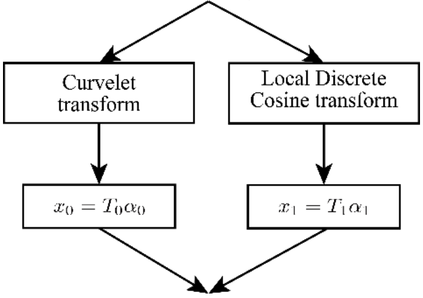


Fig.1. General flowchart of the system


Curvelet transform
Local Discrete Cosine transform
$x_0 = T_0\alpha_0$
$x_1 = T_1\alpha_1$

(a) Searial algorithm

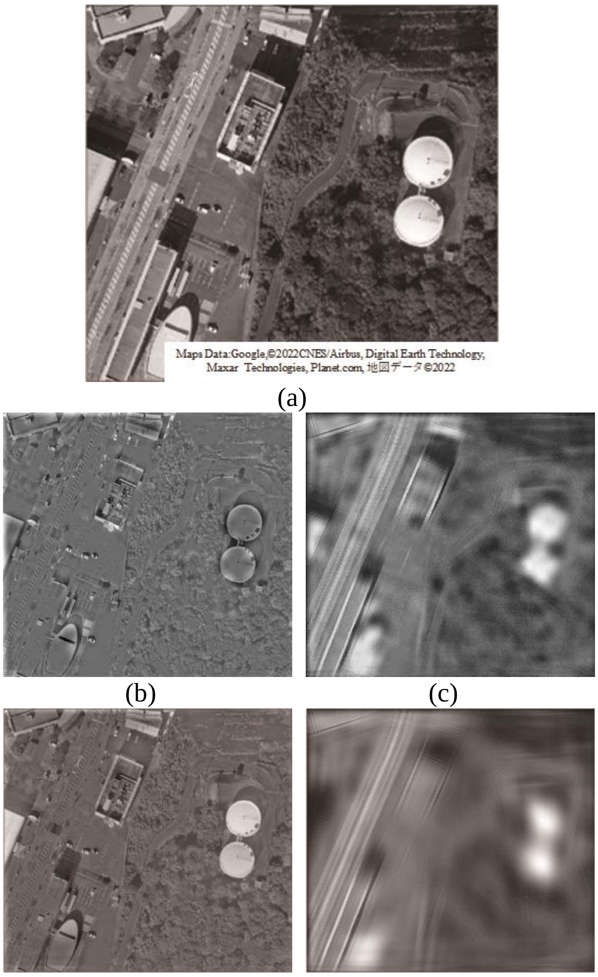

Curvelet transform
Local Discrete Cosine transform
$x_0 = T_0\alpha_0$
$x_1 = T_1\alpha_1$

(b) parallel algorithm
Fig.2. Flowchart of the system

In the first place, a flow diagram of the decomposition process is shown (Figure 1). As shown in the flow diagram, the method first performs a grayscale transformation on the color image. In the second place, each dictionary is used to obtain the coefficients, and the coefficients are repeatedly compared using a threshold value δ to perform decomposition while maintaining sparsity. In addition, the decomposition method was divided into a serial algorithm (Figure 2(a)) and a parallel algorithm (Figure 2(b)). The detection results were compared to verify the effectiveness of each method.

2.2. *Edge detection*

In the computer experiment, the result will be evaluated after the edge detection by filter processing using the Canny filter due to the clarification of how much the target structure can be extracted accurately [7]. The edge detection is characterized by comparing the luminance values between adjacent or nearby pixels to detect the outline of a structure. In this sense, it is important to detect the outline of complex structures such as forests, leaves, and other small components of the structure, making it difficult to detect the forest itself. Therefore, detecting a specific structure in multiple areas, such as forests, is not a simple problem for pattern matching scheme and learning scheme with teacher patterns. In the proposed system, fine and coarse components are decomposed, and the outline of only the structure is extracted by subtracting the fine components, which are tuned for the forest in the image.


Maps Data:Google,©2022CNES/Airbus, Digital Earth Technology,
Maxar Technologies, Planet.com, 地図データ©2022

(a)

(b) (c)

(d) (e)

Fig.3. Decomposed results by MCA with two dictionaries in algorithms as shown in Figure 2. Original image (a), LDCT image in the serial algorithm (b), CURVE image in the serial algorithm (c), LDCT image in the parallel algorithm (d), CURVE image in the parallel algorithm (e).

3. Results and Discussion

3.1. *Decompose aerial image*

The original image was converted to grayscale in Figure 3(a). The decomposition results are shown in Figure 3(b) and Figure 3(c) for a serial algorithm and in Figure 3(d) and Figure 3(e) for a parallel algorithm. The decomposition of texture and background was performed using the method shown in Method 2. The LDCT image (b) is the texture, and the Curvelet image (c) is the background (Figure 3(b) and Figure 3(c)). In this case, it can be seen that structures and the superficial structure of the forest are extracted in the LDCT image. On the contrary, no such features can be seen in the Curvelet component.

Figure 3(d) and Figure 3(e) show images decomposed by the parallel algorithm. As in Figure 3(d) and Figure 3(e), the LDCT image (d) is the texture, and the curvelet image (e) is the background.

It can be seen that the parallel algorithm extracts more structure from the texture than the serial algorithm.

This method also used "16" for the window interval used in the LDCT and 5 for the curvelet transform.

3.2. *Edge detection by using Canny filter*

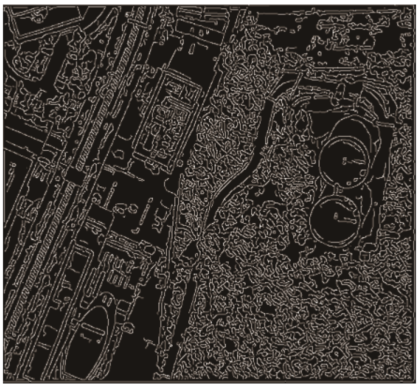
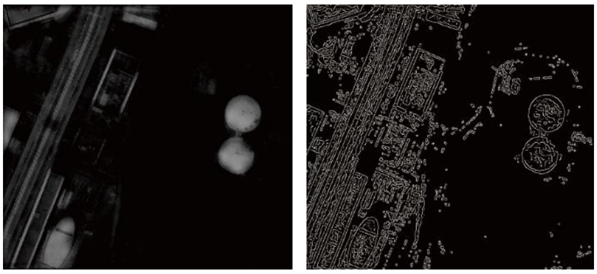


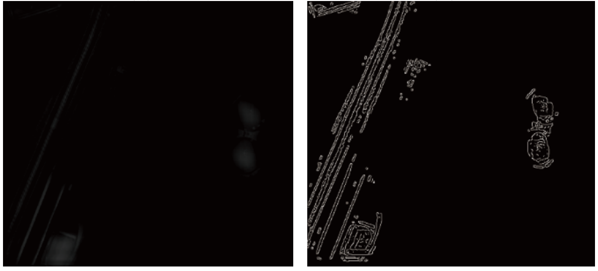
Fig.4. Edge detection from the original image (without MCA)

Figure 4 shows edge detection from the current image.

In this paper, we focus on the texture after decomposition and remove the forest area from the viewpoint of edge detection. Normally, forest areas in aerial photographs are composed of a sequence of detailed information on trees, so edge detection detects the edges of trees and leaves, making it difficult to obtain an outline of the area.



(a) (b)



(c) (d)

Fig.5. Edge detection results after MCA. Subtraction of LDCT image from the original (a) and its edge detection (b) in the serial algorithm, and Subtraction of LDCT image from the original (c) and its edge detection(d) in the parallel algorithm.

A decomposed image by MCA is used. Focusing on the decomposed image with the serial algorithm shown in Figure 5(a) and Figure 5(b), it can be seen that the decomposed image (texture) is subtracted from the original image, and edge detection is performed from Figure 5(b) to capture the outline of the structure without acquiring detailed information such as forests.

On the other hand, focusing on the decomposed image in the parallel algorithm shown in Figure 5(c) and Figure 5(d), the difference of the decomposed image (texture) is obtained from the original image, and edge detection is performed from Figure 5(d), which does not extract forest information but fails to detect other edges as well. In other words, it is necessary to determine the value and decomposition method of these features according to various situations.

4. Conclusion

This paper describes the possibility of removing the target objects by using dictionaries appropriately and focusing on the structural pattern extraction function. We believe this method has broad potential as an automated method for structural data extraction from complex images and can be applied to other fields.

Acknowledgements

This work was supported in part by JSPS KAKENHI (16H01616, 17H06383), Project on Regional Revitalization Through Advanced Robotics (Kyushu Institute of Technology/Kitakyushu city, Japan) and the New Energy and Industrial Technology Development Organization (NEDO).

References

1. S. Saito, Y. Aoki, "Building and road detection from large aerial imagery", Proceedings of SPIE, The International Society for Optical Engineering, February 2015.
2. M. Negri, P. Gamba, G. Lisini, Florence Tupin, "Junction-Aware Extraction and Regularization of Urban Road Networks in High-Resolution SAR Images", EEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, VOL. 44, NO. 10, OCTOBER 2006.
3. G. Koutaki, K. Uchimura, S. Wakisaka, "Road Extraction from Aerial Imagery based on Crossing Detection", IPSJ SIG Technical Report.
4. M. Elad, J.-L. Starck, P. Querre, D.L. Donoho, "Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA)", Applied and Computational Harmonic Analysis,Volume 19, Issue 3, November 2005, Page 340-358.
5. J.-L. Starck, Y. Moudden, J. Bobin, M. Elad, D.L. Donoho, "Morphological Component Analysis", Proceedings of SPIE – The International Society for Optical Engineering, August 2005.
6. E. Candes, L. Demanet, D. Donoho and L. Ying, "Fast Discrete Curvelet Transforms", July 2005, revised March 2006.
7. W. Rong, Z. Li, W. Zhang and L. Sun, "An Improved Canny Edge Detection Algorithm", Proceedings of 2014 IEEE, International Conference on Mechatronics and Automation, August 3 - 6, Tianjin, China.

Authors Introduction

Mr. Wataru Oshiumi

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's student at Kyushu Institute of Technology, Japan.

Mr. Ankur Dixit

He received his Master's degree in Electronics & Communication Engineering in 2014 from National Institute of Technology (NIT), Kurukshetra, Haryana, India. He worked in the Mangalayatan University as an assistant professor in 2014-2017. He enrolled at the PhD course in Kyushu Institute of Technology, Japan from 2017 and he is receiving the PhD supervision.

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.

3D Point Cloud Registration and Segmentation of Reflective Metal Objects Using Go-ICP and Improved RANSAC

Kairi Morita, Ziyue Liu, Jing Cao
Graduate School of Engineering, Kyushu Institute of Technology,
1-1, Sensui-Cho, Tobata-ku, Kitakyushu 804-8550, Japan

Seiji Ishikawa, Masuhiro Nitta, Joo Kooi Tan
Faculty of Engineering, Kyushu Institute of Technology
1-1, Sensui-Cho, Tobata-ku, Kitakyushu 804-8550, Japan
E-mail: etheltan@cntl.kyutech.ac.jp

Abstract

Registration and segmentation of 3D data are necessary in many fields, such as factory automation, automated driving, or even in the medical field. However, the technique is generally applied to non-metal objects. One of the problems of registration of a metal object is that the point clouds representing a metal object contain many outliers and missing points because of its reflective nature. This makes the accuracy of the registration and segmentation degrade. In this paper, we propose registration and segmentation techniques that are robust to outliers. For registration, we use the globally optimal Go-ICP (Global optimal - Iterative Closest Point) algorithm considering the goodness of a combination of point cloud sets to escape from convergence to a local solution. In segmentation, we address the problem of RANSAC generating false segments consisting of nearly identical multiple planar points and introduce an improved RANSAC. We use three kinds of the metal tray to show the effectiveness of the proposed technique.

Keywords: Registration, Segmentation, Reflective Metal objects, Go-ICP, RANSAC

1. Introduction

Registration and segmentation are important techniques in various fields, such as factory automation, automated driving, and the medical field. However, no effective method has been proposed for objects such as metals, which are strongly affected by reflections. The problem with metallic objects is that reflections cause a lot of noise and missing points in the point cloud, which reduces the accuracy of registration. Another problem is that segmentation generates many spurious surfaces. In this study, we investigate effective registration and segmentation techniques for trays made of metal. In Chapter 2, we describe a proposed registration technique using a threshold value. In Chapter 3, we describe a proposed segmentation technique using normals and RANSAC, and in Chapter 4, we conduct experiments to verify the effectiveness of the proposed technique.

2. Registration

Let $X = \{x_i\}$ ($i = 1, \ldots, M$) and $Y = \{y_j\}$ ($j = 1, \ldots, N$) be two point groups. Here, $x_i, y_j \in \mathbb{R}^3$ are points in a 3-dimensional space. To perform 3-dimensional registration, we find a rotation matrix $\mathbf{R} \in SO(3)$ and a translation vector $\mathbf{t} \in \mathbb{R}^3$ that minimize the distance between corresponding points in the two point groups.

The evaluation function is given by

$$E(\mathbf{R}, \mathbf{t}) = \sum_{i=1}^{M} e_i(\mathbf{R}, \mathbf{t})^2 \qquad (1)$$

$$= \sum_{i=1}^{M} \left\| \mathbf{R}x_i + \mathbf{t} - y_{j^*} \right\|^2$$

where $e_i(\mathbf{R}, \mathbf{t})$ is the pointwise residual of x_i. The point $y_{j^*} \in Y$ is represented as the optimal correspondence of x_i. j^* is then

$$j^* = \underset{j\in\{1,\dots N\}}{\operatorname{argmin}} \left\| \mathbf{R}x_i + \mathbf{t} - y_j \right\| \qquad (2)$$

In this study, the Go-ICP algorithm is used to find the optimal $\mathbf{R}, \mathbf{t}$ that minimize the evaluation function $E(\mathbf{R}, \mathbf{t})$.

2.1. *Go-ICP algorithm*

The Go-ICP algorithm [1] is a combination of the branch-and-bound method and the ICP algorithm. It is a globally optimal ICP algorithm that uses the branch-and-bound method to reduce the possibility of convergence to a local solution.

2.2. *Proposed Method*

The proposed method adds a threshold to the Go-ICP algorithm to improve the accuracy of registration of multiple point clouds. three thresholds th_1, th_2, andth_3 are set, and the process is performed in five steps. The steps are shown below. Let p_k $(k = 1, \dots, K)$ be the point clouds. r_{before} is the distance between the point clouds before registration, and r_{after} is the distance between the point clouds after registration.

1. Align p_k and p_{k+1}
   If $k = 0$ and $r_{after_k} < th_1$, then $p_k$and p_{k+1} are integrated to make P_l $(l = 1)$.
   Also, let $k = k + 1$.
   If $k \geq 1$ and $r_{after_k} < th_1$, then integrate p_{k+1} into P_l. Also, assume $k = k + 1$.
   If $k \geq 1$ and $r_{after_k} > th_1$ then $l = k$, $k = k + 1$ and go to step 2.
2. Repeat 1 for all point groups. Multiple point clouds $P_l, P_{l+k}, \dots, P_L$ are created. If $l_i \neq l_{i+1} - 1$ and $l_i \neq l_{i-1} + 1$ for subscript l of P, go to Step 3.
3. Calculate the RMSE of P_l and P_{l+n} $(n = 1)$ before alignment.
   Align P_l and P_{l+n} if $r_{before_l} < th_2$. Go to Step 4.
   If $r_{before_l} > th_2$, do not align, $n = n + 1$ (until $n \leq 2$, $l = l + 1$ if $n > 2$) and $l = l + 1$, then go to Step 5.
4. Calculate RMSE after alignment.
   If $r_{after_l} < th_3$, merge $P_l$and P_{l+n}. $l = l + 1$ and go to step 5.
   If $r_{after_l} > th_3$, do not integrate, $n = n + 1$ (until $n \leq 3$, $l = l + 1$ if $n > 3$) and $l = l + 1$, then go to Step 5.
5. Repeat step 3.
6. Output the point cloud.

Since point clouds contain noise and outliers, they significantly impact the accuracy of registration. By using a threshold value, point clouds whose registration results are below the threshold value can be excluded, and, as a result, point clouds containing noise and outliers can be excluded. By performing step 3 and subsequent steps, this method can exclude point clouds that are above the threshold value and improve the accuracy of 3D reconstruction. Therefore, if the registration results are good, the process terminates at step 2, and the result is the same as the original Go-ICP.

3. Segmentation

The point cloud of the metal tray obtained by registration is segmented to determine its shape. For this purpose, the four sides and the bottom of the point cloud tray are segmented. After segmentation, the plane equation of the bottom surface is used to obtain the top surface of the tray, and the vertices of the tray are obtained from the plane equations of each plane. The dimensions of the point cloud tray are then determined by measuring the height, width, and depth of the tray from the vertices.

3.1. *RANSAC-based Segmentation*

RANSAC is used to obtain the planes of metal trays. However, metal objects are often noisy and deficient, and if RANSAC is applied to metal trays, the problem is estimating spurious planes, as shown in Figure 1. Therefore, we propose a method of dividing the point cloud by the normal vector of the surface and the normal vector of the point, and then performing RANSAC. This method can solve the problem of spurious surfaces.

The search radius is used to calculate the normals. Let $\boldsymbol{P}_i$ be a point and r be a radius. using a KD tree, points within radius r from the point of interest are used and principal component analysis is performed to obtain the normal vector $\boldsymbol{n}_i$. This is done for all points.

RANSAC consists of the following four steps:

1. Select 3 points at random from the input points.
2. Estimate the plane from the three points.
3. Set a threshold and measure the number of inliers.
4. Repeat steps 1-3 a specified number of times to find the plane with the maximum number of inliers.

The final decision plane $\widehat{M}$ can be expressed by the following equation [2];

$$\hat{M} = \underset{M}{\operatorname{argmax}} \left\{ \sum_{P_i \in U} T(P_i, M) \right\} \qquad (3)$$

where U is the set of remaining points, P_i is a point, M is a plane, and $T(P_i, M)$ is an inlier index defined by

$$T(P_i, M) = \begin{cases} 1 & L_i < L_t \\ 0 & otherwise \end{cases} \qquad (4)$$

Here L_i is the point-to-surface distance and L_t is the corresponding threshold value.

3.2. *Proposed Method*

In the proposed method, initially, only a single plane of the metal tray is estimated using RANSAC. Then, normal vector $\vec{n} = (\mathrm{a}, \mathrm{b}, \mathrm{c})$ of the estimated plane is obtained. Similarly, normal vector $\overrightarrow{n_{P_i}} = (a_i, b_i, c_i)$ is found for each point in the point group, the point-plane angle θ is calculated, and the point group is divided into three parts as follows;

$$S_k(P_i) = \begin{cases} S_1(P_i) & \theta_i < \theta_1 \\ S_2(P_i) & \theta_1 < \theta_i < \theta_2 \\ S_3(P_i) & \theta_2 < \theta_i \end{cases} \qquad (5)$$

where $S_k(P_i)$is the set of segmented points and $\theta_k(\mathrm{k} = 1,2,3)$ is the threshold value.

The threshold value θ_i is calculated by

$$\theta_i = \frac{(a \cdot a_i + b \cdot b_i + c \cdot c_i)}{\sqrt{a^2 + b^2 + c^2} \cdot \sqrt{{a_i}^2 + {b_i}^2 + {c_i}^2}} \times \frac{360°}{\pi} \qquad (6)$$

For each point cloud that is divided, all planes of the metal tray can be estimated using RANSAC. The top surface of the tray is then determined by varying the intercept of the plane equation for the bottom surface.

$$T(P_i, M) = \begin{cases} 1 & a_i x + b_i y + c_i z + d' > 0 \\ 0 & otherwise \end{cases} \qquad (7)$$

The top surface is determined by varying d' until the total number of points counted satisfies $T(P_i, M) < T_{th}$. Here, T_{th} is the threshold value for the number of points counted. Once the required sides, bottom and top surfaces are estimated, the 8 vertices are determined and the tray's height, depth and width are calculated.

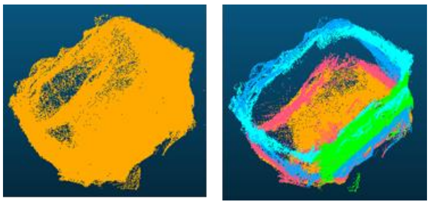

(a) (b)

Figure 1. Examples of spurious surface detection. (a) Point cloud of a metal tray viewed from an angle; (b) results of applying RANSAC five times to (a) and coloring the estimated planes; the bottom of the tray (orange) and one side (green) are estimated, but the rest planes are estimated as spurious planes.

4. Experiment

4.1. *Multiple Registration*

In this section, we register multiple point clouds of metal trays from different viewpoints. We compare the registration results of the proposed method with those of the original Go-ICP to demonstrate the effectiveness of the proposed method. The three types of trays used in the experiments are shown in Figure 2. The point cloud dataset was obtained by photographing the metal trays rotated on a rotating table with a fixed realsenseD455 (Figure 3). Only the metal tray is extracted by background differencing, and the point cloud of the tray is generated from the resulting depth image. The detailed conditions of the experiment are listed in Table 1. Stainless steel (SUS304) has noise, outliers, and defects most, so registration is performed over a 360° range. The registration results using the conventional and proposed methods are shown in Figure 4 and Figure 5, respectively. The registration results of the conventional method differ from those of the proposed method only for the stainless steel (SUS304) tray, indicating that the accuracy of 3D reconstruction is improved by the threshold value. The results for the aluminum tray and the stainless steel (SUS340) tray were the same as those of the conventional method because they did not proceed beyond step 3 of the proposed method.

4.2. *Segmentation*

In this section, we segment the point cloud after registration, measure the dimensions, and determine the difference from the original dimensions. The experimental results show a point cloud with only the trays enlarged. The result of segmenting the point cloud into three parts using the proposed method is shown in Figure 6. The pink point cloud is the first plane estimated, and it was divided into three point clouds, pink, orange, and green, based on the angle between the normal vector of the pink plane and the normal vector of each point. Applying RANSAC to each of the three divided point groups, all sides of the tray can be estimated as shown in Figure 6. RANSAC was applied multiple times for each point group, and the one that estimated the sides correctly was manually selected. The dimensions are measured from the planes estimated in Figure 7 as shown in Figure 8, and as given in Table 2,3, and 4.

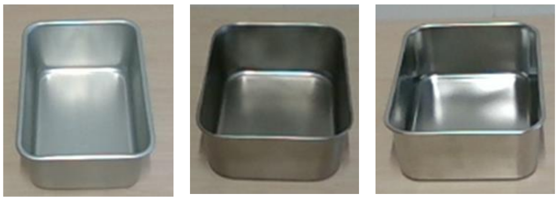

(a) (b) (c)

Figure 2. Metal trays for experiments: (a) Aluminum tray; Width: 156[mm], Depth: 227[mm], Height: 85[mm] ; (b) Stainless steel (SUS340) tray; Width: 169[mm], Depth: 227[mm], Height: 78[mm] ; (c) Stainless steel (SUS304) tray; Width: 170[mm], Depth: 225[mm], Height: 78[mm]

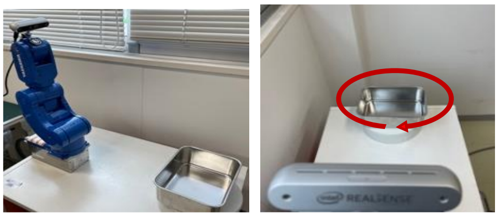
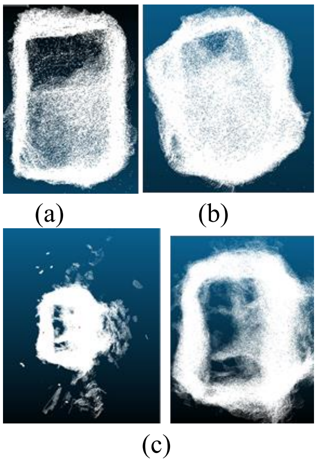


Figure 3. Experimental environment: (Left) A camera fixed on the robot and the metal tray on the rotating stand; (Right) a metal tray viewed from the camera; the metal tray is rotated clockwise and photographed.

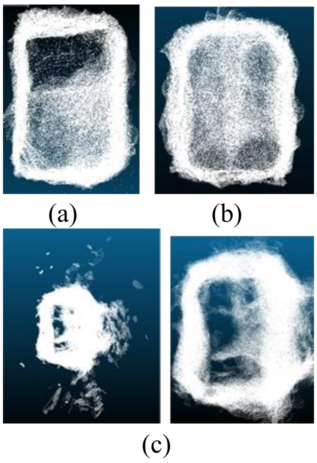
Table 1. Registration details for each point cloud

Metal Types	Aluminum	Stainless steel (SUS340)	Stainless steel (SUS304)
Registration range(°)	180	180	360
Point cloud set	28	28	110
Parallax of 1 set (°)	6.42	6.42	3.27


(a) (b)

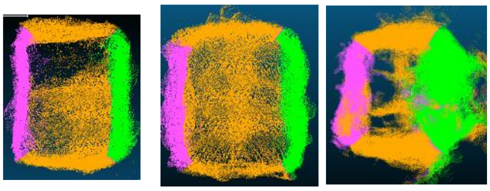
(c)

Figure 4. Registration results of conventional methods: (a) Aluminum tray; (b) Stainless steel (SUS340) tray; (c) Stainless steel (SUS304) tray: (Left) original result; (Right) enlarged tray.


(a) (b)

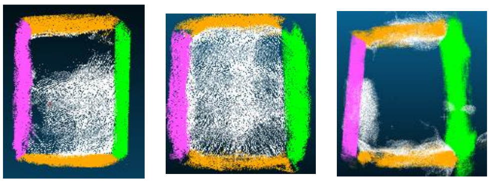
(c)

Figure 5. Registration results of the proposed method: (a) Aluminum tray, (b) stainless steel (SUS340) tray; (c) stainless steel (SUS304) tray: (Left) original result; (Right) enlarged tray.



(a) (b) (c)

Figure 6. Result of dividing the point cloud into three regions based on the normal vector of the pink side and the angle of each point: (a) aluminum tray;



(a) (b) (c)

Figure 7. Results of applying RANSAC to each point cloud in the three regions: (a) Aluminum tray; (b) Stainless steel (SUS340) tray; (c) Stainless steel

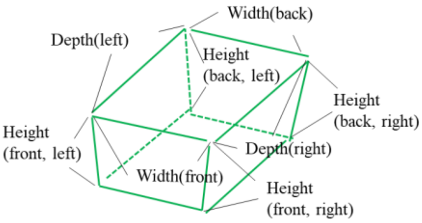


Figure 8. Dimensional measurement

Table 2. The result of dimensional measurement with an aluminum tray

	Actual size	Measurement Results		Error	
Width[mm]	156	143(front)	139(back)	-13	-17
Depth[mm]	227	218(left)	222(right)	-9	-5
Height(right)[mm]	85	97(front)	95(back)	+12	+10
Height(Left)[mm]	85	99(front)	98(back)	+14	+13

Table 3. The result of dimensional measurement with a stainless steel (SUS340) tray

	Actual size	Measurement Results		Error	
Width[mm]	169	165(front)	169(back)	-4	±0
Depth[mm]	227	214(left)	232(right)	-13	+5
Height(right)[mm]	78	81(front)	83(back)	+3	+5
Height(Left)[mm]	78	82(front)	84(back)	+4	+6

Table 4. The result of dimensional measurement with a stainless steel (SUS304) tray

	Actual size	Measurement Results		Error	
Width[mm]	170	180(front)	146(back)	+10	-24
Depth[mm]	225	208(left)	236(right)	-17	+11
Height(right)[mm]	78	101(front)	100(back)	+23	+22
Height(Left)[mm]	78	98(front)	97(back)	+20	+19

5. Conclusion

In this study, we investigated effective registration and segmentation methods for metal trays with high noise, outliers, and missing values. A threshold was introduced for registration of point clouds of metal trays viewed from multiple directions. The normal information and RANSAC were used to segment the sides and bottom of the trays.

Several improvements can be considered for future work. In registration, the threshold is set manually, and, in segmentation, the correct plane is selected manually. Therefore, it is thought that automating the process will increase its versatility. In this study, we applied RANSAC multiple times and manually selected the one that correctly estimated the sides. It is, however, possible to automatically estimate the sides and bottom by using the normal vector of the plane in which the tray is placed. Then, variants of RANSAC could be used for plane estimation.

References

[1] J. Yang, H. Li, D. Campbell, and Y. Jia, “Go-ICP: A globally optimal solution to 3D ICP pointset registration,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 38(11):2241–2254, Nov. 2016.

[2] B. Xu, W. Jiang, J. Shan, J. Zhang, and L. Li, “Investigation on the weighted RANSAC approaches for building roof plane segmentation from lidar point clouds,” Remote Sensing, vol. 8, no. 1, p. 5, 2015.

Authors Introduction

Mr. Kairi Morita

He received his B.E. degree in Department of Control Engineering in 2021 from the Faculty of Engineering, Kyushu Institute of Technology in Japan. He is acquiring the M.E. in the same University. His research interests are image processing, robot vision and robotics.

Mr. Ziyue Liu

He received B.E. degree in Mechanical Design, Manufacturing and Automation in 2014 from the School of Mechanical and Electrical Engineering, University of Electronic Science and Technology of China. He received M.E. degree in Control Engineering in 2020 from Faculty of Engineering, Kyushu Institute of Technology,Japan. His research interests are image processing, pattern recognition, and robotics.

Ms. Jing Cao

She received her 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 includes computer vision, machine learning and motion recognition.

Prof. Dr. Seiji Ishikawa

He graduated from The University of Tokyo 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 from Biomedical & Fuzzy System Association, Japan. He is a member of IEEE.

Prof. Dr. Masuhiro Nitta

He received his master's and doctoral degrees from Nara Institute of Science and Technology in 2004 and 2007, respectively. He was an assistant professor at Tokyo University of Science from 2007 to 2012. He is an assistant professor at Kyushu Institute of Technology since 2012. He received a best paper award from SICE Japan in 2007. His research interests are blind signal processing and linear systems analysis.

Prof. Dr. Joo Kooi Tan

She is currently with the Department of Mechanical and Control Engineering, Kyushu Institute of Technology, as a Professor. 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, and The Biomedical Fuzzy Systems Association of Japan.

Development of a Safe Walking Assistance System for Visually Impaired Persons Using MY VISION — Estimation of a Safe Passage from Sidewalk Information Based on Transfer Learning of VGG-16 Network

Takumi Yokote
Graduate School of Engineering, Kyushu Institute of Technology,
Kitakyushu, Fukuoka 804-8550, Japan
E-mail: yokote.takumi165@mail.kyutech.jp

Joo Kooi Tan
Faculty of Engineering, Kyushu Institute of Technology
Sensui-cho Tobata-ku
Kitakyushu, 804-8550, Japan
E-mail: etheltan@cntl.kyutech.ac.jp

Abstract

In recent years, the number of visually impaired persons has been increasing year by year, and outdoor accidents have also been increasing when they go out. It is difficult to detect hazards on sidewalks even with a currently popular technique, such as a semantic segmentation technique or YOLO, because sidewalk situations are complicated and change frequently. For this reason, we propose a method of recognizing sidewalk situations from a self-viewpoint video called MY VISION. Conventional methods detect objects surrounding the sidewalk by learning the objects' features beforehand and guiding visually impaired persons according to the position/direction of the detected object. The proposed method neither learns objects nor detects objects. We focus on sidewalk situations and use a multi-class classification technique based on transfer learning of VGG-16 to guide visually impaired persons' walk according to three kinds of sidewalk information to ensure more safety. The effectiveness of the proposed method was confirmed by experiments.

Keywords: Safe walking assistance, deep learning, visually impaired, MY VISION.

1. Introduction

According to a survey conducted by the Ministry of Health, Labour and Welfare in 2008, the number of visually impaired persons was 312,000 [1]. Regarding accidents involving visually impaired people on the street, a 2003 survey showed that approximately 42% of all respondents had experienced a walking accident involving injury while walking outdoors within past five years [2]. This indicates that walking outdoors is very dangerous for the visually impaired.

Currently, white canes, Braille blocks, and guide dogs are widely used to help the visually impaired recognize their surroundings. However, each of these methods has its own problems. This indicates that there is a need for a method that allows visually impaired people to explore their surroundings easily and extensively.

Therefore, we propose a method to recognize a surrounding environment from MY VISION by multi-class classification using deep learning. The MY VISION in this paper is based on a chest-mounted camera image. When the proposed method is completed, the visually impaired person will be able to know danger earlier. This method also guides the user to the center of the sidewalk, thus enabling the user to pass through a path with fewer hazards. The above is the significance of this research.

Previous studies include obstacle detection by object and area recognition using RGB-D sensors [3], and methods to recognize people and obstacles using YOLO and Faster-RCNN, that are excellent algorithms for object detection [4][5].

However, even if these methods produce objects and walkable areas as one task, instructions for a visually impaired person to avoid obstacles must be learned and processed as a separate task.

In this paper, by using instruction-based multiclass classification to understand the situation on the sidewalk, the learning and processing from the input to the instructions for the visually impaired person to walk can be performed as a single task. Another unique aspect of this research is that it uses images of the left and right edges of the sidewalk as well as the center of the sidewalk as a dataset, taking into account the actual situation in which a visually impaired person is walking.

2. Method

2.1 *Deep Learning model*

2.1.1 *Transfer learning*

In order to create a deep learning model for this method, we perform fine-tuning of the VGG16 model [6].

2.1.2 *VGG16 model*

VGG16 is the second-ranked model in ILSVRC in 2014. It uses ImageNet as its training dataset and classifies 1000 categories.

This method uses a model in which the output layer of this model is modified. The model is shown in Figure 1.

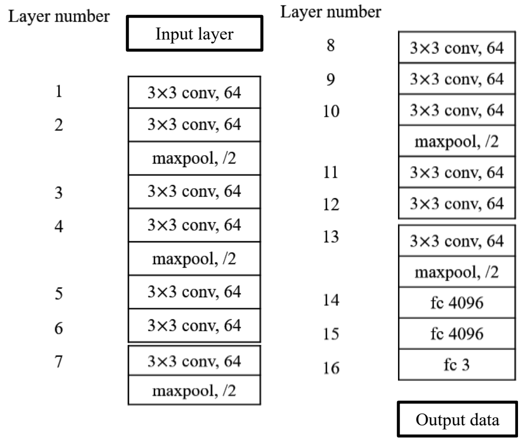


Fig.1 Structure of the model for gait instruction classification

2.2 Learning

2.2.1 *Loss function*

In this method, cross entropy is used as a loss function. Cross-entropy is a loss function for class classification, where n is the sample number, K is the number of classes, y_{nk} is the output data of the n th sample class k expressed as a probability, and t_{nk} is the teacher data of the n th sample class k expressed as 0 or 1, the error of each data is given by the following formula.

$$E_n = -\sum_{k=1}^{K} t_{nk} \log y_{nk} \tag{1}$$

When the total number of samples is N, the overall error is

$$E = \sum_{n=1}^{N} E_n \tag{2}$$

The more the output data and the teacher data match, the smaller the value.

2.2.2 Parameter optimization methods

In this study, Adam is used as a parameter optimization method. Adam's method is described below. First, let $m_{i,t}$ be the exponential moving average of the first-order moment of the gradient and $v_{i,t}$ be the exponential moving average of the second-order moment of the gradient, defined as follows.

$$m_{i,t} = \rho_1 m_{i,t-1} + (1-\rho_1)\nabla E(\boldsymbol{w}^{(t)})_i \tag{3}$$

$$v_{i,t} = \rho_2 v_{i,t-1} + (1-\rho_2)(\nabla E(\boldsymbol{w}^{(t)})_i)^2 \tag{4}$$

where the initial values are $m_{i,t} = v_{i,t} = 0$. In addition, ρ_1 and ρ_2 are decay rates, and $\nabla E(\boldsymbol{w}^{(t)})_i$ is the i th component of the gradient. However, since this equation takes an initial value of 0, the moment estimates are biased toward 0 at the beginning of the update. The moment estimates corrected for this bias are shown in the following equation.

$$\hat{m}_{i,t} = \frac{m_{i,t}}{(1-(\rho_1)^t)} \tag{5}$$

$$\hat{v}_{i,t} = \frac{v_{i,t}}{(1-(\rho_2)^t)} \tag{6}$$

The gradient descent method uses these estimates, and the equation becomes Eq. (7).

$$\Delta \boldsymbol{w_i}^{(t)} = -\eta \frac{\hat{m}_{i,t}}{\sqrt{\hat{v}_{i,t} + \epsilon}} \tag{7}$$

Here ϵ is a very small number that is introduced so that the denominator is not zero.

2.3 Multiclass classification

In this method, the gait instruction classification is multi-class classification into three classes. Multi-class classification is the process of sorting data into several categories. The following is the description of the training dataset used in this study.

The training dataset used in this study consists of 15 sidewalks in different locations, and at each location, three images were taken: an image of a person walking on the left edge of the sidewalk, an image of a person walking on the right edge of the sidewalk, and an image of a person walking in the center of the sidewalk. The images were then classified into three classes: "Go", "Avoid to the right" and "Avoid to the left.

The following is an explanation of the situations and obstacles assumed for each class. First, the obstacles are people and bicycles within 5 meters in front of the vehicle. The "Go" class is for situations where there are no obstacles on either side, such as in the center of a sidewalk, and there are no obstacles in front of you to stop. The "Avoid to the right" class is for situations where there is an obstacle on the left, such as when walking on the left side of the sidewalk, or when there is an obstacle in front of you and the sidewalk is visible to the right of the obstacle, making it safe. Finally, the "Avoid to the left" class is for the cases where there is an obstacle to the right, when walking on the right side of the sidewalk, or when the obstacle is in front of you and the sidewalk is visible and safe to the left of the obstacle.

3. Experiments and Results

3.1 *Experiments*

The number of datasets are shown in Table 1, and the wearers are shown in Figure 2. Image size is 224 x 224[pixel]. Precision was used as the evaluation function.

Table1 Number of data sets

Dataset	Learning[frame]	Test[frame]
Go	1529	505
Left	1622	421
Right	1656	405
Total	4807	1331

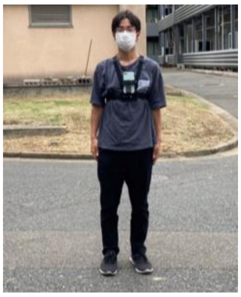


Fig.2 The wearer

3.2 *Results*

The accuracy for each class is shown in Table 2 shows how each correct answer label was estimated for each class.

Table2 Number of correct answers and Precision per correct answer label

	Go	Left	Right	Total
number of correct answers [frame]	494	276	306	1076
Precision	0.978	0.656	0.756	0.808

4 Discussion

First, Table 2 shows that the correct response rate is high for images with the correct label "Go", while the rate is low for images with the correct label "Avoid to the right" or "Avoid to the left. First, the high rate of "Go" is considered to be due to the fact that the basic "Go" label is applied when there are no obstacles on the left, right, or nearby, making it easier to make a decision. Second, the low rate of correct responses for the other two images may be due to the fact that, as shown in Figure 3, both the sidewalk and the roadway are shown in the image, and the roadway is misidentified as the sidewalk, and second, as shown in Figure 4, a passerby is in the center of the image, making it difficult to choose whether to avoid the right or left side of the road.

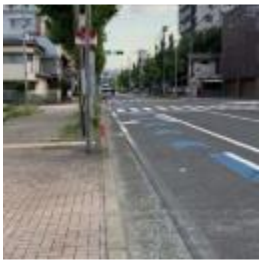


Fig.3 Example of mistaking a road for a sidewalk

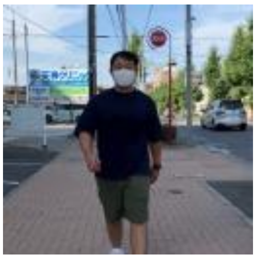


Fig.4 Examples that are difficult to discern which way to avoid

5 Conclusion

In this study, we proposed a walking instruction classification method based on multi-class classification using deep learning. The proposed method was trained on a dataset of labeled left-most, center, and right-most images taken at 15 locations using MY VISION. Using the model created by the training, the images were classified into three multi-class categories: "Go", "Avoid left" and "Avoid right".

Experiments were conducted to verify the effectiveness of the proposed method.

Test images taken at different locations from the training images were classified into multiple classes, and the percentage of correct answers and the breakdown of the estimated labels for the test images were examined. The results showed that the correct answer rate for the label "Go" was 97.8[%], the correct answer rate for the label "Avoid to the left" was 65.6[%], the correct answer rate for the label "Avoid to the right" was 75.6[%], and the average of the correct answers for the three labels was 80.8[%].

Future issues include clarification of the classification criteria and the distinction between people and objects. In order to clarify the classification criteria, it is necessary to create a small region in the image and judge whether or not there is an object in the region. In addition, to distinguish between people and objects, it is necessary to take video images and judge people based on the motion between two adjacent frames.

References

1. Ministry of Health, Labour and Welfare, Social and Support Bureau, Disability Health and Welfare Department: "Summary of the Results of the 2016 Survey on Living Conditions, etc. (National Survey of Children and Persons with Disabilities at Home)," pp1-14, 2018. (In Japanese)
2. N. Abe and N. Hashimoto: “Walking accidents National survey to Maintain Visually Handicapped Persons Walking Environment”, pp.81-92, 2005. (In Japanese)
3. A. Aladren, G. Lopez-Nicolas, L. Puig, J. J. Guerrero: “Navigation Assistance for the Visually Impaired Using RGB-D Sensor with Range Expansion”, IEEE Systems Journal, pp922-932, 2016.
4. O. Masurekar, O. Jadhav, P. Kulkarni, S. Patil: “Real Time Object Detection Using YOLOv3”, International Research Journal of Engineering and Technology (IRJET), pp3764-3768, 2020.
5. S. Ren, K. He, R. Girshick, J. Sun: “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks”, IEEE Transactions on Pattern Analysis and Machine Intelligence, pp.1137-1149, 2016.
6. K. Simonyan, A. Zisserman: ”Very Deep Convolutional Networks for Large-Scale Image Recognition”, International Conference on Learning Representations, 2015.

Authors Introduction

Mr. Takumi Yokote

He received his B.E. degree in Department of Control Engineering in 2017 from the Faculty of Engineering, Kyushu Institute of Technology in Japan. He is acquiring the M.E. in the same University. His research interests are development of a safe walking assistance system for visually impaired persons using MY VISION, estimation of a safe passage transfer learning, and neural network.

Prof. Dr. Joo Kooi Tan

Prof. Tan is currently with the Department of Mechanical and Control Engineering, Kyushu Institute of Technology, as a Professor. 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, and The Biomedical Fuzzy Systems Association of Japan.

Detection of Fallen Persons and Person Shadows from Drone Images

Taisei Ono, Haruka Egawa, Yuta Ono
Graduate School of Engineering, Kyushu Institute of Technology
Sensuicho1-1, Tobata, Kitakyushu, Fukuoka, 804-8550
E-Mail: ono.taisei774@mail.kyutech.jp, Japan

Seiji Ishikawa*, Joo Kooi Tan**
*Emeritus Professor, ***Faculty of Engineering, Kyushu Institute of Technology*
1-1 Sensui-cho, Tobata, Kitakyushu, Fukuoka, 804-8550, Japan
E-Mail: etheltan@cntl.kyutech.jp, Japan
https://www.kyutech.ac.jp/~etheltan/

Abstract

In recent years, the development of automatic search methods based on aerial images taken by drones has been attracting attention in order to prevent secondary disasters and to perform rescue operations quickly in the search for victims of natural disasters. Although various methods exist for automatic person detection for search, they are based on the assumption that the background area of a person captured by a drone camera is a uniform ground in which only those persons who require rescue exist without any shadows or trees. In this paper, we propose a method of automatic detection of both fallen persons and person shadows, or trees on the ground. The method is combination of Ri-HOG and Ri-LBP features to search for fallen persons. These features are robust to rotation. We then employ GrabCut and brightness values to detect shadows. The effectiveness of the proposed method was verified by experiment.

Keywords: Fallen person, Shadow of person, Drone images, Value histogram, GrabCut, Random forest.

1. Introduction

In the event of a disaster, the survival rate of victim declines rapidly with the passage of time, so rapid search activities are necessary. In addition, because of the risk of secondary disasters during disaster search activities, "quick" and "safe" search activities using drones have been attracting attention. As for person search, there is a person detection method using a heat source[1], but it is difficult to distinguish whether the person is standing or lying down. In addition, the fallen person detection method[2] and the fallen person head detection method[3], which focus on the shape of the person, incorrectly detect the shadow of an upright person whose shape is similar to that of a fallen person. Therefore, in this paper, we propose an automatic detection method for fallen persons and person shadows.

We define a shadow of a standing person outdoors as a person shadow, and propose a method of detecting person shadows from images captured by a camera mounted on a drone. Previous studies on shadow detection used RGB color information[4]. Since shadow colors generally have low brightness, we believe that the use of brightness is effective for shadow detection. Another study[5] used saturation as a feature for shadow detection, but saturation is easily affected by sunlight outdoors, so the color information changes significantly depending on light exposure. Other shadow detection methods include texture-based shadow detection[6,7], but these methods assume that texture exists in the shadows of all scenes in a video, and texture detection is difficult in aerial video taken with a drone due to resolution issues.

In this study, we propose a method of detecting fallen persons and person shadows using brightness as a feature. By discriminating between fallen persons and person shadows, it is possible to detect standing persons in addition to persons lying on the ground.

The proposed method uses the Rotation-invariant Histogram of Oriented Gradients (hereinafter referred to as Ri-HOG) feature[8] and the Rotation-invariant LBP (hereinafter referred to as Ri-LBP) feature[9] obtained from the circular cells to detect the shape of fallen persons and that of person shadows. After extracting the foreground by GrabCut[10] from the detected window, the brightness features are extracted to discriminate between a fallen person and a person shadow figure. Random Forest[11] is used as a discriminator for detection. There are two types of discriminators; a discriminator for person/shadow candidates that uses the images of a person falling down (a positive class) and the ground images (a negative class), and a discriminator for person/shadow candidates that uses the images of a person falling down after the GrabCut (a positive class) and the images of a person shadow after the GrabCut (a negative class). These detectors use brightness as a feature value.

2. Proposed Method

2.1. *Detection of fallen person and person shadow candidates*

In the proposed method, Ri-HOG[8] and Ri-LBP features are used for person and person shadow candidates detection. Ri-HOG features are HOG[12] features with rotation-invariance. Unlike HOG that uses rectangular cell, Ri-HOG uses cell arrangement that divides concentric circles. Ri-LBP[9] features are features that adds rotation invariance to Local Binary Pattern (LBP).

2.2. *Foreground region extraction*

GrabCut is used to extract foreground regions. In this section, we describe GraphCuts first and then GrabCut. Graph Cuts is a region segmentation method before applying GrabCut.

2.2.1. *Graph Cuts*

GraphCuts[13] is a method to efficiently find the combination of labels with the minimum cost, given a cost function that gives the minimum value when the labels are appropriately labeled. The cost function $E(L)$ used in Graph Cuts is shown below.

$$E(L) = R(L) + \lambda \cdot B(L) \tag{1}$$

$$R(L) = \sum_{p \in P} R_p(L_p) \tag{2}$$

$$B(L) = \sum_{\{p,q\} \in N} B_{\{p,q\}} \cdot \delta(L_p, L_q),$$

$$\delta(L_p, L_q) = \begin{cases} 1 & if\, L_p \neq L_q \\ 0 & otherwise \end{cases} \tag{3a}$$

$$B_{\{p,q\}} = \exp\left(-\frac{(I_p - I_q)^2}{2\sigma^2}\right) \cdot \frac{1}{\mathrm{dist}(p,q)} \tag{3b}$$

where P is the total set of pixels and p is its element. λ is a non-negative weighting of the relative importance of the domain term $R(L)$ and the boundary term $B(L)$.

The domain term $R(L)$ is a function that depends only on the labels given to the pixels, and is the sum of all pixels when a label is given to pixel p. It is computed from a region where the locations of the target and part of the background are previously taught. The boundary term $B(L)$ is a sum over all neighboring pixels regarding the continuity of neighboring pixels p and q, due to a priori knowledge of the circumstances under which the labels of neighboring pixels should be given and $B_{\{p,q\}}$is a cost function, where pixels are adjacent to each other, $\mathrm{dist}(p,q)$ is the distance between pixels p and q. When pixels p and q have the same label, the cost is set to zero by $\delta(L_p, L_q) = 0$. σ is a value that accounts for image noise.

Let $\Pr(L|D)$ be the likelihood of data D given L, and let $\Pr(L)$ and $\Pr(D)$ be the prior probabilities of L and D. Then, the Bayesian equation gives the following relationship.

$$\Pr(L|D) = \frac{\Pr(L|D)\Pr(L)}{\Pr(D)} \tag{4}$$

Since $\Pr(D)$ is a fixed value, the problem is to find L that maximizes the right-hand side. The cost-minimization problem is then solved by taking the logarithm of the numerator of the right-hand side and multiplying it by -1 to obtain E(L). Compared with equation (1), the following is obtained.

$$\mathrm{R}(L) = -\ln P(L|D) - \ln P(L) \tag{5}$$

The histograms $\Pr(I|L)$ and L= {"obj", "bkg"} are created from the specified target and the background regions, respectively. Using the probability density function, the domain terms are cost functions as follows;

$$R_p(\text{“obj”}) = -\ln\Pr(I_p/\text{“obj”}) \tag{6}$$
$$R_p(\text{“bgk”}) = -\ln\Pr(I_p/\text{“bkg”})$$

R_p(“obj”) is a function that doubles the probability density of pixel p when its label is the background and its pixel value, and the boundary term is then calculated by Eq(3b).

2.2.2. *GrabCut*

Based on the assumption that there is a difference between the pixel values of adjacent pixels of the target and the background in the input image, GrabCut[14] calculates the likelihood of the foregroundness and the backgroundness for each pixel in the specified rectangle based on color statistics using the GMM (Gaussian Mixture Model). Likelihood of the foregroundness and the backgroundness is calculated. Graph Cuts is then applied to the segmentation of the specified region. The GMM is relearned from the obtained results, and GrabCut is applied to it. This process is repeated to improve the segmentation accuracy.

2.3. *Extraction of V histogram features*

The Value histogram feature is a feature that excludes hue and saturation from the HSV (Hue, Saturation, Value: HSV) histogram. It is used to discriminate a fallen person and a person shadow.

2.3.1. *HSV Transformation*

First, the input RGB color image is converted from the RGB space to the HSV (Hue, Saturation, Value: HSV) space. The HSV transform is performed using the following equation. Here in case of $H < 0$, add 2π to H, and if $V\_max = 0$, $S = 0$ and $H = infinite$.

$$V_{max} = \max\{R, G, B\},\ V_{min} = \min\{R, G, B\} \qquad (7)$$

$$V = V_{max} \qquad (8)$$

$$S = (V_{max} - V_{min}) / V_{max} \qquad (9)$$

$$H = \begin{cases} \frac{G-B}{V_{max}-V_{min}}\frac{\pi}{3} & (\text{If } V_{max} = B) \\ \frac{B-R}{V_{max}-V_{min}}\frac{\pi}{3} + \frac{2\pi}{3} & (\text{If } V_{max} = R) \\ \frac{R-G}{V_{max}-V_{min}}\frac{\pi}{3} + \frac{4\pi}{3} & (\text{If } V_{max} = B) \end{cases} \qquad (10)$$

2.3.2. *Creating a value histogram*

The proposed method creates a two-dimensional histogram based on Value, excluding hue and saturation from the HSV histogram of a color image. The V histogram represents Value as a variable. The range of the Value is originally from 0 to 180, but it is changed to 45 levels from 0 to 44. and the intensity of the V histogram is normalized to 1 to speed up the computation.

2.4. *Detection of fallen persons and shadows*

The proposed method uses Random Forest as a discriminator, which is characterized by its ability to efficiently learn even high-dimensional features through random learning and its ability to suppress the influence of noise in the teacher signal by randomly selecting training data.

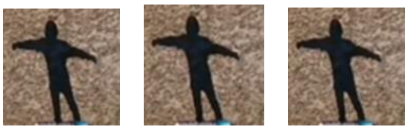


Fig. 1 Example images of the shadow of a person

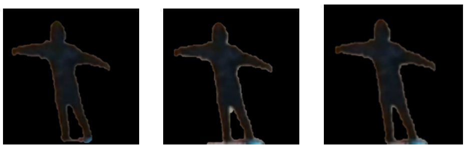


Fig. 2 Examples of a training image for the shadow of a person detector

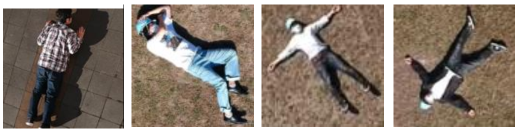


Fig. 3 Example images of a fallen person

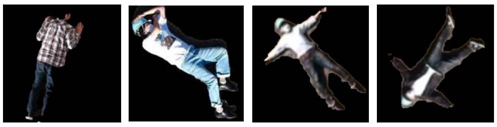


Fig. 4 Examples of a training image for a fallen person detector

First, we detect a candidate of a person who fell down based on a discriminator trained using Ri-HOG and Ri-LBP features. Next, the foreground extraction is performed by applying GrabCut to the detection window of the fell-down person candidate. Finally, the foreground image is used to discriminate between a fallen person and shadow of a person based on a discriminator trained using the value V of the extracted foreground image.

3. Experiment

In the performed experiments, images of a fallen person are used as positive class images in the image database for the detection of a candidate fallen person, and the images other than persons are used as negative class images. INRIA Person Dataset[15] was also used as the negative class image. A discriminator is constructed from these training images. All training images are 61[pixels] in height and 61[pixels] in width.

The discriminator between a fallen person and a human shadow using Value was constructed using images such as Figure 2, which is a GrabCut image of a human shadow image shown in Figure 1, as the Positive

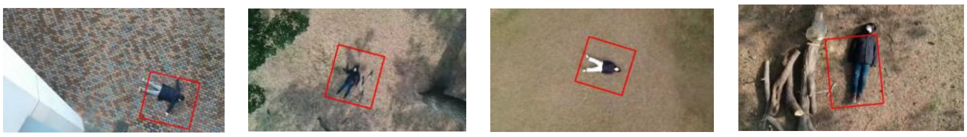


Fig. 5 Examples of detection results of a fallen person

class images in the image database, and images such as Figure 4, which is a GrabCut image of the a fallen person image shown in Figure 3, as the Negative class images. All training images are 500[pixels] in height and 500[pixels] in width.

3.1. *Experiment 1*

Four types of videos in which a person is captured by drone are used in the experiment. In all of these videos, there is only one fallen person to be detected.

The result of detection is evaluated with every detection window using Intersection over Union (IoU). The IoU threshold was experimentally set to 0.5. The experimental results showed that the detection rate was 0.80 for video 1, 0.71 for video 2, 0.78 for video 3, and 0.82 for video 4.

Examples of experimental results are shown in Figure 5. The red rectangle represents the detection result of a person who fell down. Figure 6 shows examples of the detection results of a fallen person after the GrabCut process.

3.2. *Experiment 2*

In this experiment, two types of videos in which a person shadow is captured by a drone are used. In all of these videos, there is only one person to be detected.

As in Experiment 1, detection is evaluated frame by frame using IoU. The IoU threshold is experimentally set to 0.5. The detection rate was 0.83 for video 5 and 0.95 for video 6.

Examples of the experimental results are shown in Figure 6. The blue rectangle represents the result of shadow detection after the GrabCut process. Examples of the experimental results is shown in Figure 7, where the proposed method is applied to the image before the Grab Cut. The blue rectangle represents the result of shadow detection. Figure 8 shows examples of the shadow detection after the GrabCut process.

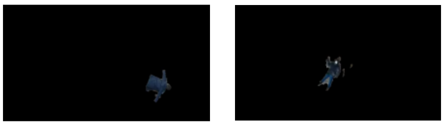


Fig. 6 Examples of detection results of a fallen person after the GrabCut process

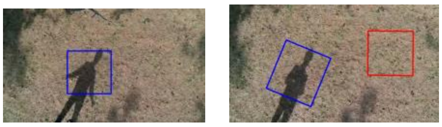


Fig. 7 Examples of a detected person shadow

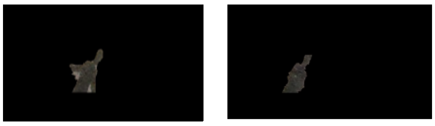


Fig. 8 Examples of a detected person shadow after the GrabCut process

4. Conclusion

In this paper, we proposed a method of discriminating between fallen persons and person shadows from aerial video images. Experimental results show that the average detection rate of a fallen person is about 78% when there is one fallen person in the frame, and the detection rate of a person shadow is about 89% when there is one person shadow in the frame. In the future, we aim at improving the detection rate in various cases, such as when the altitude is high.

References

1. A. Gaszczak, T. P. Breckona, J. Hana, “Real-time people and vehicle detection from UAV imagery”, Proceedings of the SPIE International Conference on Robots and Computer Vision XXVIII: Algorithms and Techniques, Vol.7878, pp.78780B-78780B-13, 2011.
2. H. Egawa, S. Ishikawa, J. K. Tan, “Detection of a fallen person from UAV images using rotation invariant features”, Proc. Biomedical Fuzzy Systems, pp.91-94, 2020. (In Japanese)
3. H. Egawa, S. Ishikawa, J. K. Tan, “Detection of a fallen person and estimation of the head position from UAV images”, The 2021 International Conference on Artificial Life and Robotics (ICAROB2021), pp.279-282, 2021.
4. P. Kumar, K. Sengupta, A. Lee, “A comparative study of different color spaces for foreground and shadow detection for traffic monitoring system”, Proceedings of the IEEE 5th International Conference on Intelligent Transportation Systems, pp.100-105, 2002.

5. V. Gomes, P. Barcellos, J. Scharcanski, "Stochastic shadow detection using a hypergraph partitioning approach", Pattern Recognition, Vol.63, pp.30-44, 2017.
6. A. Leone, C. Distante, F. Buccolieri, "A texture-based approach for shadow detection", Proc. IEEE Conference on Advanced Video and Signal Based Surveillance, pp.371-376, 2005.
7. A. Leone, C. Distante, "Shadow detection for moving objects based on texture analysis", Pattern Recognition, Vol.40, pp.1222-1233, 2007.
8. T. Fujimoto, Y. Kameda, I. Matsuda, S. Itoh, "Person detection from top view images using rotation invariant HOG features", Proc. The Institute of Image Information and Television Engineers, pp.17_10_1-17_10_2, 2014.
9. M. Pietikaine, T. Ojala, Z. Xu, "Rotation-invariant texture classification using feature distribution", Pattern Recognition, Vol.33, pp.43-52,2000.
10. D. Khattab, C. Theobalt, A. S. Hussein, M. F. Tolba, "Modified GrabCut for person face segmentation", Ain Shams Engineering Journal Volume 5, Issue 4, December 2014, pp.1083-1091.
11. L. Breiman, "Random forests", Machine Learning, Vol.45, No.1, pp.5-32, 2001.
12. N. Dalal, B. Triggs, "Histograms of oriented gradients for Person detection", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp.886-893,2005.
13. Y. Boykov, M-P. Jolly, "Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images", Proc. of IEEE International Conference on Computer Vision, I, pp.105-112, 2001.
14. C. Rother, V. Kolmogorov, A. Blake, "GrabCut": Interactive foreground extraction using iterated graph cuts",ACM Transacctions on Graphics, Vol, 23, No.3, pp.309-314, 2004.
15. INRIA Person Dataset, https://b.hatena.ne.jp/entry/pascal.inrialpes.fr/data/person

Authors Introduction

Mr. Taisei Ono

He received his B.E. degree in Department of Control Engineering in 2021 from the Faculty of Engineering, Kyushu Institute of Technology in Japan. He is acquiring the M.E. in the same University. His research interests are detection of fallen person, and shadow of person, image processing of drone images.

Ms. Haruka Egawa

She received B.E.and M.E. in Control Engineering from the Graduate School of Engineering, Kyushu Institute of Technology, Japan. Her research includes aerial images processing, human/fallen human detection and motion recognition.

Mr. Yuta Ono

He obtained the M.E. degrees in Department of Control Engineering from Kyushu Institute of Technology in 2017, Japan. He is presently a doctor course student in the same university. His research interests include feature extraction, machine learning, human detection and recognition.

Prof. Dr. Seiji. Ishikawa

Seiji Ishikawa graduated from The University of Tokyo 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 from Biomedical & Fuzzy System Association, Japan. He is a member of IEEE.

Prof. Dr. Joo Kooi Tan

Prof. Tan is currently with the Department of Mechanical and Control Engineering, Kyushu Institute of Technology, as a Professor. 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, and The Biomedical Fuzzy Systems Association of Japan.

A Systematic Literature Review on Emotion Recognition System In Malaysia

Muhammad Nadzree Mohd Yamin
Faculty of Business, Multimedia University, Jalan Ayer Keroh Lama, 75450 Bukit Beruang, Melaka

Kamarulzaman Ab. Aziz
Faculty of Business, Multimedia University, Jalan Ayer Keroh Lama, 75450 Bukit Beruang, Melaka Email: kamarulzaman.aziz@mmu.edu.my

Tan Gek Siang
Faculty of Business, Multimedia University, Jalan Ayer Keroh Lama, 75450 Bukit Beruang, Melaka

Nor Azlina Ab. Aziz
Faculty of Engineering and Technology, Multimedia University, Jalan Ayer Keroh Lama, 75450 Bukit Beruang, Melaka

Abstract

Artificial intelligence (AI) is an important technology that evolved from theories into tangibility with significant impacts and applications across sectors as well as borders. It is also one of the key technologies that gave rise to the fourth industrial revolution (IR 4.0). One key subcategory of AI is the automated emotion recognition system (ERS); the application of AI to recognize human emotional states. ERS is seen as an embedded technology that can be used in our daily lives and environment including the workplace. The importance of ERS will become more significant as we move towards the fifth industrial revolution (IR 5.0), where one of the key aspects identified is the enhancements in human-computer interaction (HCI). ERS has the potential to enable smart HCI, i.e. ERS can be seen as a technology to bridge us from IR 4.0 into IR 5.0. Crucial for this is good adoption or diffusion levels of ERS amongst society. Therefore, there is a need to understand the factors that affect the adoption of ERS. Specifically, this paper seeks to establish and discuss the current ERS research landscape in Malaysia by reporting findings from the systematic literature review covering works over a decade; from the year 2011 to 2022.

Keywords: Emotion recognition system (ERS), Artificial intelligence (AI), Affective computing (AC), The fourth industrial revolutionR1 (IR 4.0), The fifth industrial revolution (IR 5.0), Technology adoption

1. Introduction

Artificial intelligence (AI), is an advancement in computer science that allows a computer or machine to learn and be trained to perform and complete a task or make decisions, replicating how a human being learns from experience [1]. An emotion recognition system (ERS) is an AI that enables a system that accepts various modalities of data, learns and allows the machine to recognize the emotion of the subject. ERS is part of Affective Computing (AC), a field that was brought forward by [2]. According to Rosalind Picard, AC is a computer that relates to, arises from and understands human behaviour. With that being said, AC led to the findings of ERS, and it has been suggested by the previous researcher for ERS as embedded technology the current technology such as AI [1].

In the past decade, ERS has emerged as one of the attractive areas in the field of AC and AI with the vision of what the technology can achieve [3]. To develop a complex system such as robots that can interact and communicate with humans, a previous study has suggested a system with a function for both understanding human emotions and expressing human emotions [4].

Outcomes and expectations from the previous studies show that ERS can be an important technology in the making based on the advantages offered to individuals, society, organizations, businesses and industry via various platforms and applications. For example, ERS in healthcare [5], ERS in driving assistance in the car [6], ERS in the classroom [7] and ERS in smartwatches as the latest addition due to the modalities of smartwatches being the same ERS modalities.

In existing research on ERS, researchers adopted AI algorithms such as convolutional neural network (CNN) and deep neural network (DNN) for emotion recognition based on various data modalities like facial expression, voice intonation, heart signals (electrocardiogram), brain signal (electroencephalogram) and many others. It is important to look into the relationship between ERS and its significance in society and industry specifically in the industrial revolution.

ERS can be seen as embedded technology in various applications that can be used for professional or individual day-to-day use. Based on the modalities for ERS, and the potential applications of ERS to society and the industry, it is important to look into the relationship between ERS and its significance in society and industry specifically in the industrial revolution.

1.1 *Artificial Intelligence (AI)*

AI has transformed from just a theory to tangibility with recent applications from AI to transform businesses, industries and societies [8]. Inspired by human intelligence, AI aims to learn, reason, making decisions like humans [9]. With years of advancing and refining AI, technologies or smart devices nowadays are equipped with AI minimizing the need for human intervention [9].

In recent technological advancements, AI established trustworthiness enabling the AI systems in the aspect of beneficence, non-maleficence, autonomy, justice and explicability [9,10]. The uniqueness of AI is characterized by the integration of AI capabilities with human capabilities [9,11]. AI systems are designed to operate with varying levels as well as defined objectives, predictions and recommendations influenced by real or virtual environments.

AI technologies offer benefits for businesses and industries such as the automation of repetitive and time-consuming tasks that allow humans to focus on higher-value of work [9]. For example, massive restructured data that was once required by human expertise to further analysis, such example, data from written reports, analyzing documents, photos, images and videos can be assisted by AI, with AI considered much faster and more efficient working in completing tasks in a shorter time [9]. Moreover, AI can integrate thousands of computers [9] and other resources to resolve complex problems.

Therefore, AI capabilities are leveraging the need for technological advancement to automate the process further. AI can provide innovative solutions based on the previous example in previous studies towards societies, individuals and industries. With the addition of ERS as an emerging technology, it can enhance technology further by technology to understand human emotions and behavioural responses [9].

1.2 *The Industrial Revolution*

The history of the industrial revolution began in the 18th century with the mechanisation known as the First Industrial Revolution (IR 1.0) which was based on machines powered by water and steam. Today, we are living in the era of the Fourth Industrial Revolution (IR 4.0) where its main characteristics are digitalisation and integration of computers in co-operating with humans. IR 4.0 reshapes individuals' lives and works with digital transformation and emerging technologies that grow the industry further [12].

In IR 4.0, technologies such as artificial intelligence (AI), the internet of things (IoT) and big data services achieve sustainability and productivity [12]. IR 4.0 concept was mainly focused on the manufacturing industry to smart manufacturing but has evolved in the last decade with the involvement of digital transformation. The results create the entire industry, delivery channels become digitization [12,13]. IR 4.0 is a big part of an individual's daily life and industry. However, innovations keep being introduced, soon there will be Fifth Industrial Revolution (IR 5.0) [14].

The main characteristic of I.R 5.0 that differentiate it from I.R 4.0 is specialization where machines and computers are equipped with the ability to understand human and act accordingly [15,16]. As suggested by [16], the industry will enable humans and machines to work together by bringing humans back into the workforce, therefore, IR 5.0 will be seen to unlock the potential of human-computer interaction (HCI). To signifies human-computer interaction, communication between human-computer need to be enabled. For example in [4,16], the computer system must have the ability to communicate with humans in some form.

Since the used cases of IR 5.0 are still in their formative years, manufacturers must actively organize their way to corporate humans and machines to maximise the opportunity that can be gained in IR 5.0. Technology enablers such as ERS may enhance robots and machines to understand human emotions with the proposition of collaborative robots (Cobots), hence, [4] suggestions in the study can help achieve a better understanding of human-computer interaction.

2. Methodology

This study adopted the systematic literature review (SLR) methodology consistent with prior works [17,18,19,20,21]. This method is ideal when a complete, summary of prevalent knowledge is needed [22]. SLR allows the researchers to identify and synthesise all relevant research in building up to the research gap, that led to the research question and research objectives respectively [21].

In 2009, an international group of systematic reviewers, methodologists and journal editors publish the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) and a guideline designed to help authors prepare a complete report for SLR [23]. PRISMA consists of three stages, first one is identification. In this section, keywords or search strings, criteria and database used will be identified. For this study, the search strings were Emotion Recognition System, Youth, Malaysia, Readiness, and Acceptance. The criteria were the OR and AND Boolean operatives [17]. The term Emotion recognition system expanded with synonyms of Emotion Recognition Technology and Emotion Recognition Applications. The search was done using the Lens.org platform. The benefits of using this platform are increasingly recognized by researchers [17,24]. First launched by Cambia in 2000 as Patent Lens, an Australia-based non-profit organization. In 2013 it became The Lens and has since become more than just a patent search platform.

Initially, the search for the Emotion Recognition System, in general, shows a result of 1,050 publications found based on The Lens filter. This study limited the date range from 2011 to 2022 and the document types are journaled articles, conference proceedings articles and dissertations. In the first identification stage, we removed any duplicated publications before moving further into the next stage. Based on the total results, 2 identified publications were removed before the next stage as shown in Fig.1.

The next stage is screening, before narrowing it down to retrieval and eligibility. In this stage, we screened the 1,048 publications by language (selecting papers written in the English language only) and relevance (selecting papers on ERS in Malaysia only). Therefore 2 papers have been excluded because not in the English language. Secondly, since the collection also included technical papers on ERS from the perspective of computer science and informatics, we further filter the papers based on the details of the title and abstract. Those found not focusing on ERS were excluded. Hence, 267 publications were excluded. Thus, the total records that were sought for retrieval such as in Fig.1, were narrowed down to 779.

Then, in the retrieval stage, we need to ensure the records are accessible. A total of 156 papers were identified as not accessible due to subscription restrictions. Due to budget restrictions, it was not possible to purchase subscriptions or individual papers. The remaining records were confirmed to be accessible either under the existing institutional subscriptions or via open access. Ultimately, 623 papers remained for retrieval and were included in the final eligibility screening.

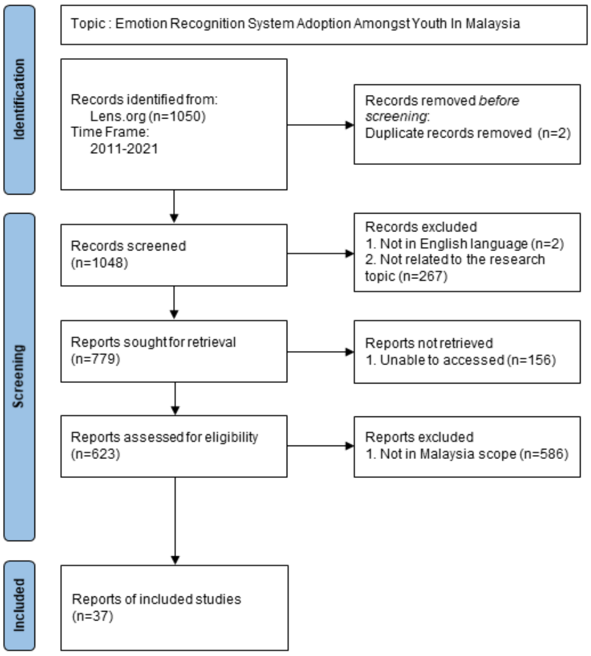

Identification
Screening
Included
Topic : Emotion Recognition System Adoption Amongst Youth In Malaysia
Records identified from:
Lens.org (n=1050)
Time Frame:
2011-2021
Records removed *before screening*:
Duplicate records removed (n=2)
Records screened
(n=1048)
Records excluded
1. Not in English language (n=2)
2. Not related to the research topic (n=267)
Reports sought for retrieval
(n=779)
Reports not retrieved
1. Unable to accessed (n=156)
Reports assessed for eligibility
(n=623)
Reports excluded
1. Not in Malaysia scope (n=586)
Reports of included studies
(n=37)

Fig.1.Prisma Flow Diagram

The final screening stage is eligibility according to the scope of the SLR. Here only works reporting from the Malaysian context are the records of interest. A total of 586 records were identified to be out of scope and thus not eligible for inclusion in this study. The final number of records confirmed to be included in this SLR is 37. The following section reports the results and analysis of the final set.

3. Results and Analysis

This section reports the descriptive analysis results of the included 37 records, focusing on the document type, year of publications, subject area, keywords related to the studies and most cited.

3.1 *Document Type*

The data that has been collected has been classified based on three categories journal article, conference proceedings article and dissertation. These types of research works are selected for the quality control process normally in place before publication. Specifically, the works go through a stringent review process by subject matter experts to verify the quality of the work. Moreover, the journal article and conference proceedings article are chosen due to journal has a better structure while the conference is short and precise that has been used in the conference [25].

Based on Table 1, the majority of the publications are journaled articles; 20 publications or 54.05%, while the rest are conference proceedings articles; 17 publications or 45.95%. The was none for the dissertation category suggesting there is potential for postgraduate research in this area.

Table 1.Document Type

Document Type	Number of publications	% (N=37)
Journal Article	20	54.05%
Conference Proceedings Article	17	45.95%
Dissertation	0	0%
TOTAL	**37**	**100**

3.2 *Year of Publication*

This study review a decade worth of works from the year 2011 to 2022, and the returned publications for each year are different. There is no definitive trend according to the no of publications over the decade (see Fig.2). It can be said that there had been a steady number of conference papers each year, but a relatively increasing trend can be seen for journal articles.

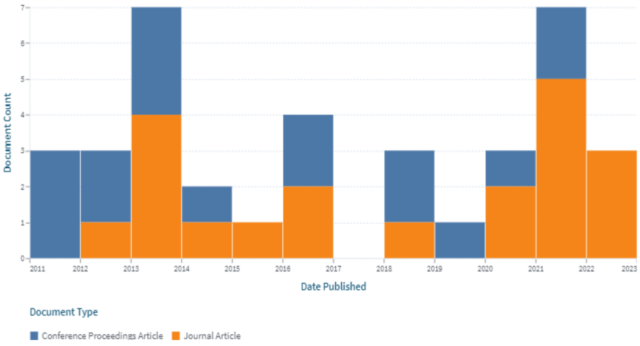


Fig.2. Year of publications

From Fig.2, we see that in the range of 2011 towards 2014 the field of affective computing, artificial intelligence and machine learning is attracting the interest of the researcher to find solutions in creating ERS, however, from 2015 to 2019, the number of publications is not constant, however, 2019 to 2022, the number of publications is increasing again. The upswing in the past 3 years can be due to the technological advancements that enable researchers to test ERS based on the modalities in the real world.

3.3. *Field of Research*

Next, the collected records were analyzed according to their field of research or subject matter (see Table 2). The highest ranking area of study is computer science with a total of 9 articles or 24.32%. Secondly, is engineering with 7 articles or 18.92%.

Table 2. Subject Matter

Field	Frequency	%(n=37)
Computer Science	9	24.32%
Engineering	7	18.92%
Artificial Intelligence	4	10.81%
Information Systems	4	10.81%
Biotechnology	3	8.11%
Computational computers	3	8.11%
Environmental Science	2	5.41%
Social Science	1	2.70%
Human-Computer Interaction	1	2.70%
Material Science	1	2.70%
Software Analytics	1	2.70%
TOTAL	37	100

The majority of the discussions focused on the potential applications, benefits, implementation or deployment, safety, and risk of ERS. However, out of the 37 articles, there is only one research truly within the field of social science; from [26] regarding the awareness and readiness of Malaysian University students for ERS. Arguably, the implications of the study can be extended from the context of Malaysian University students to Malaysian youths in general.

3.4. *Most Cited*

Next, the retrieved records that were included in the studies were analyzed in terms of citation. According to the analytics from Lens.org, the 37 publications collectively garnered 510 scholarly citations. Table 3 shows the top 10 most cited works in the collection. Ranked 1st, the paper by [27] is also one of the earliest publications in the retrieved records, thus it is expected to have a high citation. Furthermore, [27] mostly discussed the modalities that can be used for ERS. Such example, the electrocardiogram (ECG) is one of the modalities for ERS. Since the paper is a review, it provides a useful compilation of related works, and this also contributes towards a higher citation number.

Table 3. Top 10 most cited

Authors	Title	Year	Cites
Jeritta et al.	Physiological signals based human emotion recognition: A Review	2011	223
Murugappan & Murugappan	Human emotion classification using wavelet transform and KNN	2011	52
Seng et al.	A combine rule based and machine learning audio-visual emotion recognition approach	2018	43
Rashid et al.	Human emotion recognition from videos using spatio temporal and audio features	2012	37
Jeritta et al.,	Electrocardiogram based emotion recognition system using empirical mode decomposition and discrete fourier transform	2013	32
Bakhtiyari & Husain	Fuzzy model of dominance emotions in affective computing	2014	22
Wani et al.	A comprehensive review of speech emotion recognition systems	2021	21
Chew et al.	Audio emotion recognition system using parallel classifiers and audio feature analyzer	2011	13
Jeritta et al.	Emotion detection from QRS Complex of ECG signals using hurst exponent from different age groups	2013	10

Hasnul et al.	Electrocardiogram based emotion recognition applications in healthcare : A Review	2021	9

From table 3, it can be observed that the earlier works tend to focus on the technical development of ERS; from developing modalities to establishing the system. More recent works such as [5], classified the applications that can benefit ERS such as the ECG-based ERS for healthcare applications. Meanwhile, [28] reviewed works on speech-based ERS.

4. Discussion

The trend for ERS has been increasing, although there is a gap between time and studies based on Fig.2. However, the momentum of the publications returned has risen from 2019 to 2022. ERS is not seen as a new technology, instead, the previous researcher suggested ERS is seen as an embedded technology through existing devices and technologies [29]. Given the advantages and continued advancement of AI, machine learning as well as deep learning, ERS has the potential to be the enabler of new features in various applications that utilize recognition of human emotions [30].

The following sections discuss the key insights derived from the retrieved 37 publications. Specifically, the discussion focused on modalities, applications of ERS, research gaps identified, limitations, and future directions for research in ERS.

4.1. *Modalities*

ERS can be identified through various input data modalities, such example, physiological modalities, psychological modalities and data mining modalities. These modalities have been majorly developed in the field of computer science, engineering, artificial intelligence, information systems, biotechnology and computational computers.

4.1.1 *Physiological Modalities*

In physiological modalities, it is consisting of electroencephalography (EEG), electrocardiography (ECG) and photoplethysmography (PPG). EEG is frequently used in neuroscience, neural engineering, and biomedical engineering to measure human brain signals through the electromagnetic behaviour of specific components [30,31,32]. EEG is most likely the favourite for obtaining high-accuracy data for automated emotions recognitions because EEG processes use the same concept as AI systems that used convolutional neural networks for machine learning and deep machine learning [33,27]. ECG is commonly used by previous researchers in the field of AC due to its potential adoption using a wearable device. In a previous study by [5], ECG was used to detect stress and the benefit of monitoring a patient's emotional stress condition to ensure that a negative tendency is not triggered. Next for EEG, is photoplethysmography (PPG). PPG has been said to be more practical and suitable in real-life use compared to EEG, in combination with ECG as multimodal, it can be implemented as a "wearable device that can collect signals from a person without compromising comfort and privacy" [35,36]. For instance, current technology devices such as a smartwatch have a built-in ECG to detect stress levels and blood pressure, consequently identifying temporal emotional states.

4.1.2 *Physical Modalities*

For physical modalities, the modalities that have been widely used are facial recognition and speech recognition. Facial recognition has gained popularity amongst ERS practitioners due to diverse applications that can be applied in the real world such as marketing purposes, security supervision, online class and gaming experience [36]. Facial expression can determine the six basic emotions, namely, sad, anger, happiness, fear and surprise. Additionally, facial expressions can be used to detect disgust, normality, and drowsiness [37]. Speech recognition as modalities of ERS is capable of detecting human speech and interpreting conditions and emotions for real-life applications [32,38]. According to [39], speech recognition is defined based on someone's voice, the computer can learn what condition they are in, therefore concluding what is the emotion of that person. In [28], a review of the speech emotion recognition system shows that speech recognition detects and classify emotion recognition system, only the inability to handle the real-time problem and different scenario that affect human emotions as human speech is vary based on culture.

4.1.3 *Data Mining Modalities*

In data mining modalities, one paper from [40], suggested that data mining known as text mining is a learning-based algorithm to describe characteristics of text, such as word expression based on human sentiments and emotions. Such example, from social media data mining, the text “Hurray!”, can be detected as part of happy emotions [41], and “Argh!” can be detected as anger and frustration, however, there are certain texts that create multiple emotions overlapping with the intended emotions, such as “Awww”, which can indicate pleasant sentiments but sometimes it can indicate a pity sentiment [40].

4.2. *Applications*

Based on the conducted SLR, some papers presented the potential applications that can be achieved through individuals, society and industries. Firstly, the education system can be improved with the adoption of ERS [42]. Emotion has a significant influence on the relationship with a performance from the perspective of a learner and instructor as described by [7], and students with positive emotions led to increasing student interest, and student performance and “have higher chances at success”. [7] their study suggested a webcam inside a computer laboratory and using a facial recognition modality to identify students’ moods and the results show significant differences in the student’s moods.

From an individual perspective, emotions can affect driving [6]. This is due to emotions’ relationship with focus. For example, anger can lead to poor focus. A poor focus led to an increasing number of accidents, therefore, it is significant for ERS to be embedded in cars for driving safety [6]. Furthermore, [6] used a driving simulation, where a virtual driving experiment was conducted to test ECG-based ERS in identifying human emotions while driving.

In [5], ERS was implemented using ECG in the healthcare industry to “reduce stress and promote relaxation”. The ECG-based ERS helps a person with mental stress to identify negative and positive emotions. Furthermore, ERS can be suggested as a supporting aid for people with certain disorders such as down syndrome, autism, and even elderly people. It will be using facial expressions in real-time video as automated emotion recognition and a computer advisor that advises on how we can react appropriately [43,3].

ERS can be potential in marketing strategy. Such example in brand awareness, product acceptance can be analyzed based on the response of human facial expressions in a live launch or video and images [43]. For example, a small group of people in a room are introduced to a product, and their reactions through facial expressions will be recorded and evaluated by ERS whether they are happy, disgusted, or anxious.

In another aspect, in video games, the developer of games has an objective which is to “fulfil his or her dream” while playing games through immersive gameplay, interesting storyline and intensive graphics [44,45]. Implementation of emotion detection can improve engagement between entertainment agents and end-user by alternatively using multimodalities such as speech recognition and EEG in the controller which can improve understanding of human emotions towards intangibles products such as music or movies [43]. In addition, with virtual reality (VR) technology in recent years, the developer can enhance the game further with real-time ERS usage towards understanding the player.

4.3. *Gaps*

With the valuable insights based on the previous paper, there are several gaps identified that motivate this study. The objective of this study particularly is to gain insights into adopters’ perspectives. The results of performing SLR are to understand the recent study on ERS and the existing gap between previous research and studies. From the results of performing the SLR on the specified topic, in Malaysia, there is a smaller number of returned publications with only 37 papers (see figure 1), specifically, studies regarding ERS adoption. Previous research signifies that ERS contributes within the aspect of the engineers and scientists, however, in real-life problems, there will be a different situation that might not apply.

Based on the Lens search, the availability of ERS is limited due to ERS being practically known amongst researchers related to the field of computer science, engineering, artificial intelligence and information systems. And when we narrowed down the scope to Malaysia, the trend of publication is slower during the early year between 2011-2016 due to the technology is still in a growth state. However, in 2017-2021, the publications are getting momentum again along with the advancement of technology, and emotion recognition has gotten recognizability throughout Malaysia. Next, we dive deep into the adoption towards the Malaysian population. In 2021, there one paper signifies the acceptance of Malaysian university students towards ERS [26]. This paper discusses the perspective of Malaysian university student awareness towards ERS whether they know about ERS and are ready to adopt the technology or they are still unaware of the technology.

In conclusion, there is still a lack of studies related to ERS in Malaysia and ERS adoption and readiness in Malaysia. This study will help to gain better insights on the perspective of ERS adoption and readiness and contribute to the practitioners and policy making by the government to ensure that ERS is more practical to use amongst users and important determinants of user's adoption in ERS.

4.4. *Limitations*

Conducting the SLR studies, and analyzing the previous paper, some limitations have been found and highlighted by the previous researchers. Firstly, the security and privacy from the emotional recognition. As machines evolve for to better over the next few years, there is a part where for humans, exposing privacy is the last part to do. Although in marketing, there is a benefit of ERS to be implemented, however, there is a risk to privacy and security, as capturing facial recognition and analyzing text data mining can feel like invading privacy and security breach without consent.

Secondly, another study [29], suggested that there is uncertainty for ERS as a limitation due to recognizing human emotions can create fear amongst humans. Specifically, in the workplace, emotion will play an important role in determining performance, productivity and efficiency, however, there will be at times when humans do not want to be measured for their emotions and observing their behaviour, therefore, will create unpleasant surroundings.

Next, from the perspective of social science, there is a limited number of papers from SLR conducted that focuses on ERS readiness and adoption. There is a need to understand the perspective of users that will be going to face the technology and the perceived usefulness of the technology matters in predicting accurately the technology will be performing in real-life.

4.5 *Future Directions*

From the paper published in Lens, the future directions for ERS there is a need for more papers for ERS as suggested that ERS applications will benefit individuals and societies in everyday lives. In addition, ERS can play a major role in the industries as suggested by the previous researcher due to the rising of the industrial revolution as seen in IR 4.0 with AI development is significant.

Concurrently, IR 5.0 key takeout focuses on human-computer interaction (HCI), there is a significant relationship between ERS in IR 5.0 with AI can be the decision makers for robots and ERS as a subpart of AI enhancing robots further with enabling the human-computer interaction [46]. As we live in the direction of technological advancement era, understanding recent development in technologies will increase the scope and understanding more about technology and human capabilities and the achievement of human and computer relationships.

5. Conclusion

Acknowledgement

This work was supported by "TMR&D Grant 2021: Emotion Recognition System (ERS) for Smart Homes Malaysia".

References

1. Kodhai, E., Pooveswari, A., Sharmila, P., & Ramiya, N. (2020). Literature Review on Emotion Recognition System. 2020 International Conference on System, Computation, Automation and Networking, ICSCAN 2020, 18–21. https://doi.org/10.1109/ICSCAN49426.2020.9262389
2. Picard, R. W. (1999). Affective Computing for HCI. Proceedings of the 8th HCI International on Human-Computer Interaction: Ergonomics and User Interfaces, 829–833. http://dl.acm.org/citation.cfm?id=647943.742338
3. Haridas, A. V., Marimuthu, R., & Chakraborty, B. (2020). Emotion Recognition System for Specially Needed People with Optimized Deep Learning

Algorithm. In 2020 Fourth International Conference on Inventive Systems and Control (ICISC). IEEE. https://doi.org/10.1109/icisc47916.2020.9171190

4. Yoshitomi, Y. (2021). Human–computer communication using recognition and synthesis of facial expression. Journal of Robotics, Networking and Artificial Life, 8(1), 10–13. https://doi.org/10.2991/jrnal.k.210521.003
5. Hasnul, M. A., Aziz, N. A. A., Alelyani, S., Mohana, M., & Aziz, A. A. (2021). Electrocardiogram-Based Emotion Recognition Systems and Their Applications in Healthcare-A Review. Sensors (Basel, Switzerland), 21(15), 5015. https://doi.org/10.3390/s21155015
6. Wang, X., Guo, Y., Ban, J., Xu, Q., Bai, C., & Liu, S. (2020). Driver emotion recognition of multiple-ECG feature fusion based on BP network and D-S evidence. IET Intelligent Transport Systems, 14(8), 815–824. https://doi.org/10.1049/iet-its.2019.0499
7. Putra, W. B., & Arifin, F. (2019). Real-Time Emotion Recognition System to Monitor Student's Mood in a Classroom. Journal of Physics: Conference Series, 1413(1). https://doi.org/10.1088/1742-6596/1413/1/012021
8. Nishant, R., Kennedy, M., & Corbett, J. (2020). Artificial intelligence for sustainability: Challenges, opportunities, and a research agenda. International Journal of Information Management, 53(March), 102104. https://doi.org/10.1016/j.ijinfomgt.2020.102104
9. Choung, H., David, P., & Ross, A. (2022). Trust in AI and Its Role in the Acceptance of AI Technologies. International Journal of Human-Computer Interaction, March. https://doi.org/10.1080/10447318.2022.2050543
10. Thiebes, S., Lins, S., & Sunyaev, A. (2021). Trustworthy artificial intelligence. Electronic Markets, 31(2), 447–464. https://doi.org/10.1007/s12525-020-00441-4
11. Krafft, P. M., Young, M., Katell, M., Huang, K., & Bugingo, G. (2020). Defining AI in policy versus practice. AIES 2020 - Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, 72–78. https://doi.org/10.1145/3375627.3375835
12. Ghobakhloo, M. (2020). Industry 4.0, digitization, and opportunities for sustainability. Journal of Cleaner Production, 252, 119869. https://doi.org/10.1016/j.jclepro.2019.119869
13. Mohamed, N., Al-Jaroodi, J., & Lazarova-Molnar, S. (2019). Leveraging the Capabilities of Industry 4.0 for Improving Energy Efficiency in Smart Factories. IEEE Access, 7, 18008–18020. https://doi.org/10.1109/ACCESS.2019.2897045
14. Aslam, F., Aimin, W., Li, M., & Rehman, K. U. (2020). Innovation in the era of IoT and industry 5.0: Absolute innovation management (AIM) framework. Information (Switzerland), 11(2). https://doi.org/10.3390/info11020124
15. Sumi, K. (2016). Affective Human Computer Interaction. Proceedings of International Conference on Artificial Life and Robotics, 21(2), 244–248. https://doi.org/10.5954/icarob.2016.is2
16. George, A. S., & George, A. S. H. (2020). Industrial Revolution 5.0: the Transformation of the Modern Manufacturing Process To Enable Man and Machine To Work Hand in Hand. Journal of Seybold Report, 15(9), 214–234. https://doi.org/10.5281/zenodo.6548092
17. K.G, P., De Alwis, C., & I., W. (2022). The Facets of Gender Stereotypes Change: A Systematic Literature Review. SSRN Electronic Journal, June 2022. https://doi.org/10.2139/ssrn.4117570
18. Kumar, S., Kar, A. K., & Ilavarasan, P. V. (2021). Applications of text mining in services management: A systematic literature review. International Journal of Information Management Data Insights, 1(1), 100008. https://doi.org/10.1016/j.jjimei.2021.100008
19. Kuss, D. J., Kristensen, A. M., & Lopez-Fernandez, O. (2021). Internet addictions outside of Europe: A systematic literature review. Computers in Human Behavior, 115, 106621. https://doi.org/10.1016/j.chb.2020.106621
20. Toorajipour, R., Sohrabpour, V., Nazarpour, A., Oghazi, P., & Fischl, M. (2021). Artificial intelligence in supply chain management: A systematic literature review. Journal of Business Research, 122(September 2020), 502–517. https://doi.org/10.1016/j.jbusres.2020.09.009
21. Van Dinter, R., Tekinerdogan, B., & Catal, C. (2021). Automation of systematic literature reviews: A systematic literature review. Information and Software Technology, 136(March), 106589. https://doi.org/10.1016/j.infsof.2021.106589
22. Schmeisser, B. (2013). A systematic review of literature on offshoring of value chain activities. Journal of International Management, 19(4), 390–406. https://doi.org/10.1016/j.intman.2013.03.011
23. Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., & Moher, D. (2021). Updating guidance for reporting systematic reviews:

development of the PRISMA 2020 statement. Journal of Clinical Epidemiology, 134, 103–112. https://doi.org/10.1016/j.jclinepi.2021.02.003

24. Martín-Martín, A., Thelwall, M., Orduna-Malea, E., & Delgado López-Cózar, E. (2021). Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations' COCI: a multidisciplinary comparison of coverage via citations. In Scientometrics (Vol. 126, Issue 1). Springer International Publishing. https://doi.org/10.1007/s11192-020-03690-4
25. Eckel, E. J. (2009). The Emerging engineering scholar: A citation analysis of theses and dissertations at western michigan university. Issues in Science and Technology Librarianship, 56. https://doi.org/10.29173/istl2470
26. Aziz, N. A. A., Aziz, N. H. A., Ismail, S. N. M. S., Khan, C. M. T., Hasnul, M. A., Rahman, A., Rahman, T. A., & Aziz, K. A. (2021). Awareness and Readiness of Malaysian University Students for Emotion Recognition System. The International Journal of Integrated Engineering, 13(6), 299–309. https://doi.org/10.30880/ijie.2021.13.06.026
27. Jerritta, S., Murugappan, M., Nagarajan, R., & Wan, K. (2011). Physiological signals based human emotion recognition: A review. Proceedings - 2011 IEEE 7th International Colloquium on Signal Processing and Its Applications, CSPA 2011, 410–415. https://doi.org/10.1109/CSPA.2011.5759912
28. Wani, T. M., Gunawan, T. S., Qadri, S. A. A., Kartiwi, M., & Ambikairajah, E. (2021). A Comprehensive Review of Speech Emotion Recognition Systems. IEEE Access, 9, 47795–47814. https://doi.org/10.1109/ACCESS.2021.3068045
29. Landowska, A. (2019). Uncertainty in emotion recognition. Journal of Information, Communication and Ethics in Society, 17(3), 273–291. https://doi.org/10.1108/JICES-03-2019-0034
30. Alhalaseh, R., & Alasasfeh, S. (2020). Machine-learning based emotion recognition system using EEG signals. Computers, 9(4), 1–15. https://doi.org/10.3390/computers9040095
31. Craik, A., He, Y., & Contreras-Vidal, J. L. (2019). Deep learning for electroencephalogram (EEG) classification tasks: A review. Journal of Neural Engineering, 16(3). https://doi.org/10.1088/1741-2552/ab0ab5
32. Jerritta, S., Murugappan, M., Wan, K., & Yaacob, S. (2014). Electrocardiogram-based emotion recognition system using empirical mode decomposition and discrete Fourier transform. Expert Systems, 31(2), 110–120. https://doi.org/10.1111/exsy.12014
33. Fang, W. C., Wang, K. Y., Fahier, N., Ho, Y. L., & Huang, Y. De. (2019). Development and Validation of an EEG-Based Real-Time Emotion Recognition System Using Edge AI Computing Platform With Convolutional Neural Network System-on-Chip Design. IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 9(4), 645–657. https://doi.org/10.1109/JETCAS.2019.2951232
34. Suhaimi, N. S., Mountstephens, J., & Teo, J. (2020). EEG-Based Emotion Recognition: A State-of-the-Art Review of Current Trends and Opportunities. Computational Intelligence and Neuroscience, 2020. https://doi.org/10.1155/2020/8875426
35. Tong, Z., Chen, X., He, Z., Tong, K., Fang, Z., & Wang, X. (2018). Emotion Recognition Based on Photoplethysmogram and Electroencephalogram. Proceedings - International Computer Software and Applications Conference, 2, 402–407. https://doi.org/10.1109/COMPSAC.2018.10266
36. Rathour, N., Khanam, Z., Gehlot, A., Singh, R., Rashid, M., Alghamdi, A. S., & Alshamrani, S. S. (2021). Real-time facial emotion recognition framework for employees of organizations using raspberry-pi. Applied Sciences (Switzerland), 11(22). https://doi.org/10.3390/app112210540
37. Kundu, T., & Saravanan, C. (2018). Advancements and recent trends in emotion recognition using facial image analysis and machine learning models. International Conference on Electrical, Electronics, Communication Computer Technologies and Optimization Techniques, ICEECCOT 2017, 2018-Janua, 1–6. https://doi.org/10.1109/ICEECCOT.2017.8284512
38. Shinde, A. S., & Patil, V. V. (2021). Speech Emotion Recognition System: A Review. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3869462
39. El Ayadi, M., Kamel, M. S., & Karray, F. (2011). Survey on speech emotion recognition: Features, classification schemes, and databases. Pattern Recognition, 44(3), 572–587. https://doi.org/10.1016/j.patcog.2010.09.020
40. Zucco, C., Calabrese, B., Agapito, G., Guzzi, P. H., & Cannataro, M. (2020). Sentiment analysis for mining texts and social networks data: Methods and tools. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10(1), 1–32. https://doi.org/10.1002/widm.1333
41. Murthy, A. R., & Anil Kumar, K. M. (2021). A Review of Different Approaches for Detecting

Emotion from Text. IOP Conference Series: Materials Science and Engineering, 1110(1), 012009. https://doi.org/10.1088/1757-899x/1110/1/012009

42. Bouhlal, M., Aarika, K., AitAbdelouahid, R., Elfilali, S., & Benlahmar, E. (2020). Emotions recognition as an innovative tool for improving students' performance and learning approaches. Procedia Computer Science, 175, 597–602. https://doi.org/10.1016/j.procs.2020.07.086
43. Arrais, J. P., Laranjeira, A., Oliveira, G., & Ribeiro, B. (2017). Deep learning in digital marketing: brand detection and emotion recognition. International Journal of Machine Intelligence and Sensory Signal Processing, 2(1), 32. https://doi.org/10.1504/ijmissp.2017.10009120
44. Hippe, Z. S., Kulikowski, J. L., Mroczek, T., & Wtorek, J. (2014). Human-Computer Systems Interaction: Backgrounds and Applications 3. Advances in Intelligent Systems and Computing, 300, 51–62. https://doi.org/10.1007/978-3-319-08491-6
45. Vinola, C., & Vimaladevi, K. (2015). A survey on human emotion recognition approaches, databases and applications. Electronic Letters on Computer Vision and Image Analysis, 14(2), 24–44. https://doi.org/10.5565/rev/elcvia.795
46. Watada, S., Obayashi, M., Kuremoto, T., & Mabu, S. (2014). A Decision Making System of Robots Introducing a Re-construction of Emotions Based on Their Own Experiences. Journal of Robotics, Networking and Artificial Life, 1(1), 27. https://doi.org/10.2991/jrnal.2014.1.1.6
47. Kawada, K., & Tamai, T. (2018). Development of Teaching Support Material for Nurturing Cooperation through Play. Journal of Robotics, Networking and Artificial Life, 5(3), 169. https://doi.org/10.2991/jrnal.2018.5.3.6

Authors Introduction

Mr. Muhammad Nadzree Mohd Yamin

Muhammad Nadzree Mohd Yamin is a Master of Philosophy (Management) student from Multimedia University. He obtained his Bachelor of Business Management (Hons) in Management (First Class) from Multimedia University in 2021.

Assoc. Prof. Dr. Kamarulzaman Ab. Aziz

Kamarulzaman Ab. Aziz is currently an associate professor at the Faculty of Business, Multimedia University. His research interest includes Cluster Development, Technology and Innovation Management, Entrepreneurship and Commercialization.

Mr. Terence Tan Gek Siang

Tan Gek Siang is a lecturer in the Marketing Management Department, Faculty of Business, Multimedia University. He teaches business and management subjects at both undergraduate and postgraduate level. He is currently the Deputy Head of International Relations in the Faculty of Business. His research interests are in tourism marketing, user acceptance of information technology, virtual reality, augmented reality, and business management.

Assoc. Prof. Dr. Nor Azlina Ab. Aziz

Nor Azlina Ab Aziz is an associate professor in the Faculty of Engineering and Technology at Multimedia University, Melaka. She is interested in the field of soft computing and its application in engineering problems. More specifically, her focus is in the area of swarm intelligence and nature inspired optimization algorithm.

A Basic Study of Hand Eye Calibration using a Tablet Computer

Junya Sato
Department of Mechanical Engineering, Gifu University,
1-1, Yanagido, Gifu-shi, Gifu, 501-1193, Japan

Takayoshi Yamada
Department of Mechanical Engineering, Gifu University,
1-1, Yanagido, Gifu-shi, Gifu, 501-1193, Japan

Kazuaki Ito
Department of Mechanical Engineering, Gifu University,
1-1, Yanagido, Gifu-shi, Gifu, 501-1193, Japan

E-mail: jsato@gifu-u.ac.jp, yamat@gifu-u.ac.jp, kazu_it@gifu-u.ac.jp

Abstract

In this study, we describe a hand eye calibration method for calibrating an attached camera and a handmade end-effector to a robot by utilizing a tablet computer. By iterating the touching black dots displayed in the computer and optimizing robot parameters to minimize the touching error, the hand eye calibration is achieved without the end-effector information such as the dimensions and attached position. We finally achieved 1.2 mm touching error.

Keywords: hand eye calibration, differential evolution, evolutionary computation, tablet computer

1. Introduction

To pick an object using a robot and vision system, it is necessary to calibrate a camera and robot in advance. More specifically, homogeneous transformation matrices between the camera and robot base or hand should be estimated accurately via the calibration. This is known as hand eye calibration. The general approach is to use a calibration board such as a checkerboard. However, this approach does not consider information such as dimensions of an end-effector and its attached position to the robot. Therefore, manual measurements of them are required, especially for a handmade end-effector. However, because the manual measurements contain errors, adjustment of the homogeneous transformation matrices is necessary. This is a tedious work. To address this problem, this study describes a method to achieve hand eye calibration without end-effector information by utilizing a tablet computer.

2. Related works

Cao *et al*. proposed a method that uses a neural network[1]. There is a direct approach that uses a laser tracker to compensate for the absolute position errors of a robot. However, such measurement devices are expensive. For this reason, Cao *et al*. proposed an indirect approach that does not use any device, but instead use the neural network to compensate the error. Mišeikis *et al*. achieved a wide range of accurate calibration by attaching a checkerboard to a robot and captured images of it using three RGB-D cameras[2]. However, this requires space to install the three cameras, which incurs high cost.

These previous works use a calibration board with a specific pattern, such as a checkerboard. Hence, this kind of calibration is a general approach. In contrast, Lee *et al*. proposed a method that does not use any markers, such as checkerboards[3]. First, the robot is captured by an RGB camera placed next to the robot, and each joint position of the robot is recognized using deep learning. The calibration is achieved by using joint positions, forward kinematics, and camera intrinsic. Nevertheless, some space is required for the camera to be installed to capture the robot.

As described, most previous works do not consider an end-effector such as its dimensions and attached position to a robot. Thus, if a handmade end-effector is used, additional calibration is required for the accurate manipulation. In order to avoid it, a novel hand eye calibration method using a tablet computer is proposed in this study.

3. Proposed method

The overview and appearance of the proposed method are shown in Figs. 1, 2 and 3, respectively. A tablet computer was used for the proposed method. The robot is installed with an RGB-D camera and a pen of the tablet as the end-effector. First, the screen of the tablet is captured by the camera (Fig. 4), and the coordinates of the nine black dots are obtained by image processing. Because their coordinate system is the image coordinate system, the obtained coordinates are transformed to the robot's base coordinate system using homogeneous transformation matrices. Second, the robot touches the dots with the attached pen. Third, the distance errors between the original and touched coordinates are obtained. Fourth, the average of the errors is calculated, and the parameters for homogeneous transformation matrices are optimized by differential evolution (DE) to minimize this error. Finally, the optimized parameters are set to the robot, and the same process is repeated. Using this iteration, the hand eye calibration is achieved. Because this approach does not require any operation by human while calibrating, smaller effort requires than existing approaches.

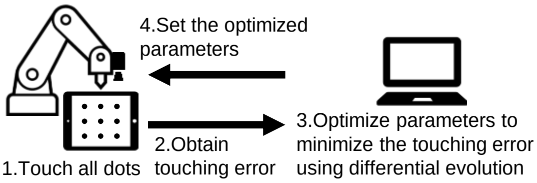

4.Set the optimized parameters
3.Optimize parameters to minimize the touching error using differential evolution
2.Obtain touching error
1.Touch all dots

Fig. 1. Overview of the proposed method.

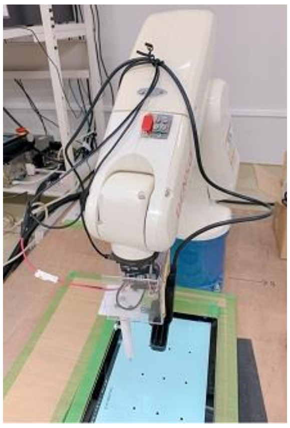


Fig. 2. Appearance of the proposed system.

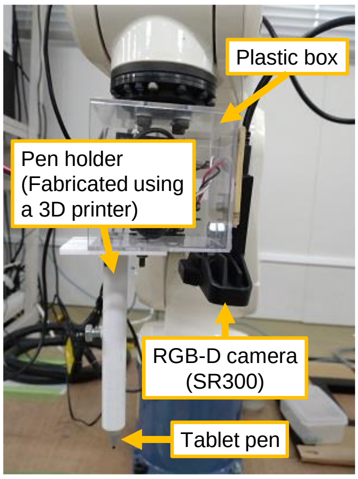


Fig. 3. Details of the end-effector in our method.

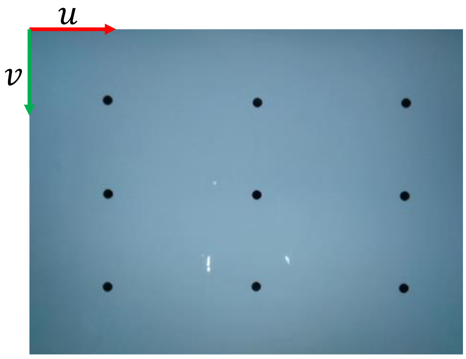


Fig. 4. Captured image by SR300, which is an RGB-D camera (640×480 pixels). The *u* and *v* axes represent image coordinate system, which is used in OpenCV for image processing.

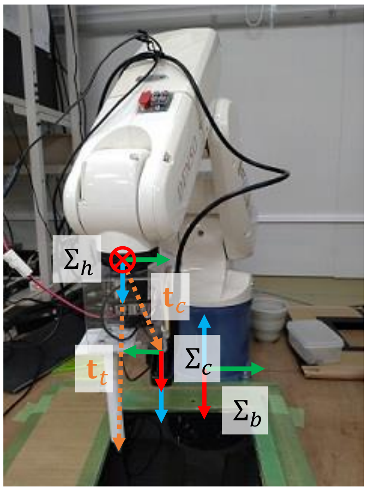


Fig. 5. Each coordinate system. The red, green, and blue arrows indicate *x*, *y*, and *z* axes, respectively.

3.1. ***Coordinate system and optimized parameters***

Figure 5 shows each coordinate system in our method. The robot base coordinate, hand coordinate system, and camera coordinate system are denoted as Σ_b, Σ_h, and Σ_c, respectively. For the transformation from Σ_b to Σ_h, rotation of -180° in y axis and translation of $(x_h, y_h, z_h)^T$ are required. Similarly, for the transformation from Σ_h to Σ_c, rotations of 180° in y axis and 180° in x axis, and translation of $\mathbf{t}_c = (x_c, y_c, z_c)^T$ are necessary. An equation to transform a position of ith black dot in Σ_c $((x_i^c, y_i^c, z_i^c)^T)$, which is measured by the camera, to the robot base coordinate system $\left((x_i^b, y_i^b, z_i^b)^T\right)$ is:

$$\begin{bmatrix} x_i^b \\ y_i^b \\ z_i^b \\ 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & x_h \\ 0 & 1 & 0 & y_h \\ 0 & 0 & -1 & z_h \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 & x_c \\ 0 & -1 & 0 & y_c \\ 0 & 0 & 1 & z_c \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_i^c \\ y_i^c \\ z_i^c \\ 1 \end{bmatrix} \quad (1)$$

Because $(x_i^c, y_i^c, z_i^c)^T$ cannot be acquired directly, they are obtained from image coordinate system using the following equation, which is based on the pinhole camera model.

$$\begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_i^c / z_i^c \\ y_i^c / z_i^c \\ 1 \end{bmatrix} \quad (2)$$

Where u_i and v_i are ith position of the black dot in the camera coordinate system. The f_x and f_y are focal length of *x* and *y* axes, respectively. The c_x and c_y are coordinate of the principal point of *x* and *y* axes, respectively. From the above equation, the positions of each black dot in the image coordinate system (x_i^c and y_i^c) can be transformed to the camera coordinate system.

Because the camera and pen are not aligned to the rotation axis of the robot hand, offset $((x', y', z')^T)$ is necessary to touch with the attached pen. It can be calculated as below.

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} \quad (3)$$

Where $\mathbf{t}_t = (x_t, y_t, z_t)^T$ is a translation vector from Σ_h to the tip of the tablet pen. The θ is a rotation angle of z axis in Σ_h. By combining Eq. (1) and (3), the position in the robot base coordinate system to touch each black dot using the tablet pen can be acquired.

$$\begin{bmatrix} x_i^{b'} \\ y_i^{b'} \\ z_i^{b'} \end{bmatrix} = \begin{bmatrix} x_i^b \\ y_i^b \\ z_i^b \end{bmatrix} - \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} \quad (4)$$

Next, the parameters to be optimized are described here. The $(x_h, y_h, z_h)^T$ in Eq. (1) represents the initial robot hand position to capture the tablet display, and this is known. The $\mathbf{t}_c = (x_c, y_c, z_c)^T$ can be roughly obtained by manual measurement. However, this is not precise. Thus, the optimization is required. Similarly, $\mathbf{t}_t = (x_t, y_t, z_t)^T$ in Eq. (3) should be optimized. The f_x, f_y, c_x, and c_y can be obtained through a manual of an RGB-D camera or SDK (Software Development Kit). Hence, the optimization for them is not required. The z_i^c represents the distance between the tablet display and the camera at the initial robot hand position. This can be known from the depth information of the RGB-D camera. The both of z_c and z' are unknown. However, they can be ignored because the robot is controlled that the attached pen always touches the tablet display to obtain the touching error. Therefore, x_c, y_c, x_t, and y_t are the parameters to be optimized.

3.2. ***Optimization by differential evolution (DE)***

In this study, differential evolution (DE), which is one of the population-based evolutionary computation methods, is adopted for the optimization because this can optimize all parameters simultaneously and is easy to use for us. The fitness function is:

$$\text{Fitness} = \sum_{i=0}^{8} \sqrt{\left(u_i - u_i^{\text{tch}}\right)^2 + \left(v_i - v_i^{\text{tch}}\right)^2} \Big/ 9 \quad (5)$$

Where $\left(u_i^{\text{tch}}, v_i^{\text{tch}}\right)$ is ith touched position by the robot. Because the parameters that minimize the touching error are required, this fitness function is used.

4. Experimental Setup

In our experiments, a 6-axis robot (DENSO VP-6242), tablet computer (Microsoft Surface Pro 7), tablet pen (Surface Pen), and an RGB-D camera (Intel SR300) were used. The set values related to Eq. (1) to (4) were $(x_h, y_h, z_h) = (320, -70, 290)$, $z_i^c = 168$, $f_x = 617.732788$, $f_y = 617.732849$, $c_x = 316.517365$, and $c_y = 242.328247$. The population and generation sizes for DE were 10 and 30, respectively. The scaling factor and crossover probability were 0.5 and 0.9, respectively. In experiments, DE/rand/1 and DE/best/1 strategies were compared. The used crossover method was binomial crossover. The set search ranges for each optimized parameter were $x_c \in [-20, 20]$, $y_c \in [20, 50]$, $x_t \in [-20, 50]$, and $y_t \in [-40, -10]$. In order to check the effectiveness of DE, random sampling (RS) was also introduced and compared.

The rotation angle of the sixth axis when the black dots were touched was gradually rotated so that the first dot was 0 degrees and the ninth was 180 degrees. In other words, the six axis was rotated by 22.5 degrees to investigate the robot can touch the dots correctly even though the hand rotates.

5. Result and Consideration

Table 1 represents the experimental results. From the S1 to S5 mean use of different random seeds because they affect the optimization results by DE. Since the results of DE/rand/1 and DE/best/1 were better than RS, the effectiveness of DE was indicated. The average mean touching error of DE/rand/1 was smaller than DE/best/1. This is because the DE/rand/1 could keep higher diversity of the population.

6. Conclusion

In this study, we proposed a hand eye calibration method using a tablet computer. Because existing previous studies do not consider dimensions of the end-effector and attached position to the robot, additional calibration for it could be required if a handmade end-effector is used. On the other hand, since our approach calibrates based on the touching error, which is obtained from the tablet computer, the hand eye calibration is achieved without the information of the end-effector. In the experiment, our method achieved minimum touching error of 1.2 mm and mean touching error of 1.7 mm.

Table 1. Mean touching error in [mm].

Method	S1	S2	S3	S4	S5	Ave.
DE/rand/1	1.5	1.2	1.3	3.3	1.2	1.7
DE/best/1	3.8	1.3	1.8	1.9	2.6	2.3
RS	4.7	4.8	4.9	3.2	2.7	4.1

References

1. C. T. Cao, V. P. Do, B. R. Lee, “A novel indirect calibration approach for robot positioning error compensation based on neural network and hand-eye vision”, applied sciences, Vol. 9, No. 9, pp. 1-17, 2019
2. J. Mišeikis, K. Glette, O. J. Elle, J. Torresen, “Automatic calibration of a robot manipulator and multi 3d camera”, proceedings of IEEE/SICE international symposium on system integration, pp. 2474-2325, 2016
3. T. E. Lee, J. Tremblay, T. To, J. Cheng, T. Mosier, O. Kroemer, D. Fox, S. Birchfield, “Camera-to-robot pose estimation from a single image”, proceedings of IEEE international conference on robotics and automation, pp. 9426-9432, 2020

Authors Introduction

Dr. Junya Sato

He received the B.E., M.E., and Ph.D. degrees in Engineering from Iwate University, Japan, in 2012, 2014, and 2017, respectively. Since 2017, he has been an Assistant Professor with the Department of Mechanical Engineering, Faculty of Engineering, Gifu University, Japan.

Dr. Takayoshi Yamada

He received the B.E., M.E., and Ph.D. degrees in mechanical engineering from the Nagoya Institute of Technology, Japan, in 1991, 1993, and 1995, respectively. He is currently a Professor with the Department of Mechanical Engineering, Gifu University, Japan.

Dr. Kazuaki Ito

He received the B.S., M.S., and Ph.D. degrees in electrical and computer engineering from the Nagoya Institute of Technology, Japan, in 1998, 2000, and 2003, respectively. He moved to the Department of Mechanical Engineering, Gifu University, Japan and has been a Professor since 2022.

Human Detection with Uprisen Angle of a Camera for the Service Robot

Watcharin Tangsuksant[1], **Amornphun Phunopas**[2], **Pornthep Sarakon**[3], **Aran Blattler**[4]
Department of Production and Robotics Engineering,
Center of Innovative Robotics and Advanced Precision Systems: iRAPs,
Faculty of Engineering, King Mongkut's University of Technology North Bangkok,
1581 Pracharat 1 Road, Wongsawang, Bangsue, Bangkok, 10800, THAILAND
watcharin.t@eng.kmutnb.ac.th [1], amornphun.p@eng.kmutnb.ac.th[2], pornthep.s@eng.kmutnb.ac.th[3], aran.b@eng.kmutnb.ac.th[4]

In order to improve the intelligent service robot, the visual perception is a crucial. This paper presents the human detection for a service robot. The camera is installed at the robot by uprisen angle around 30 degree. Two feasible algorithms for the real-time detection between Haar cascade and Single Shot Detector (SSD) algorithm are compared. This research collects the training data as 1,000 images by the uprisen angle and the different views of the human, within 60 cm of user range. between a camera and human. The result shows that our proposed model of SSD method is higher performance than another by 0.933 of the average of IoU. Therefore, the proposed method is suitable to apply for the service robot.

Keywords: Human detection, Uprisen angle, Service robot, Single shot detection, Haar cascade.

1. Introduction

The service robot can see like a human using the camera. The camera position is initially fixed with an uprisen angle to see people at a 60 cm range distance. The field of view is designed for people with a height minimum of 120 cm to see the upper body in the frame grabber. However, it is just the first procedure of data acquisition. Next, the robot needs to understand the semantics of acquired images. This paper focuses on distinguishing humans from other objects. The robot must prioritize safety for every movement to beware of whether a human is in front of it. Moreover, the robot can interact with humans with different behaviors, such as greeting and playing with the people around. Many vision-based human detection techniques [1] are reviewed in various machine learning techniques and vision-based on revolutionary methods such as SIFT, SURF, PCA, and HOG. The human detection data sets are accessible publicly, such as MS COCO, ImageNet, and WiderPerson. However, this paper prepares its own data sets for the experiment.

The human face is a biometric feature or identity for feature extraction. Anirudha B Shetty et al. in 2021 [2] have compared two facial recognition techniques, Local Binary Pattern and Haar Cascade. The results showed that Haar Cascade has higher accuracy than the Local Binary Pattern, and Haar Cascade uses more execution time than the Local Binary Pattern. Many researchers use some primary object detection techniques with good performance, such as Feature extraction using stepwise convolutional self-encoder, Faster R-CNN, YOLO, and SSD. Jintong Cai et al., in 2022 [3], selected the SSD technique to analyze bee behavior to detect honeybees. The result exhibited increased accuracy on multiple datasets, but the model's performance in complex scenarios is still limited. Rohit Raj et al., in 2020 [4], mentioned that Haar Cascade is advantageous over deep learning speed. In deep learning, such CNNs perform much better results at object detection, but they need a high-performance chipset like GPU or TPU for efficient inference. There are improvements in DSSD, FSSD, and ASSD [5] from SSD-based detectors. The result showed that ASSD has better the accuracy of SSD by a large margin at a small extra computation cost.

2. Proposed method

Although, the first version of service robot can detect the obstacle around itself with sensors co-working to other systems, but it cannot distinguish the obstacle types such as human and others.

According to the existing service robot function, the robot cannot appropriately response to the human or other obstacles. For example, the robot tries to avoid the human who want to interact with the robot instead of stopping itself or interacting to the user. Therefore, the visual perception is a crucial part for solving these situations. For the existing robot design, the webcam is placed on the body of the robot which is uprisen angle around 30 degree and high 120 cm. from the floor as same as the surface screen of the robot as shown in Fig 1.

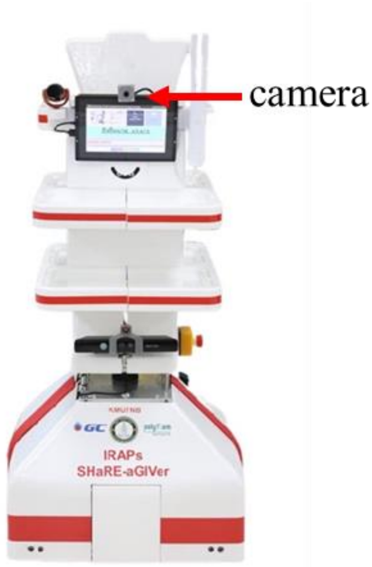


Fig.1 Service robot and webcam installed position.

In order to classify the human and other objects appearing in front of the robot, this research focuses to two main algorithm that is suitable for a real-time system. First is Haar Cascade algorithm which is the machine learning base for object detection [6]. Haar features is used for extracting the value of each subarea in the images. Each feature is calculated as a single value by subtracting the number of pixels under white and black rectangles. Fig.2 is shown the Haar features such as edge features, line features and rectangle features.

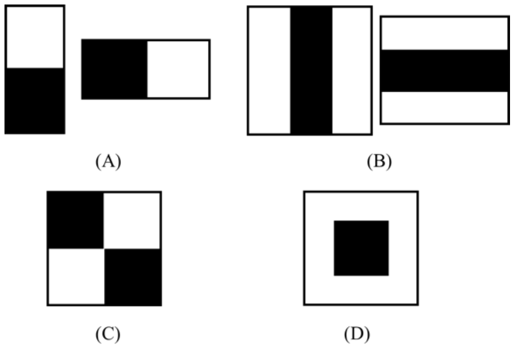


Fig.2 (A) Edge Feature, (B) Line Feature, (C) Rectangle Feature, (D) Center-Surround Feature.

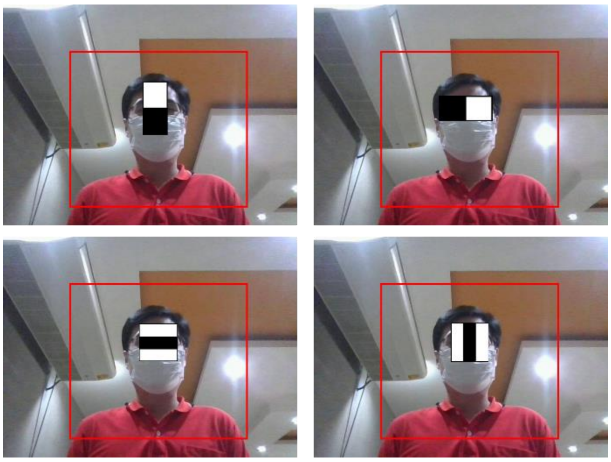


Fig.3 Example of features scanning on an image

Then, Haar features are processed as similar as the convolutional kernel. These will be scanning from top-left to the bottom right of the images for the several times with the various features of Haar as shown in Fig.3. The concept of integral images is applied for calculating rectangular features quickly. Instead of calculating at every pixel, this concept creates sub-rectangles and array references for each those sub-rectangles. Next, Adaboost training come into play for selecting the best features from the more 100,000 features. Final step is to use the cascade classifier for detecting the sub-region that contain the human face.

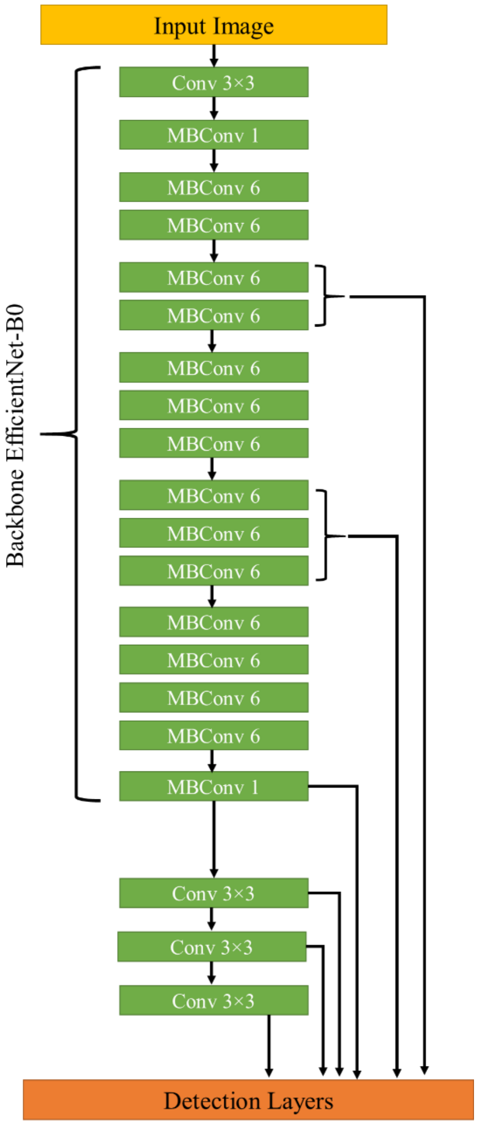


Fig.4 Single Shot Detector (SSD) of Effcientnet-B0 architecture.

Second is Single Shot Detector (SSD) which is the deep learning base [7]. SSD is a backbone network of Efficientnet-B0 as shown in Fig.4., that is adjustable the fully connected layer (FC). There are two different sizes of images between 300×300 px. and 512×512 px. applying with SSD, namely SSD300 and SSD512. Generally, the final output of SSD will show the boundary box of detected object and its category name with predicting score.

According to the original architecture, 20 classes of output are provided. This research changes the transferring trained model to 2 classes output between human and others. Moreover, the SSD300 is selected for detection.

3. Experiments and Results

Based on these two possible algorithms, which are Haar cascade and SSD, for a real-time application. This section shows the experiment for comparing the performance of human detection.

Although, both of them already have the pre-trained data of human, but the uprisen angle of a camera still has some mistakes. Therefore, this research will compare and show the improvement between these algorithms.

As the position of the camera on the robot's body, it is installed around 120 cm. from the floor and the uprisen angle about 30 degrees. In addition, the user range is set by 60 cm. from the camera approximately. Therefore, the experiment compares the accuracy of the detector algorithm between Haar cascade and SSD which are shown before and after training the images of above condition.

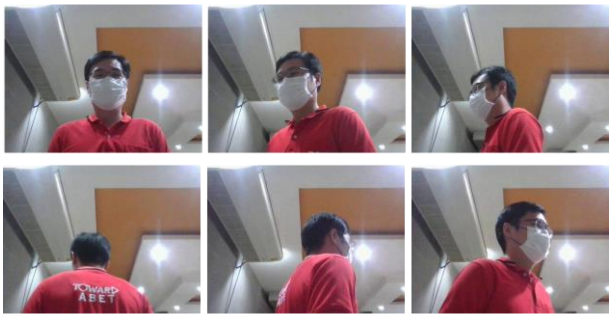


Fig.5 Training images in the various views of people.

A thousand images are collected from 9 people with the various angles such as front side, right side, left side, back side of each person as shown in Fig.5. Then, all images are trained by Haar cascade and SSD with two classes of human and other objects. The experiment tests the 99 unknown images with various side of people, the images resolution is 640×480 px.

For the Haar cascade technique, this research uses pre-trained of upper-body for detection with scale factor of 1.05 and min-neighbor of 2. Our trained model of Haar cascade uses scale factor of 1.2 and min-neighbor 2.

Moreover, the SSD of pre-trained model is used for two classes of output, compared to our trained model. The configuration of our trained model defines the IoU (intersection of union) threshold, center variance and size variance as 0.5, 0.1 and 0.2, respectively.

Table 1. The Comparison of Average IoU between pre-trained and proposed models.

	Pre-trained model	Proposed
Haar cascade	0.120	0.607
SSD	0.868	0.933

Table 1. shows the comparison results of the average IoU for human detection as the user range with uprisen angle (30 degree) of a camera between Haar cascade and SSD. Eq. 1 shows the average of IoU which N is the number of testing images which the intersection and union areas are shown in Fig.6. As the showing result, Haar cascade presents the average IoU value as 0.120 and 0.607 for upper-body (traditional model) and our proposed. In addition, SSD technique shows the average of IoU of 0.868 and 0.933 for pre-trained and proposed models.

$$Average\ IoU = \frac{1}{N}\sum_{1}^{N}\frac{Area\ of\ Intersection}{Area\ of\ Union} \tag{1}$$

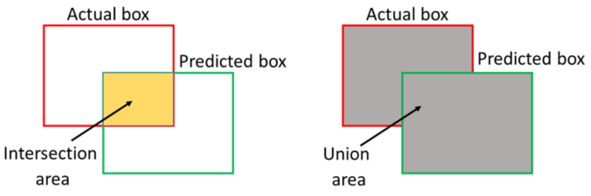


Fig.6 Intersection and Union area for actual and predicted boxes.

4. Conclusion and Discussion

This research proposes the human detection method that is feasible for the service robot with the uprisen angle of a digital camera around 30 degrees. Two possible techniques are selected for real-time detection such as Haar cascade and Single Shot Detector.

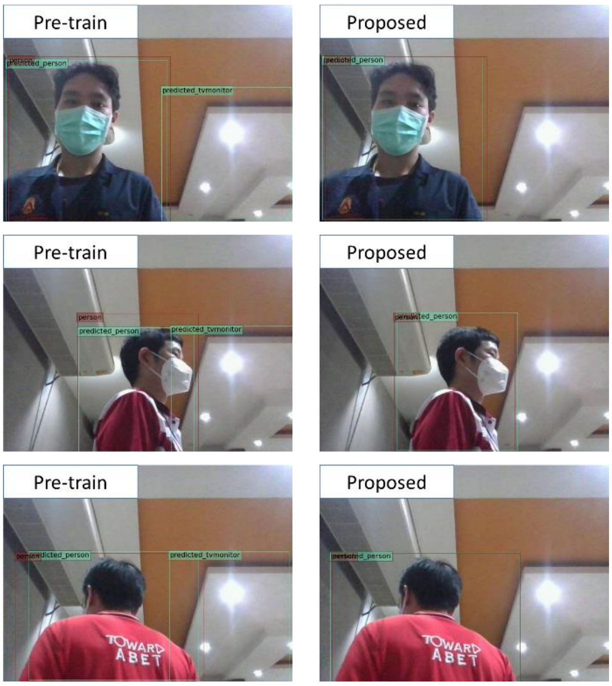


Fig.7 Example human detection between pre-trained and our proposed models of SSD.

As the results, both of Haar cascade and SSD shows the better performance in average of IoU for our proposed method. In addition, the highest performance is shown by SSD with our proposed model as 0.933 of average for IoU value. Fig.7 shows the comparison between pre-trained and our proposed model for SSD. Several times of pre-trained detection shows some mistakes, for example, tv_monitor or other classes are detected at the right side of the images while those areas have no any TV. These errors are calculated around 54% for pre-trained model, but the proposed model can extremely reduce the error. Furthermore, our proposed can detect the human with the various viewpoints.

Therefore, this research ensures that the proposed model of SSD is feasible for applying for our service robot with the uprisen angle (around 30 degrees) and 120 cm. from the based. Future work will implement to the other systems in the service robot.

Acknowledgements

This research was funded by the Ministry of Higher Education, Science, Research and Innovation, Thailand (Grant No.175613)

References

1. Sumit, Shahriar Shakir, Dayang Rohaya Awang Rambli, and Seyedali Mirjalili. "Vision-based human detection techniques: a descriptive review." IEEE Access 9 (2021): 42724-42761.
2. Shetty, Anirudha B., and Jeevan Rebeiro. "Facial recognition using Haar cascade and LBP classifiers." Global Transitions Proceedings 2.2 (2021): 330-335.
3. Cai, Jintong, et al. "Single shot multibox detector for honeybee detection." Computers and Electrical Engineering 104 (2022): 108465.
4. Raj, Rohit, et al. "Feature based video stabilization based on boosted HAAR Cascade and representative point matching algorithm." Image and Vision Computing 101 (2020): 103957.
5. Yi, Jingru, Pengxiang Wu, and Dimitris N. Metaxas. "ASSD: Attentive single shot multibox detector." Computer Vision and Image Understanding 189 (2019): 102827.
6. Paul Viola and Michael J. Jones, “Robust real-time face detection”, International Journal of Computer Vision, 57(2):137–154, 2004.
7. Liu, Wei, et al. "Ssd: Single shot multibox detector." *European conference on computer vision*. Springer, Cham, pp.21-37, 2016.

Authors Introduction

Dr. Watcharin Tangsuksant

He received B.Eng. degree in Biomedical Engineering from Srinakharinwirot University, Bangkok, Thailand in 2013. His M.Eng. degree in Biomedical Engineering from King Mongkut’s institute of technology Ladkrabang, Bangkok, Thailand in 2015. His D.Eng. degree in Department of life science and system engineering from Kyushu institute of technology, Wakamatsu campus, Japan, in 2019. He is currently a lecturer in Department of production and robotics engineering at King Mongkut’s University of Technology North Bangkok. His research interest is imageprocessing and machine learning.

Asst.Prof.Dr. Amornphun Phunopas

He received his B.S. degree in Electronics Physics in 2005 from the Faculty of Science, Thammasat University in Thailand. His M.Sc. degree in Robotics and Automation from Institute of FIeld robotics (FIBO), Thailand. His D.Eng degree in Department of Computer Science and Systems Engineering from Kyushu institute of technology, Iizuka campus, Japan, in 2012. He is currently an instructor in Department of production and robotics engineering at King Mongkut’s University of Technology North Bangkok. His research interest is robotics.

Dr. Pornthep Sarakon

He is a lecturer in the Department of Production and Robotics Engineering, King Mongkut's University of Technology North Bangkok, Thailand. He received a B.Eng. (Hons.) degree in Biomedical Engineering from Srinakharinwirot University, Nakhon Nayok, Thailand in 2015 and an M.Eng. degree in Information and Communication Technology for Embedded Systems from the Sirindhorn Inter-national Institute of Technology, Thammasat University, Thailand in 2017. He received a Ph.D. degree in Electrical and Electronic Engineering from Kyushu Institute of Technology (Kyutech), in 2021. His research interests include artificial intelligence, model compression, computer vision, and 3D body data.

Dr. Aran Blattler

He received B.Eng. degree in Production Engineering from King Mongkut’s University of Technology North Bangkok, Bangkok, Thailand in 2013. His M.Eng. degree in Materials and Production Engineering (TGGS) from King Mongkut’s University of Technology North Bangkok, Bangkok, Thailand in 2016. His D.Eng. degree in Department of Computer Science and Systems Engineering from Kyushu institute of technology, Iizuka campus, Japan, in 2021. He is currently a lecturer in Department of production and robotics engineering at King Mongkut’s University of Technology North Bangkok. His research interest is high-precision measurement and high-precision robotics.

Synthesis of Drive Systems of Flapping and Feathering Motions for Bird-like Robot using Twist Drive Mechanism

Jun Iwao
Department of Interdisciplinary Informatics, Kyushu Institute of Technology,
680-4 Kawazu, Iizuka, Fukuoka, 820-8502, Japan

Hiroshi Ohtake
Department of Intelligent and Control Systems, Kyushu Institute of Technology,
680-4 Kawazu, Iizuka, Fukuoka, 820-8502, Japan
Email: iwao.jun100@mail.kyutech.jp, hohtake@ics.kyutech.ac.jp

Abstract

In the research field of flying robots, many studies on flapping-wing aircraft have been conducted in recent years. In our previous research, we have developed a robot that mimics the musculoskeletal structure of an actual bird using a twist drive mechanism, and have achieved two types of wing motions, which are flapping and feathering, independently of each other. In this study, we have synthesized the drive systems of flapping and feathering, and have succeeded to interlocked two motions. In addition, we have covered the wings with a membrane to verify the wing motion under the condition of a membrane wing, which is strongly affected by air resistance. The flapping motion with twisting of the hands, such as actual birds do in flight, has been achieved.

Keywords: Twist Drive, Flapping robot, Feathering robot, Biomimetics

1. Introduction

Currently, a lot of research in the field of flying robots is focused on rotary-wing aircraft such as drones. These aircrafts use propellers attached to the fuselage to generate lift for flight, and are used in a wide variety of fields in our daily lives, such as transportation [1] and aerial photography [2]. Flying robots that differ from rotary-wing aircraft include fixed-wing aircraft and flapping-wing aircraft. Fixed-wing aircraft include airplanes which already play important roles in our daily lives as same as rotary wing aircraft. On the other hand, flapping-wing aircraft have been the focus of attention in recent years, but compared to the two types of flying robots mentioned above, there are still few examples of research and very only a few examples of use.

Flapping-wing robots mainly imitate the motions of actual flying creatures, such as birds and insects. In this research, we focus on a bird-like flapping wing aircraft. In research on bird-like robots, most of them use crank mechanisms or servo motors in the body of the robot to achieve the wing motion. The crank mechanism converts the rotational motion of the motor into the swinging motion of the wings via gears. In the case of servo motors, the wing is directly attached to the rotary shaft of the servo motor, and the angle of rotation of the servo motor is adjusted to achieve the swinging motion of the wing. However, actual birds move their wings by contracting and relaxing their muscles, and the mechanisms of the above two types of mechanisms are very different from those of actual birds. Therefore, in this study, we focused on the imitation of the bird's motions by employing the twist drive mechanism [3].

The twist drive mechanism utilizes the tensile force generated when twisting a thread. As shown in Fig.1, the rotating shaft of the motor is connected to the object to be moved by a thread, and rotating the motor causes the thread to twist and contract, pulling the object. Rotating the motor in reverse causes the thread to untwist, weakening the pulling force. The feature of being able to control the movement of an object by contraction and relaxation of a thread is very similar to the actual muscular motion of a bird. In our previous research [4], we have developed a robot that mimics the musculoskeletal structure of a real bird using a twist drive mechanism, and confirmed that flapping can be achieved. In addition, birds perform various wing motions in flight in parallel with flapping, such as feathering. Feathering is a twisting action of the wing tip hand, and has the role of increasing lift and thrust in flight and reducing aerodynamic resistance by changing the angle of attach. In our research [5], we tried to achieve feathering in addition to flapping by using a twist-drive mechanism, and succeeded in developing a robot that can achieve the two types of motions independently.

In this paper, we integrate and link the flapping and feathering drive systems, which were independently achieved in the literature [5], to realize a flapping motion with a twisting motion of the hand part, as actual birds do in flight. Next, a membrane is attached to the wing, and the motion is verified under the condition of a membrane wing, which is strongly affected by air resistance.

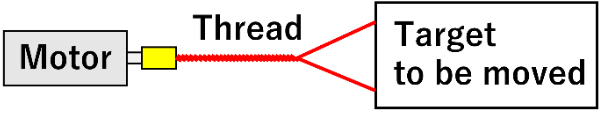


Fig.1.Twist drive mechanism

2. Achievement of flapping with feathering

2.1. *Interlocking flapping and feathering drive systems*

A CAD drawing of the developed flapping wing robot is shown in Fig.2. Regarding the motors used in the twist drive mechanism, two ECX SPEED 13M brushless motors made by Maxon were used for flapping, and two geared motors made by Pololu were used for feathering [5]. Motors for flapping are controlled using the MAXON EPOS4 Compact50/5 control unit with Windows PC. Motors for feathering are controlled using the Arduino Uno and the TA7291P motor drivers.

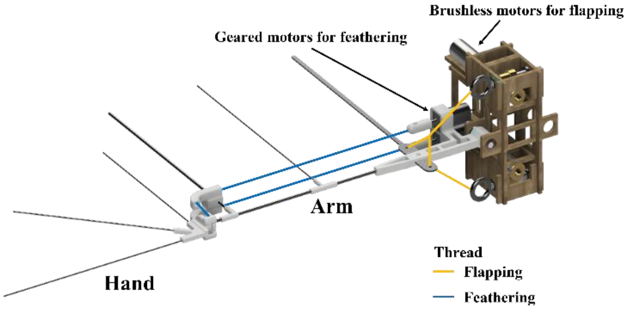


Fig.2.CAD model of the developed flapping wing robot

The wings of the flapping wing robot are divided into arm and hand sections, and only the hand parts are moved for feathering. Two motors of the twist drive mechanism for feathering are fixed at the base of the wing, and two motors for flapping are fixed inside the body of the robot. Focusing on the parameters of the wing only, the wing length is 380.00 mm, the wing chord length is 161.00 mm, and the wing weight, including the weight of the two motors for feathering, is 29.94 g. Fig.3 shows the developed flapping wing aircraft flapping and Fig.4 shows feathering. The flapping frequency for flapping was 1 Hz, and the maximum angle for both swinging up and down was 34 °. Fig.4 shows the image of twisting the front edge of the hand (the front side of Fig.4) in the upward and downward directions. The direction of the front edge of the hand can be switched to the upward and downward directions in as little as about 0.13 seconds.

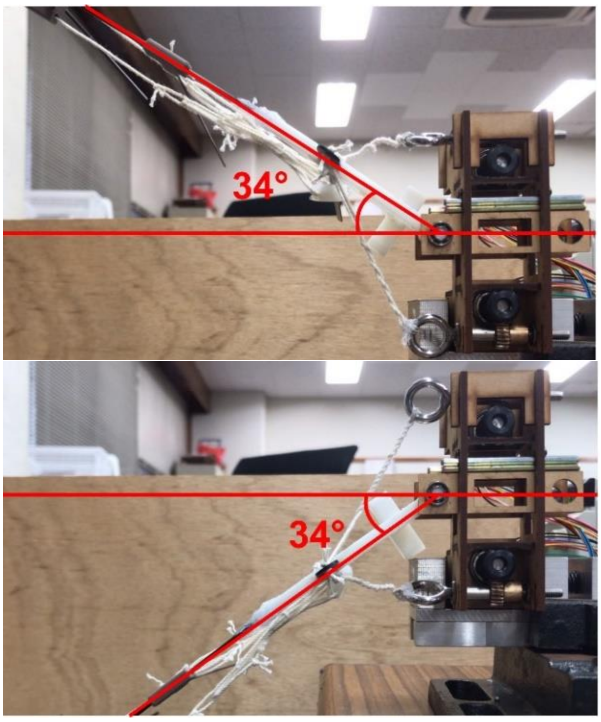


Fig.3.Flapping in progress

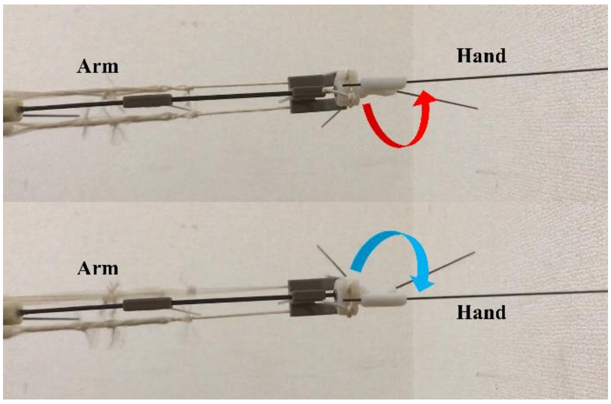


Fig.4.Feathering in progress

In our previous research [5], each operation in Fig.3 and Fig.4 was performed separately. In this research, we tried to achieve a combined flapping and feathering motion by linking the drive systems of the flapping and feathering motions through serial communication. Referring to the actual flapping behavior of birds, we decided that the timing of feathering when combining the two types of motion should be performed at the moment when the wings switch between flapping up and down. The flow of the program is as shown in the flowchart in Fig.5. First, immediately after the wing is swung up, the values for twisting the front edge of the hand in the downward direction are sent to the Arduino Uno via serial communication. Then, immediately after the wing is swung down, the values for twisting the front edge of the hand in the upward direction are sent to the Arduino Uno via serial communication in the same manner. The above process assumes that the front edge of the hand turns upward while the wing is being swung up, and that the front edge of the hand turns downward while the wing is being swung down. By repeating the process in a series of steps, we aimed to achieve flapping with feathering.

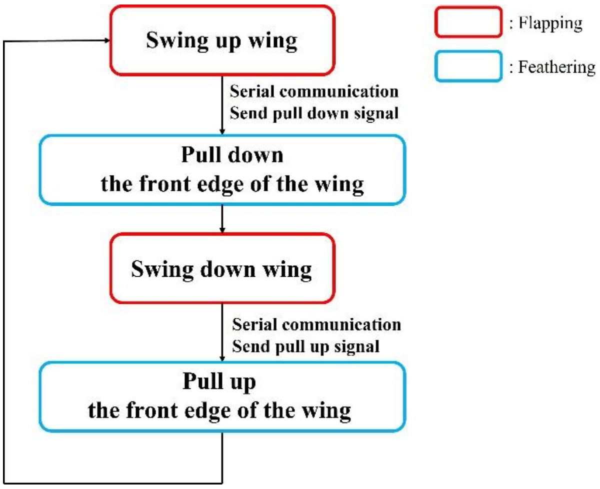
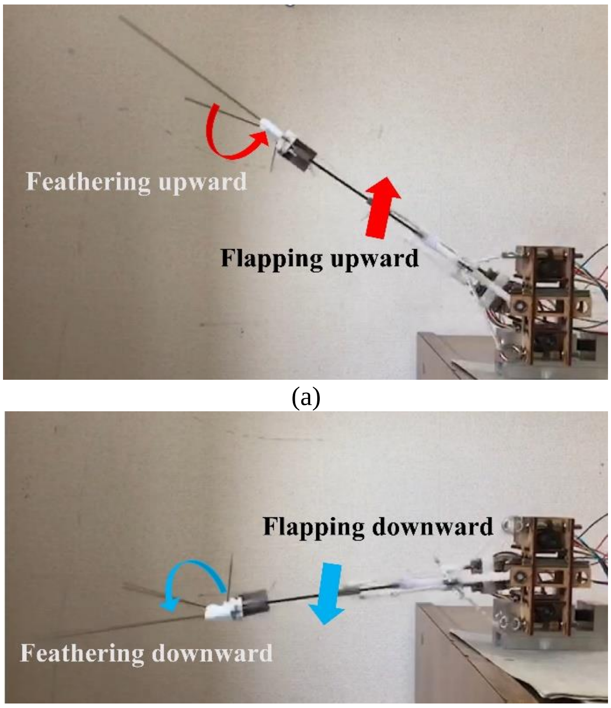


Fig.5.Program flowchart for linking flapping and feathering

2.2. *Operation verification results*

A program according to the flowchart in Fig.5 was created and the operation was verified. Fig.6(a) shows the hand part while swinging up the wing, and it was confirmed that the front edge of the hand part was turning upward. Similarly, Fig.6(b) shows the hand with the wing swinging down, and it was confirmed that the front edge of the hand was turning downward. Thus, we have succeeded to link each drive system of flapping and feathering using serial communication.



(a)

(b)

Fig.6.View of the hand section when the wing is flapping up and down

3. Verification of Operation with Membrane Wing

3.1. *Membrane Attachment to a Skeleton Wing*

In the operation verification in Section 2, the wing of the flapping-wing aircraft was only the frame, and the effect of air resistance was very small. Since the flapping wing machine is a flying robot, achieving autonomous flight is one of its future goals. Therefore, it is necessary to have wings that can generate lift and thrust, which are important for flight, and that can flap even under the condition of high air resistance. Therefore, in this study, we attached a membrane to the wing and verified whether flapping and feathering are possible even when the wing is strongly affected by air resistance. Fig.7 shows the flapping wing aircraft with the membrane attached. Ripstop fabric was used for the wing membrane. The wing membrane is attached only to the upper surface of the wing, and the hand and arm parts are separated at the joints so as not to interfere with feathering. The weight of the membrane wing, including the weight of the two motors for feathering, was 32.90 g, an increase of 2.96 g from the skeleton wing state in Section 2.

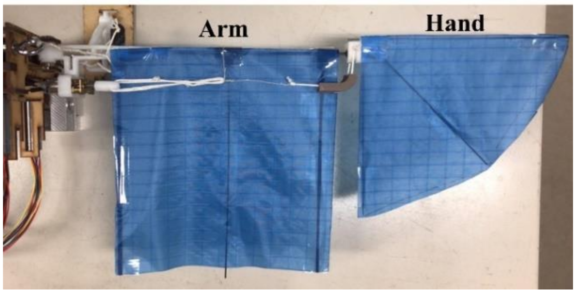
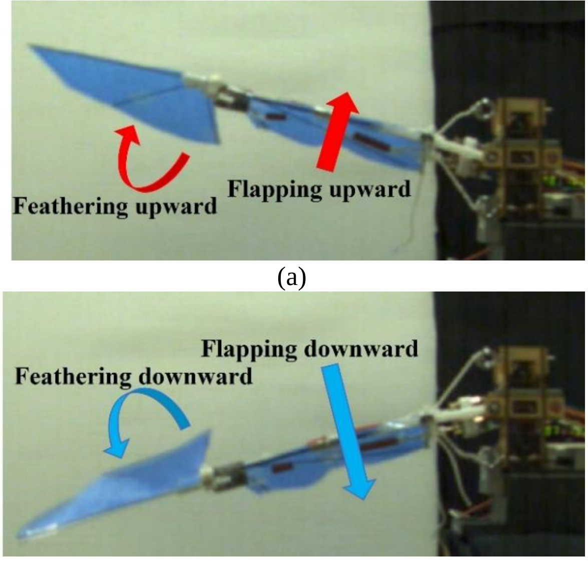


Fig.7.Flapping wing aircraft with membrane on the wing

3.2. *Operation verification under membrane wing*

Fig.8 shows the results of the operation verification in the state shown in Fig.7. Fig.8(a) shows that the front edge of the hand is turned upward while the wing is being swung up. Similarly, Fig.8(b) shows that the front edge of the hand was intentionally directed downward while the wing was being swung down. The flapping motion with twisting of the hands has been achieved.



(a)

(b)

Fig.8.View of the hand section when the membrane wing is flapping up and down

4. Conclusion

In this paper, we have linked the flapping and feathering drive systems to achieve flapping motion with feathering. We have also verified the operation of the membrane wing, which is strongly affected by air resistance, and have confirmed that it performs as well as the skeleton wing. Our future work is to measure lift and thrust forces of the membrane wing.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number JP20K04364.

References

1. Robin Kellermann, Tobias Biehle, Liliann Fischer, "Drones for parcel and passenger transportation: A literature review", Transportation Research Interdisciplinary Perspectives, 2020, 4, 100088.
2. Scott Hamilton, Jason Stephenson, "Testing UAV (drone) aerial photography and photogrammetry for archaeology", Technical Report, Lakehead University, 2016.
3. Ivan Godler, Takashi Sonoda, Kazuo Sakurai, "Modeling and Evaluation of a Twist Drive Actuator for Soft Robotics", Advanced Robotics, Vol.26, No.7, pp.765-783, 2012.
4. Ryuichiro Yoshinaga and Hiroshi Ohtake, "Realization of Flapping Motion Using Twist Drive Mechanism for Bird-like Robot", Proceedings of the 2019 JSME Conference on Robotics and Mechanism, 2A2-R10, 2019, in Japanese.
5. Jun Iwao and Hiroshi Ohtake, "Development of flapping robot achieving both flapping and feathering motions using twist drive mechanism", Proceedings of the 2022 SICE Annual Conference, 2022, 855-860.

Authors Introduction

Mr. Jun Iwao

He received his bachelor's degree from Kyushu Institute of Technology, Japan, in 2021. He is currently a graduate student at Kyushu Institute of Technology, Japan. His research interests include Flying robots.

Dr. Hiroshi Ohtake

He received his PhD degree from The University of Electro-Communications, Japan, in 2006. He is currently an Associate Professor, Kyushu Institute of Technology, Japan. His research interests include Flying Robot, Intelligent Control and Brain Machine Interface (BMI).

Optimization Algorithm for Balancing QoS Configuration in Aggregated Robot Processing Architecture

Abdul Jalil[*], **Jun Kobayashi, Takeshi Saitoh**
Kyushu Institute of Technology,
680-4 Kawazu, Iizuka-shi, Fukuoka, 820-8502, Japan
E-mail: malla.abdul-jalil545@mail.kyutech.jp, jkoba@ics.kyutech.ac.jp, saitoh@ai.kyutech.ac.jp.
www.kyutech.ac.jp

Abstract

Quality of Service (QoS) manages the data traffic to reduce packet loss, latency, and jitter in the network. This study aims to design an optimization algorithm to find the balance of QoS configuration to set the rates and buffer size while the robot data processes are communicated in the Aggregated Robot Processing (ARP) architecture. This study implements optimization to manage the DEPTH and DEADLINE QoS configuration in Robot Operating System 2 (ROS 2) node communication. Unbalancing DEPTH and DEADLINE configurations can affect the high latency time of message data transmission and packet loss in RELIABLE connections. The results of this study show that the optimization algorithm can determine the optimal value of DEPTH and DEADLINE by balancing the QoS configuration to improve the robot data transmission in the ARP architecture.

Keywords: Optimization, Quality of Service, ROS 2, Aggregated Robot Processing

1. Introduction

Robot data processing flows generally have three components: sensing, planning, and actuation [1]. These components can be connected as node communication in the network and use Quality of Service (QoS) policies to manage the quality of data transmission between components. QoS manages data traffic on the network to reduce packet loss, latency, and jitter. However, the unbalanced QoS configuration on the network can influence the performance of robot data transmission between components, such as packet loss and latency.

This study aims to develop an optimization algorithm to find the balancing QoS configuration of the rates and buffer size while the robot data processes are communicated in the Aggregated Robot Processing (ARP) architecture. We implement this optimization to find the optimal value of DEPTH and DEADLINE when the robot data processes between sensing, planning, and actuation components are communicated using ROS 2 (Robot Operating System 2). ROS 2 is built on top of Data Distribution Service (DDS) and uses a set of QoS policies to tune node communication [2].

Some researchers have analyzed the effectiveness of QoS policies configuration in ROS 2. They analyzed latency [2][3][4][5][6], throughput [2][4], packet loss [4][5][6][7], and memory consumption [2]. Several studies show that when DEPTH configures the buffer with a small size and DEADLINE configures the rates with high frequency, some of the packets will be lost in ROS 2 node communication [4][5][6][7]. Furthermore, when DEPTH configures the buffer with a large size, some memory space will be used in that configuration [4][8], and when DEADLINE sets the rates with a low frequency, the data transfer rate becomes low, and this will affect the real-time of the message data transfer between nodes.

The contribution of this paper is to improve robot data transmission by balancing the DEPTH and DEADLINE QoS configuration when the ROS 2 nodes transmit the data using the RELIABLE and KEEP_LAST options. We analyze it because strict reliability is not guaranteed if the DDS uses the RELIABLE and KEEP_LAST options for data transmission between nodes [9].

2. Methods

Figure 1 shows the ARP architecture developed in our study [10]. In ARP architecture, the robot data processes for localization and path planning are executed in the Computer Environment Dedicated to Data Processing (CEDDP). The robot computer function reads the sensor data in the sensing component and then sends them to the CEDDP, which also drives the robot actuator based on the path-planning result sent from the CEDDP in the actuation component. The robot computer and CEDDP can exchange data through a wireless network.

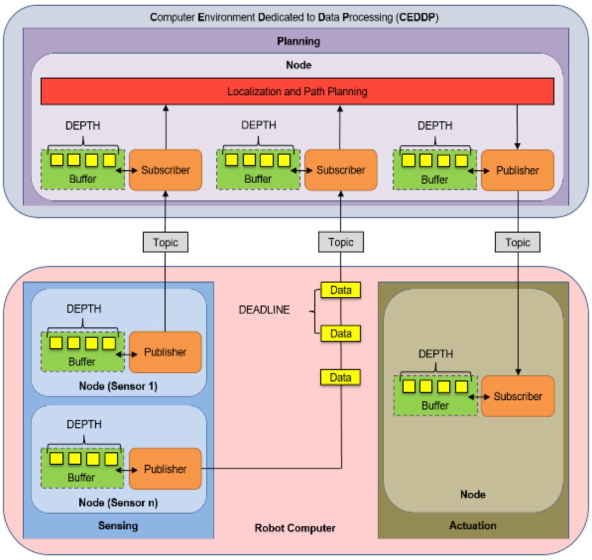


Fig.1.Aggregated Robot Processing architecture

Based on the illustration shown in Fig.1, the sensor node in the sensing component transmits sensor data to a node in the planning component through a topic SPt. Then the node in the planning component transmits the robot localization and path planning result to a node in the actuation component through a topic PAt. Therefore, our idea here to find the optimal value of DEADLINE R is to divide the maximum data transmission rate $Rmax$ by the total topic used to transmit sensor data from a node in the sensing component to a node in the planning component $\sum SPt$, added with the total topic used to transmit the data from a node in the planning component to a node in the actuation component $\sum PAt$.

$$R = \frac{Rmax}{\sum SPt + \sum PAt} \tag{1}$$

In Eq. (1) we divide $Rmax$ by $\sum SPt + \sum PAt$ to balance the rate configuration in DEADLINE R with all topics when transmitting the message data from the publisher to the subscriber. In this study, $Rmax$ is the maximum rate when only one topic is used to transfer message data from the publisher to the subscriber. Furthermore, based on the idea of Eq. (1) we create the first constraint in our optimization with the following:

$$\frac{Rmax}{\sum SPt + \sum PAt} - R \geq Rmin \tag{2}$$

In Eq. (2) the value of R should be greater than or equal to the minimum transmission rate $Rmin$, which means that the communication of message data between the publisher and the subscriber is satisfied when R is greater than or equal to $Rmin$.

Next, find the optimal value of DEPTH D to determine the buffer size. Our idea here to find the optimal value of the DEPTH is to balance it with a DEADLINE tune. If the DEADLINE tune is large and close to the maximum rate of the DEADLINE, the DEPTH tune will also be high and close to the maximum value of DEPTH $Dmax$. Otherwise, if the DEADLINE tune is low, the DEPTH tune will also be low and close to the minimum value of the DEPTH $Dmin$. Here, $Dmax$ is the maximum queue size for storing data samples in the buffer when the KEEP_LAST option is chosen, that is, 5000 [8]. Based on this idea, we create the second constraint to find the optimal value of DEPTH with the following:

$$Dmax \frac{R}{Rmax} - D \geq Dmin \tag{3}$$

For the next constraints, bound the variable R not to be greater than or equal to the maximum rate and not less than or equal to the minimum rate $Rmin \leq R \leq Rmax$. After that, bound the variable D not to be greater than or equal to the maximum DEPTH and not less than or equal to the minimum DEPTH $Dmin \leq D \leq Dmax$. Finally, create the optimization equation to find the optimal value of DEPTH and DEADLINE with the following:

$$\begin{aligned} & max\ R + D \\ s.t.\quad & \frac{Rmax}{\sum SPt + \sum PAt} - R \geq Rmin \\ & Dmax \frac{R}{Rmax} - D \geq Dmin \\ & Rmin \leq\ R \leq Rmax \\ & Dmin \leq\ D \leq Dmax \end{aligned} \tag{4}$$

In this study, we used CVXPY to implement the optimization algorithm. CVXPY is an open-source Python-embedded modeling language that solves the problem of convex optimization [11]. The following algorithm is our proposed optimization to find the optimal value of DEPTH and DEADLINE based on the optimization shown in equation 4.

```
Algorithm: Optimization algorithm
Require: $\sum SPt, \sum PAt, Rmax, Rmin, Dmax, Dmin$
Variable: $R, D$
Constraints:
    [$\frac{Rmax}{\sum SPt + \sum PAt} - R \geq Rmin$,
    $Dmax \frac{R}{Rmax} - D \geq Dmin$,
    $Dmin \leq R \leq Dmax$,
    $Dmin \leq D \leq Dmax$]
Objective_Function: $max\ R + D$
Problem: (Objective_Function, Constraints)
Problem.solve()
```

3. Results and Discussion

Figure 2 shows the experimental illustration of our study. To perform the analysis, we used Raspberry Pi 4, which has a processor of 1.5GHz with 8 GB memory for the robot computer, and a laptop computer with an Intel Core i5 @ 2.60GHz x 4 with 12 GB memory as a CEDDP. The operating system installed on these computers was Linux Ubuntu 20.04 LTS, Foxy Fitzroy for the ROS 2 distribution, and Fast-RTPS for the DDS middleware. Furthermore, we used a wireless access point with a frequency of 2.4 GHz to exchange the message data between the robot computer and CEDDP.

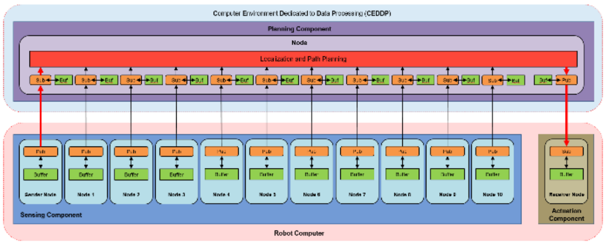


Fig.2. Experimental illustration

Based on the illustration shown in Fig.2, we analyze the performance of our optimization when the node between the sensing, planning, and actuation components is communicated based on the QoS configurations shown in Table 1.

Table 1. QoS configurations in experiments.

RELIABILITY	RELIABLE
HISTORY	KEEP_LAST
DEPTH	1, 5000, Optimization (D)
DEADLINE	100 Hz, 200 Hz, 500 Hz, 1000 Hz, Optimization (R)
DURABILITY	VOLATILE
LIVELINESS	AUTOMATIC

In the experiment, we evaluated the optimization algorithm if each node in the sensing, planning, and actuation component exchanges the message data with a size of 10 bytes, 100 bytes, and 1000 bytes. Furthermore, Table 2 shows the requirement value entered into the optimization algorithm.

Table 2. Input value to optimization algorithm.

	1^{st} Exp	**2^{nd} Exp**	**3^{rd} Exp**	**4^{th} Exp**
$\sum SPt$	11	11	11	11
$\sum PAt$	1	1	1	1
$Rmax$	100 Hz	200 Hz	500 Hz	1000 Hz
$Rmin$	1 Hz	1 Hz	1 Hz	1 Hz
$Dmax$	5000	5000	5000	5000
$Dmin$	1	1	1	1
Opt (D)	365	390	405	410
Opt (R)	7 Hz	15 Hz	40 Hz	82 Hz

Based on the information shown in Fig.2 and Table 2, $\sum SPt = 11$ is the total of topics used to transmit data from 11 nodes in the sensing component to a node in the planning component. $\sum PAt = 1$ is the total number of topics used to transmit data from a node in the planning component to a node in the actuation component. $Rmax$ = 100 Hz, 200 Hz, 500 Hz, and 1000 Hz is the maximum rate when only one topic is used to transmit the data between the publisher and the subscriber, respectively. $Rmin$ = 1 Hz is the minimum data transmission rate, $Dmax = 5000$ is the maximum queue to store the data sample in the buffer, and $Dmin = 1$ is the minimum queue in the buffer. Furthermore, Opt (D) and Opt (R) are the optimization results of DEPTH D and DEADLINE R.

In this study, we analyze the optimization efficiency by measuring the latency time of message data transmission and calculating the packet loss while the sender node in the sensing component sends the message data to the receiver node through a node in the planning component. We measured the latency by calculating the time that elapsed until the receiver node in the actuation component received the message data sent from the sender node. Figure 3 shows the latency result in this experiment.

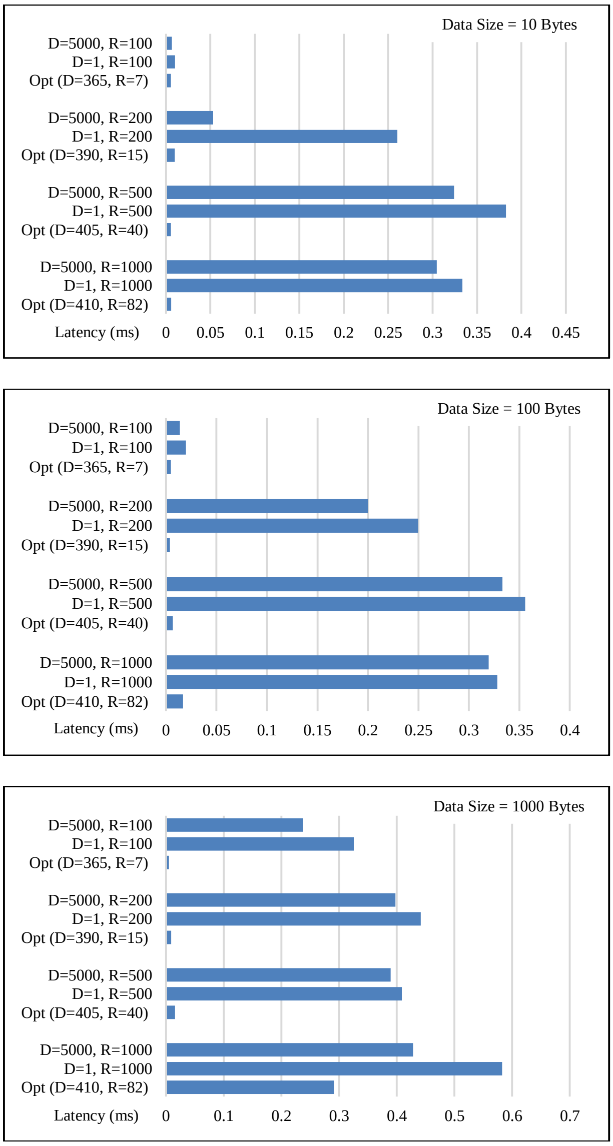


Fig.3. Latency results.

Based on the analysis results shown in Fig.3, the balance of the QoS configuration can improve the latency of message data transmission compared to when the node transmits the data with the maximum rate in the DEADLINE and the maximum/minimum configuration in the DEPTH. Furthermore, Fig.4 shows the packet loss when the sender node transmits the data to the receiver node through a node in the planning component.

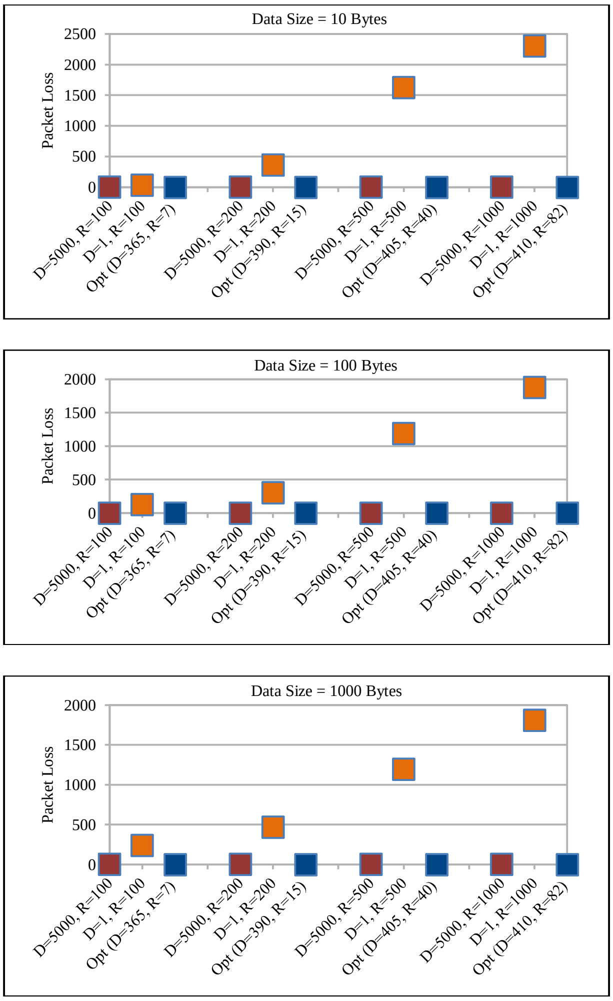


Fig.4. Packet loss results.

Based on the information results shown in Fig.4, the unbalanced QoS configuration increased the packet loss in node communication. Furthermore, the balance of QoS configuration can reduce packet loss.

4. Conclusion

This study has designed an optimization algorithm to find the balance of DEPTH and DEADLINE QoS configuration to improve ROS 2 node communication in the ARP architecture. The balance of DEPTH and DEADLINE QoS policies shown in this study can improve the latency time of message data transmission and reduce packet loss in a RELIABLE connection. Next, we will implement this optimization algorithm to improve the multi-robot data transmission in ARP architecture.

References

1. J. Staschulat, I. Lütkebohle, and R. Lange, "The rclc Executor: Domain-specific deterministic scheduling mechanisms for ROS applications on microcontrollers: work-in-progress", International Conference on Embedded Software, EMSOFT, IEEE, 2020. DOI: 10.1109/EMSOFT51651.2020.9244014.
2. Y. Maruyama, S. Kato, and T. Azumi, "Exploring the Performance of ROS2", International Conference on Embedded Software, EMSOFT, IEEE, 2016. DOI: 10.1145/2968478.2968502.
3. T. Kronauer, J. Pohlmann, M. Matthé, T. Smejkal, and G.~Fettweis, "Latency Analysis of ROS2 Multi-Node Systems", International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI), IEEE, 2021. DOI: 10.1109/MFI52462.2021.9591166.
4. J. Fernandez, B. Allen, P. Thulasiraman, and B. Bingham, "Performance Study of the Robot Operating System 2 with QoS and Cyber Security Settings", International Systems Conference (SysCon), IEEE, 2020. DOI: 10.1109/SysCon47679.2020.9275872.
5. P. Thulasiraman, Z. Chen, B. Allen, and B. Bingham, "Evaluation of the Robot Operating System 2 in Lossy Unmanned Networks", International Systems Conference (SysCon), IEEE, 2020. DOI: 10.1109/SysCon47679.2020.9275849.
6. Z. Chen, "Performance Analysis of ROS 2 Networks Using Variable Quality of Service and Security Constraints for Autonomous Systems", Naval Postgraduate School, 2019.
7. J. Park, R. Delgado and B.W. Choi, "Real-Time Characteristics of ROS 2.0 in Multiagent Robot Systems: An Empirical Study", IEEE Access, vol. 8, pp. 154637-154651, 2020. DOI: 10.1109/ACCESS.2020.3018122.
8. eProsima, "Fast DDS Documentation", Release 2.8.1, December 2022.
9. https://community.rti.com/static/documentation/connext-dds/5.2.0/doc/manuals/connext_dds/html_files/RTI_ConnextDDS_CoreLibraries_UsersManual/Content/UsersManual/HISTORY_QosPolicy.htm.
10. A. Jalil and J. Kobayashi, "Experimental Analyses of an Efficient Aggregated Robot Processing with Cache-Control for Multi-Robot System", 20th International Conference on Control, Automation and Systems (ICCAS), IEEE, pp. 1105-1109, 2020. DOI: 10.23919/ICCAS50221.2020.9268225.
11. S. Diamond and S. Boyd, "CVXPY: A Python-Embedded Modeling Language for Convex Optimization". Journal of Machine Learning Research, Vol. 17, Num 83, pp. 1-5, 2016.

Authors Introduction

Mr. Abdul Jalil

He received Master's and Bachelor's degree from the Department of Computer Systems, STMIK Handayani Makassar, Indonesia in 2015 and 2012, respectively. Currently, he is a Doctoral student at Department of Computer Science and Systems Engineering, Kyushu Institute of Technology, Japan.

Dr. Jun Kobayashi

He is an Associate Professor at Department of Intelligent and Control Systems, Kyushu Institute of Technology. He obtained his PhD degree at Kyushu Institute of Technology in 1999. Currently, his main interests are Robotics, Physiological Computing, and Cybernetic Training.

Prof. Takeshi Saitoh

He is Professor of Department of Artificial Intelligence, Kyushu Institute of Technology. He received B.Eng., M.Eng., and D.Eng. degrees from Toyohashi University of Technology, Japan. His research field includes image processing and pattern recognition.

Research on robotic assembly of gear motors (Stator recognition using keypoint matching and stator insertion using contact position estimation)

Yasumoto Imai, Takayoshi Yamada, Junya Sato
Department of Mechanical Engineering, Gifu University, Yanagido 1-1, Gifu 501-1193, Japan

Toshiki Hayashi, Shota Aono
Tsubakimoto Chain Co., Nakanoshima 3-3-3, Kita-ku, Osaka 530-0005, Japan

E-mail: yamat@gifu-u.ac.jp, jsato@gifu-u.ac.jp

Abstract

This paper describes a robotic assembly system for gear motors. Currently, automation through the introduction of robots is being actively implemented in various fields. However, assembly tasks, which require dexterous skills, are still often performed manually. Humans use information such as vision and hand sensation to accomplish these tasks. Therefore, research is being conducted to develop assembly robots using cameras and force sensors. In this paper, two main methods, key point matching and contact position estimation, are used to realize the gear motor assembly work, which is the insertion of the stator into the reducer. Keypoint matching is used to recognize the position and orientation of the part using a camera. Contact position estimation is used to detect contact between parts in the insertion process using a force sensor to prevent failure of the insertion operation. An experimental system using these methods is developed to achieve automatic stator insertion.

Keywords: Robotic assembly, Keypoint matching, Contact point estimation, Gear motors, Image sensor, 6-axis force sensor

1. Introduction

A gearmotor (Fig. 1) is a motor packaged with a reduction gear connected to the rotating shaft of a DC or AC motor and is widely used in cranes and transport equipment. Currently, gearmotors are assembled manually, which requires high level skills, such as the sense of dexterity and visual information of skilled workers. Tanaka et al. [1] discussed a grasping system of a small nut, based on force or current measurements. Satoh et al. [2] discussed precision assembly tasks of a polygon mirror, based on forces and moments and the orientation of the part.

This paper describes a method of automatic stator insertion in the assembly sequence of gearmotors, on the basis of which the rotor is assembled in the reduction gear unit. Section 2 describes an overview of the automatic assembly equipment as a hardware configuration. Section 3 describes key point matching for stator angle recognition as a vision system and contact position estimation for preventing failures in insertion operations. Section 4 then examines the effectiveness of the proposed automated assembly system using the methods in Section 3.

Fig. 1 Gearmotor

2. Setup of the experiment system

In this section, we describe our experiment system for this study. An articulated robot arm Denso VS-050 is used for assembling parts. A 6-axis force sensor Leptrino PFS055YA251, an RGB-D camera Intel RealSense D435, a 2-finger robot hand Robotiq 2F-140 are mounted on the end of the robot arm (Fig. 2). A reducer and motor manufactured by Tsubakimoto Chain Co. is used as the assembly target, and experiments is conducted in the assembly environment shown in Fig. 3. The reduction gear is set on a fixture created by a 3D printer. In this environment, the stator is grasped and assembled into the reducer.

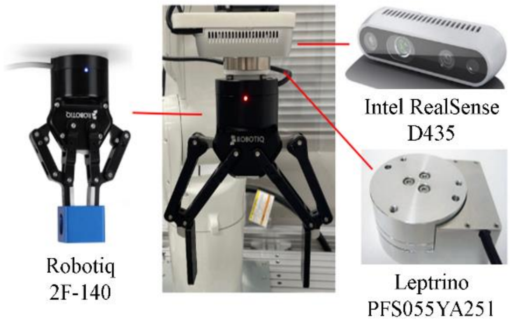


Fig. 2 Handling System

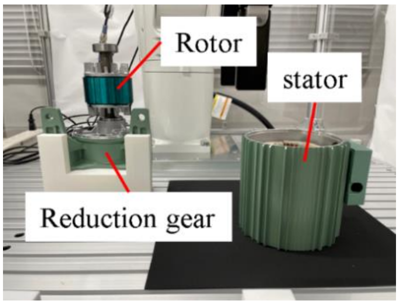


Fig. 3 Assembly Environment

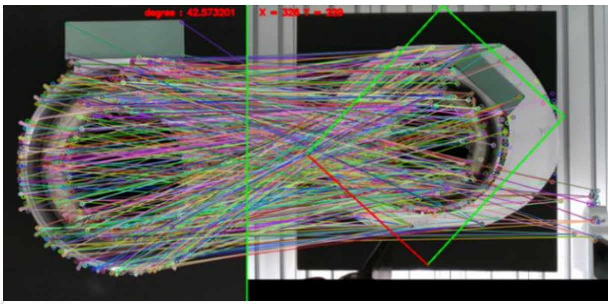
3. Assembling method

3.1. *Keypoint matching*

Keypoint matching first extracts keypoints from each of the template and target images that are robust to translation, scaling, rotation, and brightness changes. See Ref. [3] for details. This study uses Hamming distance to compute similarity. If both character strings are x and y, the Hamming distance $\mathrm{Ham}(x,y)$ can be defined by the following Eq. (1).

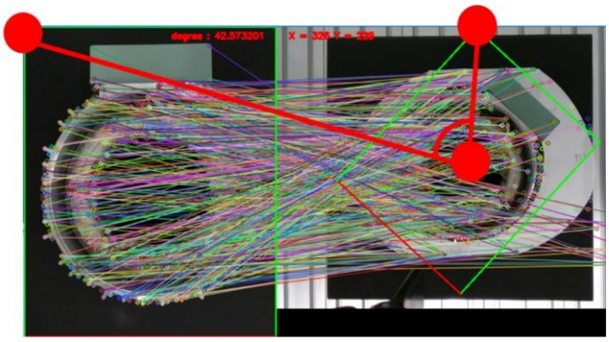
$$\mathrm{Ham}(x,y) := \sum_{i=1}^{m} \begin{cases} 1 & (x[i] \neq y[i]) \\ 0 & (x[i] = y[i]) \end{cases} \qquad (1)$$

Next, the top k keypoints with the highest similarity among the feature points extracted from each of the template and target are matched (Fig. 4). This is called k-NN matching. The affine transformation matrix (2 × 3) is estimated by RANSAC from a set of keypoints extracted from both the template and target images. The angle is then calculated from a total of three points: the center coordinates of the bounding rectangle of the target image and the upper left vertex coordinates before and after affine reconstruction (Fig. 5). The purpose of using this method is because the stator must be inserted at an angle that matches the threaded holes in the reducer and stator.



(a) Template Image (b) Target Image

Fig. 4 Keypoint Matching



(a) Template Image (b) Target Image

Fig. 5 Angle Detection Image

3.2. *Estimation of the contact point*

Since stator insertion is not always smooth, the optimal insertion position must be detected. For this purpose, the optimal insertion position is detected by repeating the corrective action of moving in the direction to avoid from the contact position using contact position estimation. Force and moment $(\boldsymbol{f}_i, \boldsymbol{m}_i)$, $(i = 1,2,\cdots,k)$ are measured from the 6-axis force sensor. In order to eliminate the offset of the force sensor, the deviation force $(\bar{\boldsymbol{f}}_i, \bar{\boldsymbol{m}}_i)$, $(i = 1,2,\cdots,k)$ from the average are utilized. Let $\boldsymbol{p}_c$ be the contact point position, we have $\bar{\boldsymbol{m}}_i = \boldsymbol{p}_c \times \bar{\boldsymbol{f}}_i$ in pure and ideal case. Considering that the measurement force and moment is contaminated with noise, the following evaluation function $J(\boldsymbol{p}_c)$ is defined by Eq. (2).

$$J(\boldsymbol{p}_c) := \frac{1}{k} \sum_{i=1}^{k} \left\| \bar{\boldsymbol{m}}_i - \boldsymbol{p}_c \times \bar{\boldsymbol{f}}_i \right\|^2 \qquad (2)$$

Minimizing the function, the position $\boldsymbol{p}_c$ is estimated by the following formula:

$$\widehat{\boldsymbol{p}}_c = \left(\frac{1}{k}\sum_{i=1}^{k}[\bar{\boldsymbol{f}}_i \times]^2\right)\left(\frac{1}{k}\sum_{i=1}^{k}\bar{\boldsymbol{m}}_i \times \bar{\boldsymbol{f}}_i\right) \qquad (3)$$

The function $J(\boldsymbol{p}_c)$ has the minimum value $s_k := J(\widehat{\boldsymbol{p}}_c)$. Uncertainty region of the estimate $\widehat{\boldsymbol{p}}_c$ is defined as the region satisfying the following condition.

$$J(\boldsymbol{p}_c) \leq 2s_k \qquad (4)$$

The region is given by an ellipsoid centered at the estimate. See Ref. [4] for the derivation of the region.

4. Experiments

4.1. *Procedure of experiments*

We perform the following steps.

Step 1: Recognize the stator by Hough transform.
Step 2: Recognize the angle of the stator by keypoint matching.
Step 3: The robot hand grasps the stator and moves it to just above the reducer.
Step 4: Perform the insertion operation. If contact is made with the rotor, then corrective action is taken until the proper insertion position is reached.
Step 5: Once the proper position has been detected, the insertion should be completed.

The details of **Steps 4** and **5** are shown in Fig. 6.

Success condition: After inserting the stator, screws can go through both screw holes of stator and reducer manually.

Number of trials: 10 trials for each of the 8 conditions (0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°) where the stator contacts the rotor as seen in the robot coordinates, 80 trials in total. The robot coordinates and angular orientation are shown in Fig. 7. In the initial condition of the experiments, the stator is shifted in advance about 2 mm from the nominal aligned position.

4.2. *Results of the experiments*

Experimental success rate was 100 % in all conditions. Table 1 shows the average time when the condition is 45°. Fig. 8 and Fig. 9 show the estimates of the contact position and its uncertainties of XY components are also shown in the case of 90° Fig. 10 shows force measurement of z component. The sampling time of the force sensor is 18.6 msec and the last 20 data ($k = 20$) are used in Eq. (3). From Fig. 10, the contact repeatedly appears in the samples from 226 to 316. If the uncertainty is small, the estimate $\widehat{\boldsymbol{p}}_c$ is valid. Note that Figures 8, 9 and 10 focus on the valid range of estimation.

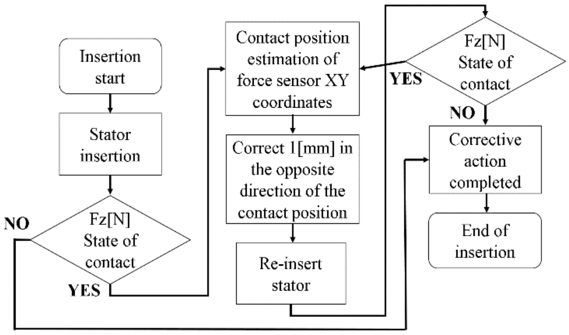

Insertion start
Stator insertion
Fz[N] State of contact
NO
YES
Contact position estimation of force sensor XY coordinates
Correct 1[mm] in the opposite direction of the contact position
Re-insert stator
Fz[N] State of contact
YES
NO
Corrective action completed
End of insertion

Fig. 6 Flowchart of the insertion process

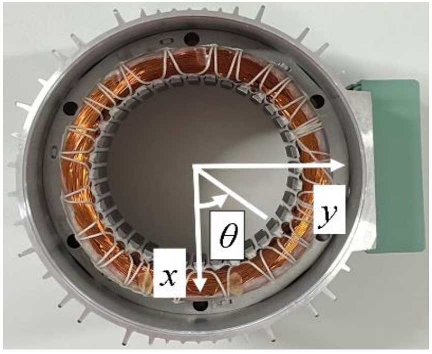


Fig. 7 Image of contact position by angle

Table 1 Average time in case of 45°

Test time	171932.8 msec
Adjustment operation time	69422.8 msec

5. Conclusions

In this paper, the robotic assembly of the gearmotor was discussed. The process that the stator was recognized, grasped and inserted into the reducer was performed using two main methods (keypoint matching and contact position estimation). The validity of these method was verified by experimentation.

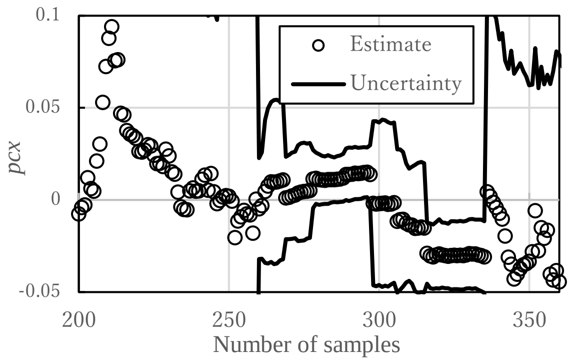


Fig. 8 Estimates of p_{cx} and its uncertainty

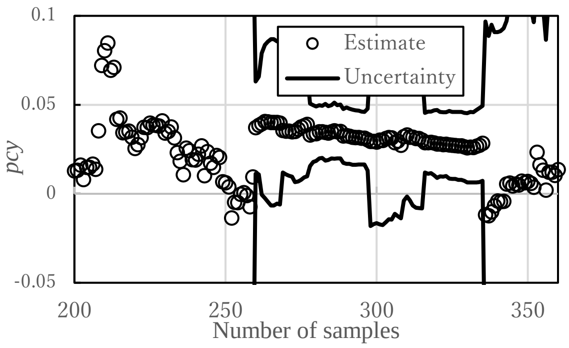


Fig. 9 Estimates of p_{cy} and its uncertainty

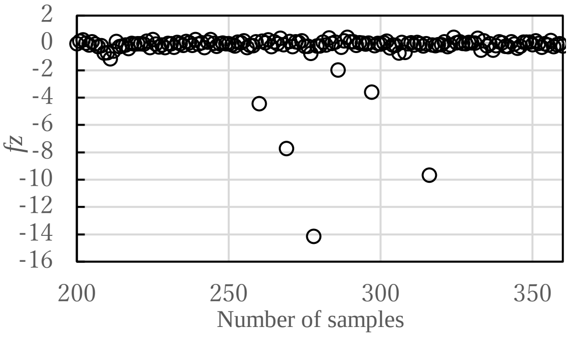


Fig. 10 Measurement force f_z

References

1. W. Tanaka, K. Joukouchi, Y. Nawa, E. Shinbo, K. Nonami, Study on Robotarm Control (Verification of Automatizing Work Using Image Recognition and Multiple Joint Robot Arm), Chiba Prefectural Institute of Industrial Support Technology, ed. (7), pp. 3-6, 2009. (in Japanese)
2. M. Satoh, Y. Yamamoto, T. Yoneyama, T. Hashimoto, Measure and Analysis of Physical Quantities in Precision Assembly Tasks, Proceedings of the Academic Lecture at the Society of Precision Engineering Conference, Vol. 2002, pp. 456, 2002. (in Japanese)
3. P. F. Alcantarilla, J. Nuevo, A. Bartoli, "Fast Explicit Diffusion for Accelerated Features in Nonlinear Scale Spaces", Proceedings of British Machine Vision Conference, pp. 1-11, 2013.
4. T. Yamada, T. Mouri, N. Mimura, H. Yamamoto, Identification of Contact Conditions by using 6-Axis Force Sensing and Robot Intelligence, Science of Machine, Vol. 69, No. 9, pp. 752-761, 2017. (in Japanese)

Authors Introduction

Mr. Yasumoto Imai

He received the B.S. degree in Mechanical Engineering from Gifu University, Japan, in 2021. He is a student at the Graduate School of Natural Science and Technology, Gifu University.

Dr. Takayoshi Yamada

He received the B.E., M.E, Ph.D. degrees in Mechanical Engineering from the Nagoya Institute of Technology, Japan, in 1991, 1993, 1995, respectively. He is a Professor with the Dept. of Mech. Eng., Gifu University. His research interests include grasping, manipulation, sensing, and automation systems.

Dr. Junya Sato

He received the B.E., M.E., and Ph.D. degrees in Eng. from Iwate University in Japan, in 2012, 2014, and 2017, respectively. He is an Assistant Professor with the Dept. of Mech. Eng., Faculty of Eng., Gifu Univ., Japan. His research interests include applications of evolutionary computations to computer vision.

Mr. Toshiki Hayashi

He received the B.E. degree in Mechanical Engineering from Kanazawa Institute of Technology, Japan, in 2013. He is a member of the staff of Tsubakimoto Chain Co. in Japan.

Mr. Shota Aono

He received the B.E. degree in Electronics from Doshisha University, Japan, in 2018. He is a member of the staff of Tsubakimoto Chain Co. in Japan.

Robot Arm Operating Interface for Easy Grasping by Specifying the Gripping Width of End-effector

Rio Takeuchi, Laijun Yang, Norihiko Kato
Department of Mechanical Engineering, Graduate School of Engineering, Mie University
1577 Kurimamachiya-cho, Tsu-city, Japan
E-mail: 421M131@m.mie-u.ac.jp, 418DB52@m.mie-u.ac.jp, nori@mach mie-u.ac.jp

Abstract

In Japan, cervical cord injuries affect roughly half of all patients with spinal cord injuries. High-Level spinal cord injuries make it particularly difficult for patients to pick up objects from the ground because of the functional impairment of the trunk. To help these people become more independent, welfare robot arms have been developed recently. In our lab, an interface has been proposed to control the robotic arm by drawing a line on a touch screen in order to grasp an object on the floor and deliver it to the user. This technique has a fixed closing width for grasping, making it challenging to use when the object's sizes change. Therefore, we proposed a method that allows the adjustment of the closing width according to the drawn line on the touch panel interface. As a result, we found that the grasping was still possible even the object was altered.

Keywords: Vision-based interface, Remote control, Eye in hand system, Welfare robot arm.

1. Introduction

Nowadays, there are many disabled people in Japan whose daily lives are limited by physical disabilities; in 2018, there were 4.36 million disabled people in Japan [1]. Which increasing year by year [2].

The spinal cord is primarily hurt when the spinal column is subjected to a powerful external force. Most of these injuries are caused by car accidents or falls from height, and many patients are injured at the C5 level [3]. Because the paralysis of C5 patients extends over almost their entire body from the chest down, it is difficult for them to pick up objects from the floor, grasp, and so on. In recent years, robotic arms mounted on wheelchairs have been introduced in nursing care situations to assist physically disabled people in their daily lives [4] [5].

For example, there are automatic operation interfaces using AR markers and interfaces that automatically grasp objects by recognizing them through the installation of multiple cameras [6] [7]. However, there is a possibility of maloperation or misidentification, and the use of multiple cameras limits the situations in which they can be used depending on the operation method. Our laboratory also proposed an operation interface in which a single 2D camera is mounted on the end-effector of a robot arm to perform grasping operations on objects in a work environment. However, the object is limited to the size because the closing width of the end-effector is constant during grasping.

In this study, we propose a new manipulation interface to extend our previous one and conduct an experiment with a healthy person.

2. Visual-based operating interface

In this study, a touch panel such as a tablet device is used as the input device. Since there is no need to grip the interface, even C5 level cervical cord injured patients can operate the device. In addition, they are accustomed to operating the interface because they use smartphones and other devices on a daily basis.

As it is relevant to the proposed method, we will briefly describe the previous methods. As a simple method for specifying a line indicating the target position is drawn on the touch panel. Consider the workspace with the origin on the floor shown in Fig. 1. A line is drawn on the operation screen to specify the hand target. The line and touched point on the screen are transformed into those on the floor. In Eq. (1) $\boldsymbol{x_c}$ is the coordinates of the camera attached to the end-effector of the robot arm in the workspace. $\boldsymbol{v}$ is the vector from the operation screen to the specified point and $\boldsymbol{x_i}$ is the point on the floor.

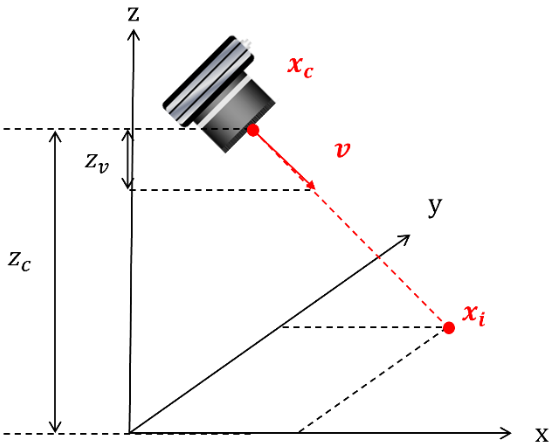


Fig.1 Working-space coordinates

$$\boldsymbol{x_i} = \frac{z_c}{z_v}\boldsymbol{v} + \boldsymbol{x_c} \qquad (1)$$

In the previous method of our laboratory, the grasping closure width is constant. Therefore, if the size or characteristics of the target object change, the robot arm may not be able to grasp it.

This method includes a function that adjusts the end-grasping effector's width based on the length of the drawn line.

3. Experiment on healthy subjects

3.1. *Experimental apparatus*

The experimental setup is shown in Fig.2. The robot arm used in the experiment was myCobot280M5 manufactured by Elephant Robotics Co. The electric wheelchair was an EMC-250T manufactured by Imasen Electric Co.

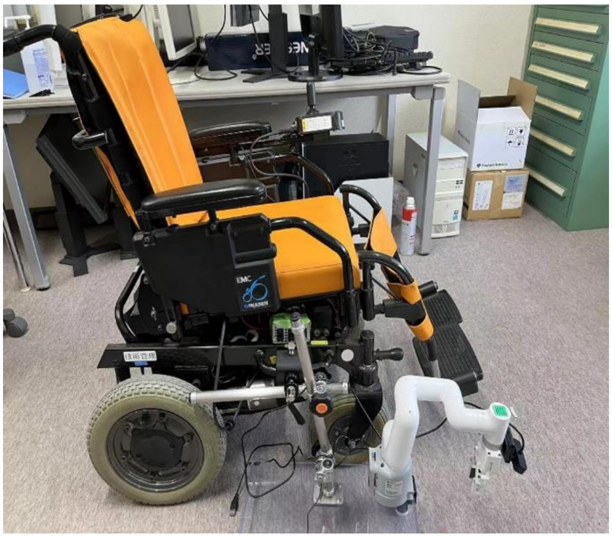
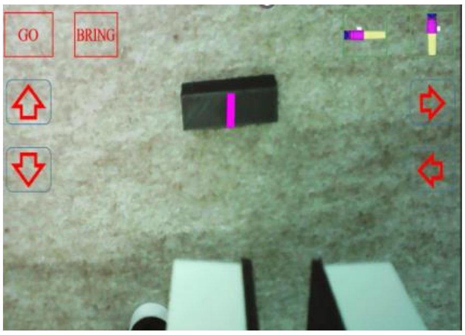


Fig.2 Experimental equipment.

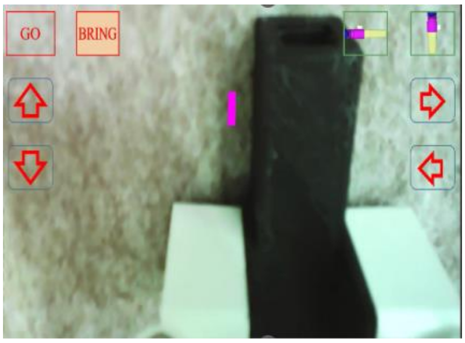
3.2. *Experimental outline*

In this study, the experimental subject was a healthy person. In the experiment, two types of objects of different sizes were placed. The experiment's goal was to manipulate the end-effector by the proposed interface to grasp objects of different widths. The reason for conducting this test on healthy subjects is to confirm that grasping is possible when the size of the target object differs from that of the subject. The primary goal of this experiment is to demonstrate that grasping is indeed possible even when the targets are of different sizes. We also want to make sure that the calculated value in the program matches the actual value of the grasping closure width. As a result, an experiment was carried out in this paper with a healthy subject.

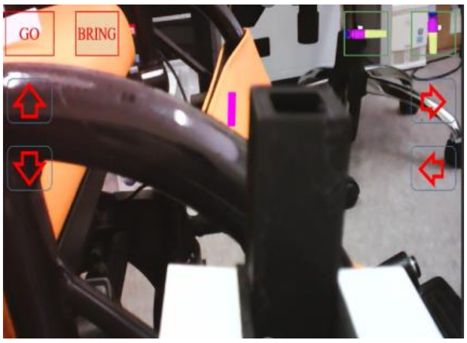
Because C5 level injured subjects have contractures in their fingers, subject was instructed to operate the touch panel with the second joint of their pinky finger for reproduction purposes. Since the maximum opening width of the end-effector of the robot arm used in the experiment is limited to 40 mm, the grasping objects in the horizontal plane are two objects with a length × width of 50 mm × 22 mm and 50 mm × 15 mm.



(a) By drawing a line along the object, the position and grasping posture are created.



(b) By pressing the GO button, the device is brought within 60mm of the object, and by pressing the BRING button, it is grabbed.



(c) Deliver the grasped object to the wheelchair passenger

Fig.3 View of grasping a 50 mm × 15 mm object

3.3. *Experimental results*

The grasping scenery of each object in the experiment is shown in Fig.3 and Fig.4 The operator specifies the target position and closing width by drawing a line as shown in Fig.3 and Fig.4 at the point where he/she wants to perform grasping.

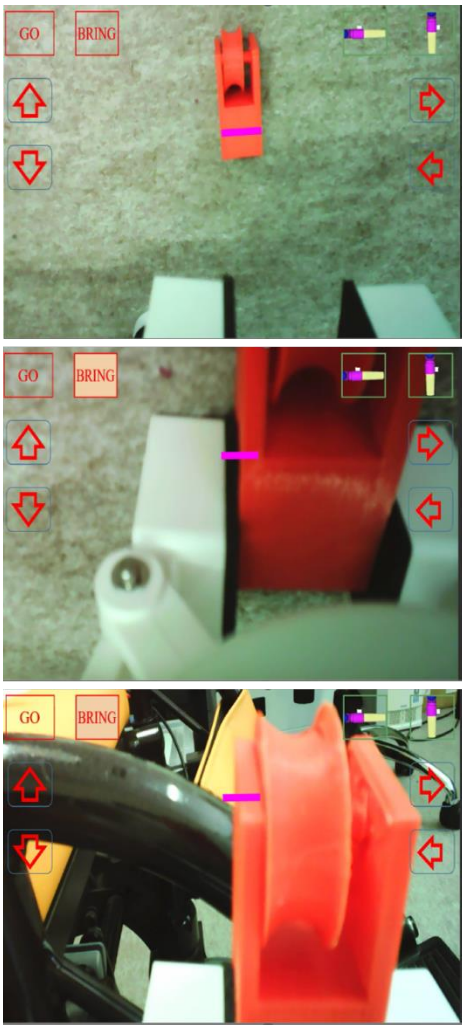


Fig.4 View of grasping a 50 mm × 22 mm object

The command values that specify the closing width of the hand calculated by the line drawn for each object are shown in Table.1.

Since the experiment was conducted twice for each object, the average value was calculated. D is the length of the drawn line, and v is the command value. For the 50 mm × 15 mm object, the program outputs a line length of 15.3 mm, which is close to the object; for the 50 mm × 22 mm object, the program outputs a line length of 22.6 mm, which is also close. Therefore, it was confirmed that grasping of the object is possible, and furthermore, the program outputs a value that is close to the actual value of the closing width for grasping calculated in the program.

Table.1 Experimental results for each object

Object	Command value v[-]	Length of line drawn D[mm]
50 mm × 15 mm	29	15.3
50 mm × 22 mm	43	22.6

4. Conclusion

Since many spinal cord injury patients are forced to live in wheelchairs or in bed, a robotic arm is needed to perform grasping of objects in daily life instead. We proposed an interface to specify the closing width of the robot arm's end-effector by drawing a line on the touch panel to a target object on the floor using a single camera. specifying the target position, grasping posture, and closing width of the robotic arm. We verified a method for grasping objects of different sizes by changing the closing width of the arm's grasp according to the length of the line drawn on the touch panel. On healthy subject, we performed experiments using targets of various sizes, and we confirmed that both the grasping width and the targets could be specified.

It has been demonstrated that even when the size of the target object varies, it is still possible to specify the closing width for grasping and to grasp an object. The closing width command to the robot arm must be adjustable in an environment with multiple objects and varied target hardness. Among other factors, it must be decided whether to specify the closing width as the robot arm approaches the object or just before it is grasped. As a result, taking into account the environment with its various objects, materials, etc. is also necessary in order to grasp an object more appropriately.

References

1. Cabinet Office, "annual Report on Government Measures for Persons with Disabilities", pp213, 2021.(in Japanese).
2. Cabinet Office,"Annual Report on Government Measures for Persons with Disabilities", pp.1, 2012. (in Japanese).
3. National Spinal Cord Injury Statistical Center, "Traumatic Spinal Cord Injury Facts and Figures at a Glance", 2022, (in Japanese).
4. K. Nagata., Y. Wakita, N. Yamanobe, N. Ando, "Development of RT-Component for Assistive Robotic Arm and Man-Machine Interface Using OpenRTM-aist", Transactions of the Japan Society of Mechanical Engineers Series C, Volume 76, Issue 766, pp. 1419- 1426, 2010. (in Japanese).
5. V. Maheu, P. Archambault, J. Frappier, F. Routhier, "Evaluation of the JACO robotic arm: Clinico-economic study for powered wheelchair users with upper-extremity disabilities", IEEE International Conference on Rehabilitation Robotics, 2011.
6. Siddhartha S. Srinivasa, Dave Ferguson, Casey J. Helfrich, Dmitry Berenson, Alvaro Collet, Rosen Diankov, Garratt Gallagher, Geoffrey Hollinger, James Kuffner, Michael Vande Weghe, "HERB : a home exploring robotic butler", Auton Robot, 2010.
7. J. Kofman, Xianghai Wu, T.J. Luu, S.Verma, "Teleoperation of a Robot Manipulator Using a Vision-Based Human-Robot Interface", IEEE Transactions on Industrial Electronics ,2005.

Authors Introduction

Mr. Rio Takeuchi

He received his Bachelor's degree in Mechanical Engineering, Faculty of Engineering, Mie University, Japan in 2021. He is currently a Master's course student in Mie University, Japan.

Mr. Laijun Yang

He received the B.E. degree in mechanical and control engineering from Xi'an University of technology, Shan'xi, China, in 2014, the M.E. degree in Dept. of mechanical engineering from the Mie University, Mie, Japan, in 2018, where he is currently pursuing the Ph.D. degree in system engineering in Mie University. His current research is focus on redundant manipulators.

Dr. Norihiko Kato

He received his Dr. Eng. from Nagoya University, JAPAN. He is currently an Associate Professor in Mie University. His research interest is robotics.

3D Real-Time Conversational Agents: Do Facial Expressions and Camera Angles Persuade Human?

Pinkie Anggia
Graduate School of Systems Information Science, Department of Media Architecture, Future University Hakodate
Hakodate, Hokkaido 041-8655, Japan

Kaoru Sumi
Department of Media Architecture, Future University Hakodate
Hakodate, Hokkaido 041-8655, Japan
Email: pinkie.anggia@outlook.jp, kaorus@fun.ac.jp
www.fun.ac.jp

Abstract

This paper aims to uncover state-of-the-art on persuasion through a dialogue system between human and virtual agents. 37 participants interacted with six virtual agents in different conditions of facial expressions and camera angles. Through empirical measurements using the Big Five theory, we discover that virtual agents' friendliness becomes a solid persuading factor to our predominantly extroverted participants. The experiment deduces that openness, conscientiousness, and extraversion personalities are easily persuaded by our virtual agents, unlike reversed-conscientiousness and neuroticism personalities.

Keywords: Persuasive Technology, Human-Agent Interaction, Digital Storytelling, Virtual Agents, Big Five

1. Introduction

Computing technologies can be utilized to help people in performing their daily tasks in different situations and setups. Persuading and motivating people through the usage of computing technology leads to the so called “persuasive technology.” Persuasive technology effectiveness can be measured by how successful it is in changing peoples’ opinion and attitude and leading the peoples’ mindset into a well-defined state. Persuasive technology is about incorporating computing technology with persuasion and social influence, not coercion nor [1].

Persuasion can be accomplished by influencing the emotions of trust, which has been an attractive topic for the research community. Thus, persuading and altering decisions relying on different emotions of trust have a variety of impacts. Hence, trust is a key factor in persuasion [2], [3]. It's important to differentiate between persuasion and coercion as the former has a positive meaning and impact, unlike the latter. In our work, a main guideline that we follow is to encourage and motivate positively from our agent to the user.

Developing an agent depends highly on the goal in mind that the agent is supposed to accomplish. Goals of agents vary massively depending on the environment, setup, culture, etc. An intelligent friendly agent iParrot goal was to influence family members to conserve energy at their homes [4]. An Embodied Conversational Agent (ECA) that promotes physical activity among Chinese adults that went through cultural adaptation for the Chinese community [5]. Agent embodiment can change the behavior of subjects into conserving more energy, thus feedback from an embodied agent is better than a computer [6].

Designing an effective agent can be achieved by following the best practices from former research. An ECA or an agent with a social dialogue is shown to be more effective than a text-based one [7]. Moreover, the attractiveness of the agent was proven to be an important factor on how persuasive the agent can be and the amount of behavioral change was very noticeable [8]. The number of agents is shown to play an important role in being more persuasive to the subject as in [9], where a multiple agent setting was shown to be more persuasive than a single agent setting. Moreover, the use of animation utilizing its expressive nature can enhance the intuitive understanding as it involves direct movement was also investigated previously [25]. Evaluating the user's impression of the agent was also studied comprehensively across 96 patterns of content ranging from emotional scenarios, facial expressions, and words used by the agent [26].

The Big Five personality traits are a grouping for personality traits that identifies five factors typically referred to as openness, conscientiousness, extraversion, agreeableness, and neuroticism [10]. Incorporating characters design and portraying personalities with Big Five Personality traits was investigated [11]. Moreover, personalizing the persuasion techniques based on the Big Five personality traits has been shown to be very effective [12].

In our work, we design virtual agents by using Unreal Engine [13] while considering the Big Five personality traits as a method to obtain interesting insights on how to persuade subjects more effectively. Moreover, we change the camera angle and the personality of the agents to study how such changes will affect the subjects. We perform our quantitative numerical analysis through two-way ANOVA test without replication [14]. Our numerical results show a strong relation between persuasiveness and the agents' personality. It also shows how related the personality of the subjects to how easy it is for them to be persuaded.

The rest of the paper is organized as follows. In Section II, we provide a solid background about Big Five personality traits. Section III deals with the methodology that we followed in our experiment. Section IV shows our numerical results and insights. Finally, we conclude the paper in Section V.

2. Related Research

The golden standard for exploring personality in the academic world is known as the Big Five, which classifies personality across five dimensions: Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism (OCEAN). Big Five personality traits affect how people respond to persuasive appeals, and how core personality traits shape responses to various aspects [24]. Hence, we choose the Big Five theory to understand whether the personality of our subjects make them easily persuaded by our agents.

2.1. *Openness*

A past study on impulse buying behavior examined that openness personality affects impulse buying [15]. Openness to new experience individuals will probably trust more to online retailers because they will enjoy purchasing from online sites which provide variety of products [16]. In marketing, for high scorers in openness, aesthetic and intellectual pursuits are particularly motivating, and people with this trait will respond well to words such as innovation, intelligence, sophistication, imagination, and creativity.

2.2. *Conscientiousness*

Conscientious individuals are probably to buy online, as online shopping is convenient and happens in a comfortable environment [17]. The study suggests that conscientious individuals are probably more careful in considering the benefits of online purchase intention, which can build their trust in online retailers [4]. In marketing, conscientious individuals will be more driven by a desire for order, efficiency, and achievement.

2.3. *Extraversion*

Highly extroverted individuals are expected to engage more in online purchasing activities because online customers can share their shopping experience and information on the online platform [18]. Past studies also found that there is a correlation between trust and extroversion [19]. In marketing, people who score highly in extraversion tend to be triggered by a desire for excitement and social rewards and will find words such as strong, outgoing, active, excitement, and attention more persuasive.

2.4. *Agreeableness*

A past research indicated that individuals with a high degree of agreeableness would have a high level of technology acceptance [20]. People who have agreeableness trait pay extra attention to preventing conflict. Therefore, they are eager to trust service providers [21]. In marketing, agreeable individuals tend to be motivated more strongly by compassion, interpersonal harmony, and a sense of belonging.

2.5. *Neuroticism*

The previous study revealed that neuroticism personality individuals are associated positively with computer anxiety [22]. Neurotic individuals invariably conceive themselves in inauspicious attitudes [10]. In marketing, if our target audience is particularly neurotic, they will be more sensitive to threats and uncertainty and will respond to messages that promise safety and security.

3. Experiment of an Interaction Between a Subject and the Virtual Agents

There were 37 students participated in our research whose age range is from 18 to 24 years old. These 37 subjects are added randomly into three groups. There were 11 subjects in Group 1 [Up], 11 subjects in Group 2 [W], and 16 subjects in Group 3 [none]. The system was created using the Unreal Engine, displaying the agents' movements and facial expressions, and the voice was presented using Japanese speech synthesis. To analyze the effect of persuasion, we use two-way ANOVA without replication method. To generate the Big Five personalities, we measure the means of personalities and perform one-way ANOVA.

3.1. *Experimental method*

3.1.1. *Virtual agents*

The 3D model of virtual agents in Fig. 1 was created using the Unreal Engine MetaHuman 4.27. They show two facial expressions, which are smile and grumpy. We use two different camera angles both from near and far.

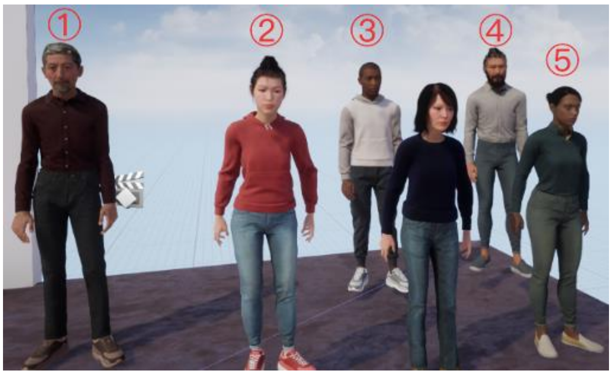


Fig. 1. The six virtual agents, named as (1) Stephan, (2) Pia, (3) Kellan, (middle) Danielle, (4) Aoi, and (5) Kendra.

3.1.2. *Grouping the conditions*

We categorize three conditions, which are shown on Fig. 2 of Group 1 [Up], Fig. 3 of Group 2 [W], and Fig. 4 of Group 3 [none]. Each subject only interacted to one group. It means that there is no subject who interacted twice in more than one groups.

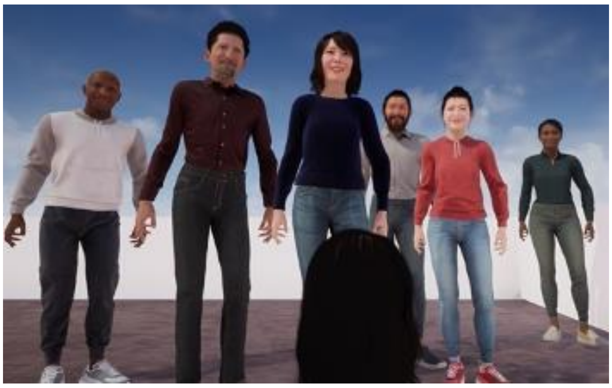


Fig. 2. Group 1 [Up]: smiling agents with the camera angle from below.

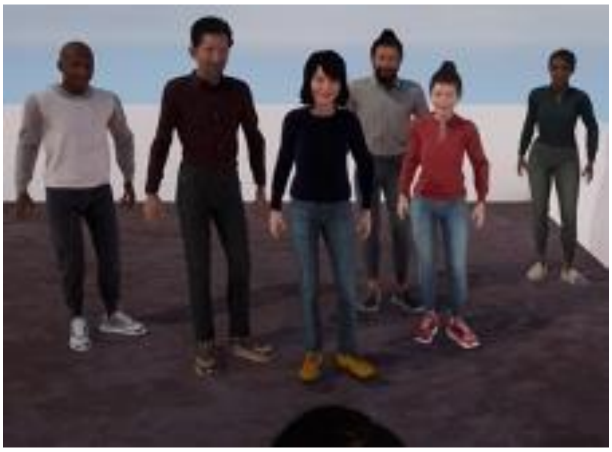


Fig. 3. Group 2 [W]: smiling agents with the normal camera angle.

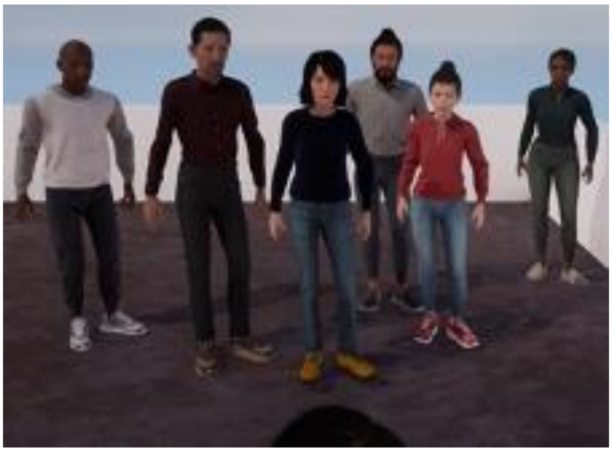


Fig. 4. Group 3 [none]: non-smiling agents with the normal camera angle.

3.1.3. *Procedures*

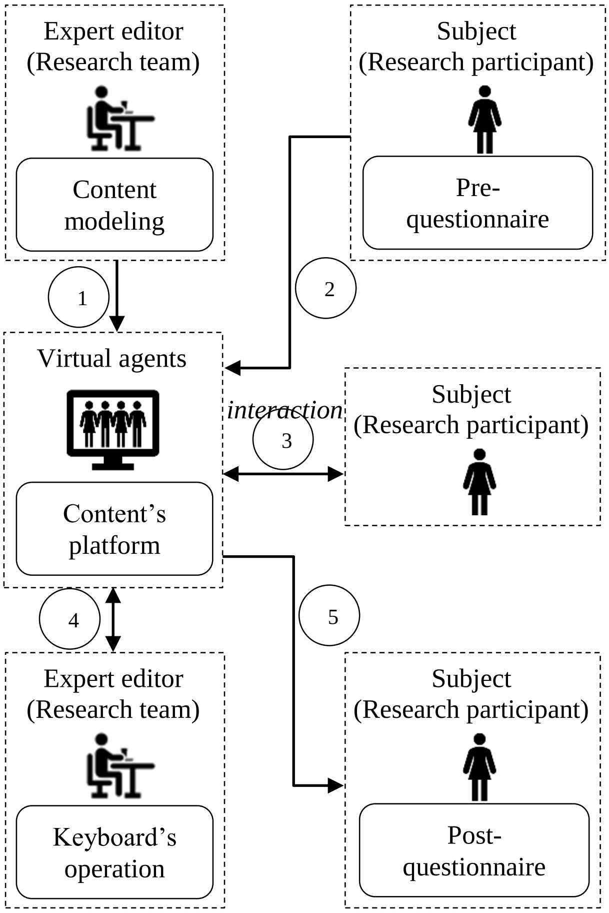


Fig. 5. Diagram of experiment.

The sequence of the experiment's procedures in Fig. 5. above is elaborated in detail as follows:

1. We developed a system that displays the virtual agents using Unreal Engine 4.27, with speech and facial expressions assigned to the keyboard. The system consists of a keyboard key with voice, camera angles, facial expressions, and actions. The experimenter operated the keyboard while a subject interacted with virtual agents in the system.
2. Prior to the interaction with virtual agents, each subject informed about a consent form, and signed it. Then, he/she fills a pre-questionnaire, which aims to understand the knowledge about a topic persuaded by the agents, that Royal Jelly is either good or bad for health. We choose this topic because there are advantages and disadvantages of nutritional foods, and we use that as a persuasive factor.
3. A subject was instructed to put on headphones, sit in front of the computer display, and interact with virtual agents. The virtual agent, Danielle, speaks to a subject. All virtual agents' voice played back from a synthesized voice recording. The detailed dialogue by virtual agents is mentioned in Appendix A. Table of Conversation. The virtual agents of (1) Stephan, (2) Pia, (3) Kellan, (4) Aoi, and (5) Kendra speak one by one after Danielle's voice.
4. The experimenter operates the system with a keyboard and asks a subject to respond interactively to the virtual agents by speaking to the system using a headphone set with a microphone. Although this system is not an actual interactive dialogue system, subjects were instructed by the experimenter to interact with the virtual agent.
5. After the interaction, each subject fills the post-questionnaire, which aims to investigate the level of persuasiveness from the contents presented by the agents. At the end of interaction, all of the subjects asked why did they give the certain answers.

3.2. *Experiment's evaluation of subjects who change their opinion during experiment: (yes → no) and (no → yes) variables*

We aim to reach two main objectives, which are finding the effect of persuasion on subjects who change their minds (*yes* → *no*) and (*no* → *yes*) from three different camera angles and facial expressions of the agents, as well as generating the Big Five personalities that the subjects and agents have in common.

Table 1. Total of subjects who change their opinion

	Group 1 [Up]	Group 2 [W]	Group 3 [none]
(yes → no)	1 person	5 people	6 people
(no → yes)	0	1 person	2 people

According to Table 1, we have two independent variables, which are (*yes* → *no*) (i.e., when the subjects change their mind from yes to no), and (*no* → *yes*) (i.e., when the subject change their mind from no to yes). We also have three conditions of camera angles and agents' facial expressions, which are grouped by into three categories, namely Group 1 [Up], Group 2 [W], and Group 3 [none]. The proper method to analyze their significance is two-way ANOVA without replications for the following reasons:

- To make comparisons between the means of our three groups of data, where two independent variables are considered.

- To allow studying the effect of multiple groups of two independent variables on a dependent variable and on each other.
- To measure each independent variable that has multiple samples.
- To find significance results from unequal sample sizes (*n*) in the subgroups.

Hence, this method has two null hypotheses tested:

- H1: The means of observations grouped by one factor are the same.
- H2: The means of observations grouped by the other factor are the same.

3.2.1. *Hypothesis testing results from subjects who change their opinion during experiment: (yes → no) and (no → yes) variables*

Table 2. ANOVA: two-factor without replication

Source of Variation	*SS*	*dF*	*MS*	*F-val*	*P-val*	*F crit*
Rows	4.46	36.00	0.12	0.70	0.86	1.74
Columns	1.09	1.00	1.09	6.15	0.02	4.11
Error	6.41	36.00	0.18			
Total	11.96	73.00				

As a result of null hypothesis H1 on Table 2, *F-value* (0.70) is less than the *F-crit* (1.74), and *P-value* (0.86) is more than the *Alpha-level* (0.05). Thus, null hypothesis H1 was rejected, and the means of observations grouped by one factor are the same. In conclusion, there are no significant differences in personalities between each subject participating in our research.

As a result of null hypothesis H2, *F-value* (6.15) is more than the *F-crit* (4.11), and *P-value* (0.02) is less than the *Alpha-level* (0.05). Thus, we successfully reject the null hypothesis H2, and the means of observations grouped by one factor are not the same. In conclusion, there is a significant difference in participants who change their minds during the experiment (*yes* → *no*) and (*no* → *yes*), which are affected by, specifically, two camera angles and agents' expressions of conditions from the Group 1 [Up] and Group 3 [none].

To generate the Big Five personalities from subjects who change their opinion during experiment, we perform these steps:

(i) Measure the independent variables (*yes* → *no*) and (*no* → *yes*) separately. Therefore, there will be two results from two variables.
(ii) Measure the means of personalities.
(iii) Perform one-way ANOVA.

3.2.2. *Big five personalities from subjects who change their opinion during experiment: (yes → no) variable*

In this stage, there are a total of 12 subjects who change their opinion (*yes* → *no*) during the experiment. There are 1 person in Group 1 [Up], 5 people in Group 2 [W], and 6 people in Group 3 [none]. After generating their response with 5-likert scale, we measure the means from each group, and find the Big Five personalities from the highest value of means, which are:

- Group 1 [Up]: openness, reversed-conscientiousness
- Group 2 [W]: reversed-conscientiousness
- Group 3 [none]: openness

Then, we perform one-way ANOVA to find whether their personality values are significantly different. This method tests one hypothesis, which is H1: The means of observations grouped by one factor are the same.

Table 3. ANOVA: single factor of (*yes* → *no*) variable

Source of Variation	*SS*	*dF*	*MS*	*F-val*	*P-val*	*F crit*
Between groups	48.03	9.00	5.34	4.98	0.00	1.97
Within groups	117.83	110.00	1.07			
Total	165.87	119.00				

Table 3 shows that *F-value* > *F crit* and *P-value* < *Alpha-level* (0.05), which means that it is successfully rejected the null hypothesis H1. Hence, there is a significant difference between each subject's personality values.

3.2.3. *Big five personalities from subjects who change their opinion during experiment: (no → yes) variable*

In this stage, there are a total of 3 subjects who change their opinion (*no* → *yes*) during the experiment. There are no subject in Group 1 [Up], 1 person in Group 2 [W], and 2 people in Group 3 [none]. After generating their response with 5-likert scale, we measure the means from each group, and find the Big Five personalities from the highest value of means, which are:

- Group 1 [Up]: no subject
- Group 2 [W]: extraversion, conscientiousness, openness
- Group 3 [none]: neuroticism

Then, we perform one-way ANOVA to find whether their personality values are significantly different. This method tests one hypothesis, which is H1: The means of observations grouped by one factor are the same.

Table 4. ANOVA: single factor of *(no → yes) variable*

Source of Variation	*SS*	*dF*	*MS*	*F-val*	*P-val*	*F crit*
Between groups	18.53	9.00	2.06	1.17	0.37	2.39
Within groups	35.33	20.00	1.77			
Total	53.87	29.00				

Table 4 shows that *F-value* < *F crit* and *P-value* > *Alpha-level* (0.05), which means that it is failed to reject the null hypothesis H1. Hence, there is no significant difference between each subject's personality values.

3.3. *Experiment's evaluation of subjects who do not change their opinion during experiment: (yes → yes) and (no → no) variables*

Table 5. Total of subjects who do not change their opinion

	Group 1 [Up]	Group 2 [W]	Group 3 [none]
(yes → yes)	9 people	5 people	4 people
(no → no)	1 person	0	3 people

Table 5 shows the total subjects who change their opinion during experiment from two variables (*yes → yes*) and (*no → no*). The method selected for testing the significance is two-way ANOVA without replication method. It tests two null hypotheses, which are:

- H1: The means of observations grouped by one factor are the same.
- H2: The means of observations grouped by the other factor are the same.

3.3.1. *Hypothesis testing results from subjects who change their opinion during experiment: (yes → yes) and (no → no) variables*

Table 6. ANOVA: two-factor without replication

Source of Variation	*SS*	*dF*	*MS*	*F-val*	*P-val*	*F crit*
Rows	4.46	36.00	0.12	0.53	0.97	1.74
Columns	2.65	1.00	2.65	11.42	0.00	4.11
Error	8.35	36.00	0.23			
Total	15.46	73.00				

As a result of null hypothesis H1 on Table 6, *F-value* (0.53) is less than the *F-crit* (1.74), and *P-value* (0.97) is more than the *Alpha-level* (0.05). Thus, null hypothesis H1 was rejected, and the means of observations grouped by one factor are the same. In conclusion, there are no significant differences in personalities between each subject participating in our research.

As a result of null hypothesis H2, *F-value* (11.42) is more than the *F-crit* (4.11), and *P-value* (0.00) is less than the *Alpha-level* (0.05). Thus, we successfully reject the null hypothesis H2, and the means of observations grouped by one factor are not the same. In conclusion, there is a significant difference in participants who change their minds during the experiment (*yes → yes*) and (*no → no*), which are affected by, specifically, two camera angles and agents' expressions of conditions from the Group 1 [Up] and Group 2 [W].

To generate the Big Five personalities from subjects who do not change their opinion during experiment, we perform the same steps as mentioned in section 3.2.1.

3.3.2. *Big five personalities from subjects who change their opinion during experiment: (yes → yes) variable*

In this stage, there are a total of 18 subjects who do not change their opinion (*yes → yes*) during the experiment. There are 9 people in Group 1 [Up], 5 people in Group 2 [W], and 4 people in Group 3 [none]. After generating their response with 5-likert scale, we measure the means from each group, and find the Big Five personalities from the highest value of means, which are:

- Group 1 [Up]: reversed-conscientiousness
- Group 2 [W]: neuroticism
- Group 3 [none]: openness, reversed-conscientiousness

Table 7. ANOVA: single factor of (*yes → yes*) variable

Source of Variation	*SS*	*dF*	*MS*	*F-val*	*P-val*	*F crit*
Between groups	32.80	9.00	3.64	3.16	0.00	1.94
Within groups	196.00	170.00	1.15			
Total	228.80	179.00				

Table 7 shows that *F-value* > *F crit* and *P-value* < *Alpha-level* (0.05), which means that it is successfully rejected the null hypothesis H1. Hence, there is a significant difference between each subject's personality values.

3.3.3. *Big five personalities from subjects who change their opinion during experiment: (no → no) variable*

In this stage, there are a total of 4 subjects who do not change their opinion (*no → no*) during the experiment. There are 1 person in Group 1 [Up], no subject in Group 2 [W], and 3 people in Group 3 [none]. After generating their response with 5-likert scale, we measure the means from each group, and find the Big Five personalities from the highest value of means, which are:

- Group 1 [Up]: neuroticism, reversed-conscientiousness, reversed-openness
- Group 2 [W]: none
- Group 3 [none]: openness

Table 8. ANOVA: single factor of (*no* → *no*) variable

Source of Variation	*SS*	*dF*	*MS*	*F-val*	*P-val*	*F crit*
Between groups	18.03	9.00	2.00	1.34	0.26	2.21
Within groups	44.75	30.00	1.49			
Total	62.78	39.00				

Table 8 shows that *F-value* < *F crit* and *P-value* > *Alpha-level* (0.05), which means that it is failed to reject the null hypothesis H1. Hence, there is no significant difference between each subject's personality values.

4. Discussion

We conducted the experiment on how the interaction from the virtual agents are able persuade a subject. Three conditions of camera angles and virtual agents' facial expressions are selected. Three methods are applied to measure the effect of this persuasion and to generate the Big Five personalities.

We found out that there is a significant difference between conditions of Group 1 [Up] and Group 3 [none], which affects the opinion of people. Thus, influencing them to perform certain actions. Group 1 [Up] is a state where the camera angle is from nearby, and the agents look up with smiling and laughing facial expressions. Group 3 [none] is a state where the camera angle is from afar while the agents are looking straight at the camera and neither smiling nor laughing.

Our experiment shows that Group 1 [Up] makes a subject to be persuaded easily by the virtual agents than the other conditions. The reasons are as follows:

(i) The virtual agents' personalities.
After the experiment, we asked all subjects of what made them change their mind. Most of them said that, "the characters seem kind." Hence, how nice and friendly the virtual agents appear has a positive effect on a subject to be persuaded. Thus, correlating with results in previous study where virtual agents with good impressions were effective in persuasion [26].

(ii) The matching personalities between a subject and the virtual agents, especially extraversion.
In our experiment, most of the subjects answer their personality with: "I think I am sloppy and careless" (72.9%), "I think I am a kind person who cares about others" (48.6%), and "I think I am lively and extroverted" (43.24%). These answers define an extraversion personality. People with high extroversion cares highly about social interaction, but careless in other aspects [23]. Extroverted people care about others and what they think. Thus, the matching personality between the virtual agents and a subject, specifically extraversion, made a subject persuaded easily.

(iii) The subject's personalities.
The subjects who change their opinion from no to yes mostly have personalities of openness, conscientiousness, and extraversion. These personalities have similarities that pronounced engagement with the external world. After the experiment, we asked what made them change their mind during the experiment. They answered that, "I wanted to feel 'safe' and did not want to make any waves in the conversation." Since they have a tendency to think of their social consequences of behavior, it makes them easily influenced or persuaded by the agents.

The participants who change their opinion from yes to no (*yes* → *no*) have personalities of reversed-conscientiousness and neuroticism. Both of these personalities are often perceived as being stubborn, focused, strive for achievement against measures or outside expectations, and highly critical. Therefore, they did not get easily influenced or persuaded by the agents.

To design virtual agents that will persuade people in the future, we need to pay attention to the quality of animation and speech, their appearance, the naturalness aspect of their way of storytelling, as well as the matching storytelling and other aspects of the participants' culture.

Despite the precise outcomes, which could be very beneficial for further research, this study has potential limitations. First, it needs to increase the sample size, which involves more participants in different age ranges. Second, to achieve balanced data, the sample size in each subgroup should be uniform.

When developing dialogue agents' system, we need to precisely target the participants and adjust the content that has a middle ground between the participant's culture and personality. It is essential because the results reveal that matching the personality between a subject and the virtual agents becomes a persuading factor.

5. Conclusion

This research incorporates the Big Five theory in a model where camera angles and friendliness of virtual agents are being studied in relation to persuading subjects. We investigate the relation between a subject's personality and how it can affect them being persuaded by the virtual agents. The camera angles and how friendly the agents seem to be had a positive effect on a subject in being persuaded. Through this observation, we conclude that subjects with the personalities of reversed-conscientiousness and neuroticism are hard to be persuaded, unlike the subjects with the personalities of openness, conscientiousness, extraversion, who can be easily persuaded.

References

1. B. J. Fogg, "Persuasive technology: using computers to change what we think and do," Ubiquity, vol. 2002, no. December, p. 2, 2002.
2. W. W. Ahmad, "Development of emotion-based trust model for designing persuasive application," in 9th International Conference on Persuasive Technology, PERSUASIVE, vol. 2014, 2014, pp. 21–23.
3. W. N. W. Ahmad and N. M. Ali, "A study on persuasive technologies: the relationship between user emotions, trust and persuasion," 2018.
4. A. Al Mahmud, P. Dadlani, O. Mubin, S. Shahid, C. Midden, and O. Moran, "iParrot: Towards designing a persuasive agent for energy conservation," in International Conference on Persuasive Technology. Springer, 2007, pp. 64–67.
5. S. Zhou, Z. Zhang, and T. Bickmore, "Adapting a persuasive conversational agent for the Chinese culture," in 2017 international conference on culture and computing (culture and computing). IEEE, 2017, pp. 89–96.
6. S. Vossen, J. Ham, and C. Midden, "Social influence of a persuasive agent: the role of agent embodiment and evaluative feedback," in Proceedings of the 4th International Conference on Persuasive Technology, 2009, pp. 1–7.
7. D. Schulman and T. Bickmore, "Persuading users through counseling dialogue with a conversational agent," in Proceedings of the 4th international conference on persuasive technology, 2009, pp. 1–8.
8. R. F. Khan and A. Sutcliffe, "Attractive agents are more persuasive," International Journal of Human-Computer Interaction, vol. 30, no. 2, pp. 142–150, 2014.
9. R. B. Kantharaju, D. De Franco, A. Pease, and C. Pelachaud, "Is two better than one? effects of multiple agents on user persuasion," in Proceedings of the 18th International Conference on Intelligent Virtual Agents, 2018, pp. 255–262.
10. M. R. Barrick and M. K. Mount, "The big five personality dimensions and job performance: a meta-analysis," Personnel psychology, vol. 44, no. 1, pp. 1–26, 1991.
11. S. Thomas, Y. Ferstl, R. McDonnell, and C. Ennis, "Investigating how speech and animation realism influence the perceived personality of virtual characters and agents," in 2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). IEEE, 2022, pp
12. W. Xu, Y. Ishikawa, and R. Legaspi, "The utility of personality types for personalizing persuasion," in International Conference on Persuasive Technology. Springer, 2022, pp. 240–254.
13. Epic Games, "Unreal engine." [Online]. Available: https://www.unrealengine.com
14. H. J. Seltman, "Experimental design and analysis," 2012.
15. D. W. Rook and R. J. Fisher, "Normative influences on impulsive buying behavior," Journal of consumer research, vol. 22, no. 3, pp. 305–313, 1995.
16. T. K. Chan, C. M. Cheung, and Z. W. Lee, "The state of online impulse- buying research: A literature analysis," Information & Management, vol. 54, no. 2, pp. 204–217, 2017.
17. A. Shahjehan and J. A. Qureshi, "Personality and impulsive buying behaviors. a necessary condition analysis," Economic research-Ekonomska istrazivanja ˇ, vol. 32, no. 1, pp. 1060–1072, 2019.
18. C. Hui, "Personality's influence on the relationship between online word- of-mouth and consumers' trust in shopping website," journal of software, vol. 6, no. 2, pp. 265–272, 2011.
19. D. Zell, C. McGrath, and C. M. Vance, "Examining the interaction of extroversion and network structure in the formation of effective informal support networks," Journal of Behavioral and Applied Management, vol. 15, no. 2, pp. 59–81, 2014.
20. V. Ozbek,¨ U. Alnıac ̧ık, F. Koc, M. E. Akkılıc ̧, and E. Kas ̧, "The impact¨ of personality on technology acceptance: A study on smart phone users," Procedia-Social and Behavioral Sciences, vol. 150, pp. 541–551, 2014.
21. S. S. Webber, S. C. Payne, and A. B. Taylor, "Personality and trust fosters service quality," Journal of business and psychology, vol. 27, no. 2, pp. 193–203, 2012.
22. L. P. Maricutoiu, "A meta-analysis on the antecedents and consequences of computer anxiety," Procedia-Social and Behavioral Sciences, vol. 127, pp. 311–315, 2014.
23. R. E. Bawack, S. F. Wamba, and K. D. A. Carillo, "Exploring the role of personality, trust, and privacy in customer experience performance during voice shopping: Evidence from sem and fuzzy set qualitative comparative analysis," International Journal of Information Management, vol. 58, p. 102309, 2021.
24. A. S. Gerber, G. A. Huber, D. Doherty, C. M. Dowling, and C. Panagopou- los, "Big five personality traits and responses to persuasive appeals: Results from voter turnout experiments," Political Behavior, vol. 35, no. 4, pp. 687–728, 2013.
25. K. Sumi, "Learning story marketing through practical experience of story creation system," in Joint International Conference on Interactive Digital Storytelling. Springer, 2010, pp. 98–110.
26. K. Sumi and M. Nagata, "Evaluating a virtual agent as persuasive technology, psychology of persuasion, janos csap ́o and andor magyar eds," 2010.

Appendix A. Table of Conversation

Audio No.	Conversation	Virtual agent(s)	Subject/ Participant
Subject filled out a pre-questionnaire before interaction with the virtual agents.			
1	"Hello. I am Danielle. Nice to meet you. Recently, more and more people are taking supplements to supplement nutrients other than meals. Do you usually take supplements?"	Danielle	-
			Answers something
2	"I am interested in royal jelly. Royal jelly is a special food that only the queen bee is allowed to eat."	Danielle	-
The experimenter pressed either 3-1 or 3-2 randomly from the keyboard.			
3-1	"According to what I heard, royal jelly is good for health by boosting immunity, and it can cause anaphylactic shock."	Danielle	-
3-2	"According to what I heard, there are two theories about royal jelly: it can cause anaphylactic shock, and it is good for your health by boosting your immune system."	Danielle	-
From the pre-questionnaire, there was a highlight question of, "Do you think Royal Jelly is good for your health?" It determined the keyboard operation. *If he/she answers:* *Yes → an experimenter pressed 4-1 from the keyboard* *No → an experimenter pressed 4-2 from the keyboard*			
4-1	"I think the theory that royal jelly can cause anaphylactic shock is correct."	Danielle	-
	"I think so, too."	(1) Stephan, (2) Pia, (3) Kellan, (4) Aoi, (5) Kendra	
			Answers something
4-2	"I think the theory that royal jelly is good for health by boosting the immune system is correct."	Danielle	-
	"I think so, too."	(1) Stephan, (2) Pia, (3) Kellan, (4) Aoi, (5) Kendra	
			Answers something
5	"I agree. Do you agree?"	Danielle	-
			Answers something

Authors Introduction

Mrs. Pinkie Anggia

Mrs. Pinkie Anggia is a Doctoral student in the Graduate School of Systems Information Science, Future University Hakodate, Japan. She received her Master's degree from the University of Indonesia. Formerly, she worked at the U.S. Embassy Jakarta, Goethe-Institut Indonesien, and the University of Indonesia.

Prof. Kaoru Sumi

Prof. Kaoru Sumi is a professor in the Department of Media Architecture, Future University Hakodate, Japan. Prof. Sumi received her Ph.D. from the University of Tokyo. She is currently working on Artificial Intelligence in Education, Affective Computing, Interactive Digital storytelling. Formerly, she worked at ATR MI&C Research Laboratories, Communications Research Laboratory, and Osaka University, National Institute of Information and Communications Technology (NICT), and Hitotsubashi University.

Quasi-static Stability Analysis of Frictionless Planar Enveloping Grasps
(Analysis of curvature effects at contact points)

Takayoshi Yamaada, Junya Sato, Kazuaki Ito, Hidehiko Yamamoto
Dept. of Mechanical Engineering, Gifu University,
Yanagido 1-1, Gifu 501-1193, Japan
E-mail: yamat@gifu-u.ac.jp

Abstract

In this paper, we discuss grasp stability of frictionless planar enveloping grasps. The grasp stability is based on potential energy of the grasp system replaced with an elastic system. A stiffness matrix of the grasp system is derived. The grasp stability is evaluated by the eigenvalues of the matrix. We show that the matrix depends on grasp positions, grasp forces, local curvatures, etc., at contact points. Moreover, we analyze curvature effects on the grasp system by differentiating the matrix by the curvatures. We show that the derivative is negative semi definite.

Keywords: Grasp stability, Enveloping grasp, Grasp stiffness matrix, Curvature effects

1. Introduction

Humans can grasp and manipulate various types of objects dexterously. In various fields like as production lines including handling tasks, picking tasks, assembly tasks, etc., manual hand works are remained. Therefore, the human-like dexterous skills are required for mechanical functions of robots. In recent years, in order to obtain appropriate grasps roughly, deep learning methods are introduced. On the other hands, in traditional ways, grasp and manipulation are analyzed from the viewpoint of mechanics, kinematics and dynamics in detail. As one of the methods, grasp stability based on potential energy of the grasp system replaced with an elastic system is investigated in order to derive grasp evaluation. In grasp forms, pinching grasp by fingertips, enveloping grasp by finger link surfaces and so on can be considered.

Ref. [1] analyzed grasp stability of pinching grasps. Local curvature effects on the stability were also investigated. Ref. [2] investigated grasp position effects on the stability, proposed an automatic generation of optimal grasp. Ref. [3] analyzed frictionless enveloping grasps by replacing joint displacement and finger surface displacement with elastic model (Figure 1).

In this paper, we discuss grasp stability of frictionless planar enveloping grasps. The stiffness matrix of the grasp system is derived. The grasp stability is evaluated by the eigenvalues of the matrix. We show that the matrix depends on grasp positions, grasp forces, local curvatures, etc., at contact points. Moreover, we analyze curvature effects on the grasp system by differentiating the matrix by the curvatures.

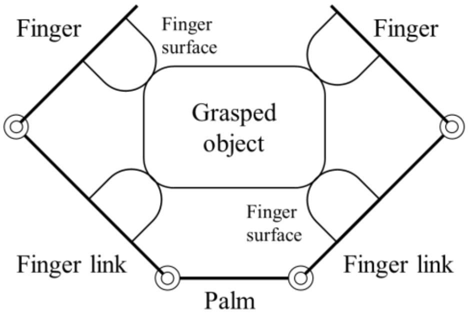


Figure 1: A planar enveloping grasp

2. Problem Formulation

2.1. *Symbols*

Symbol i is finger number, j is joint and link number. Some coordinates are defined as shown in Figure 2. Symbol Σ_b denotes a base coordinate frame on the system. Symbol Σ_o is an object coordinate frame, Σ_{bo} is its initial pose. Symbol Σ_{i0} is a i-th finger base coordinate frame, Σ_{ij} is the *j*-th link coordinate frame of the *i*-th finger, Σ_{bij} is the initial pose of Σ_{ij}. Symbols Σ_{Cfij} and Σ_{Coij} are contact coordinate frames on the finger and the object, Σ_{Lfij} and Σ_{Loij} are the initial pose of Σ_{Cfij} and Σ_{Coij} , respectively. Symbol Σ_{ijp} is a deformed coordinate frame of Σ_{Lfij} . Vectors and matrices are shown in Appendix.

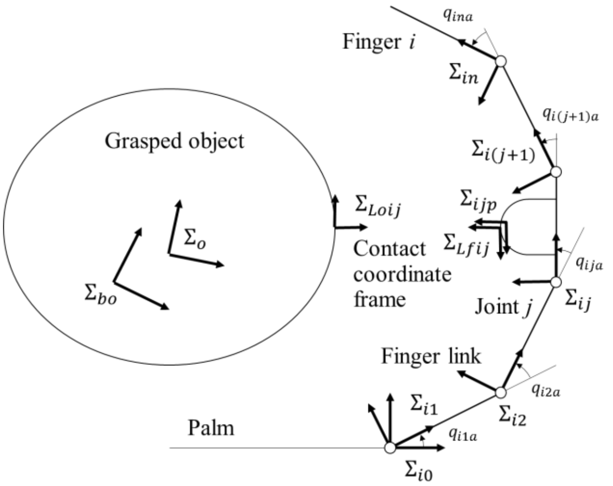


Figure 2: Coordinate frames on the object and the finger

2.2. *Joint and finger surface elastics*

We consider that the *i*-th finger is constructed of serial links with *n* joints. Position of the *j*-th joint (active joint), q_{ija}, is represented as

$$q_{ija} = q_{nija} + q_{cija} + q_{dija} \in \mathbb{R} \tag{1}$$

$$\begin{cases} q_{nija} & \text{Natural position} \\ q_{cija} & \text{Compression by the initial force} \\ q_{dija} & \text{Displacement by the object displacement} \end{cases}$$

The homogeneous transformation matrix of the *j*-th link is represented as

$$\begin{aligned} &{}^{i(j-1)}A_{ij}(q_{ija}) \\ &= \begin{cases} {}^{i(j-1)}A_{bij}(q_{nija}, q_{cija})A_r(q_{dija}) & \text{for revolute joint} \\ {}^{i(j-1)}A_{bij}(q_{nija}, q_{cija})A_t(q_{dija}\boldsymbol{u}_1) & \text{for prismatic joint} \end{cases} \end{aligned} \tag{2}$$

The link surface contacts on the object. Deformation of the link surface is represented as the pose displacement of the surface coordinate (passive joint) and represented by the following form.

$$\boldsymbol{q}_{ijp} = \boldsymbol{q}_{nijp} + \boldsymbol{q}_{cijp} + \boldsymbol{q}_{dijp} \in \mathbb{R}^3 \tag{3}$$

$$\begin{cases} \boldsymbol{q}_{nijp} & \text{Natural position} \\ \boldsymbol{q}_{cijp} & \text{Compression by the initial force} \\ \boldsymbol{q}_{dijp} & \text{Displacement by the object displacement} \end{cases}$$

The displacement is represented by the following homogeneous transformation matrix.

$$\begin{gathered} {}^{ij}A_{Lfij}(\boldsymbol{q}_{ijp}) \coloneqq {}^{ij}A_{ijp}\,{}^{ijp}A_{Lfij}(\boldsymbol{q}_{dijp}) \\ {}^{ijp}A_{Lfij}(\boldsymbol{q}_{dijp}) \coloneqq A_{tr}(\boldsymbol{q}_{dijp}), \qquad {}^{ijp}A_{Lfij}(0_{3\times 1}) = I_3 \\ \lim_{\boldsymbol{q}_{dijp}\to 0} \Sigma_{ijp} = \Sigma_{Lfij} \end{gathered} \tag{4}$$

Displacement of the *j*-th link is represented by

$$\boldsymbol{q}_{dij} \coloneqq \left[q_{dija}, \boldsymbol{q}_{dijp}^T\right]^T \in \mathbb{R}^4 \tag{5}$$

2.3. *Potential energy of the j-th link*

Elastic coefficient of the joint displacement is defined as $s_{ija} \in \mathbb{R}$. The potential energy of the *j*-th joint is obtained by the following form:

$$U_{ija}(q_{dija}) \coloneqq \frac{1}{2}s_{ija}(q_{cija} + q_{dija})^2 \tag{6}$$

Initial joint torque τ_{ija} is obtained by

$$\tau_{ija} \coloneqq s_{ija}q_{cija} \tag{7}$$

Symbol $S_{ijp} \in \mathbb{R}^{3\times 3}$ represents stiffness coefficient of the elasticity of the link surface displacement. The potential energy of the *i*-th finger link is given by

$$U_{ijp}(\boldsymbol{q}_{dijp}) \coloneqq \frac{1}{2}\left[\boldsymbol{q}_{cijp} + \boldsymbol{q}_{dijp}\right]^T S_{ijp}\left[\boldsymbol{q}_{cijp} + \boldsymbol{q}_{dijp}\right] \tag{8}$$

Initial contact force $\boldsymbol{\tau}_{ijp}$ is obtained by

$$\boldsymbol{\tau}_{ijp} \coloneqq S_{ijp}\boldsymbol{q}_{cijp} \tag{9}$$

The potential energy of the *j*-th joint is given as

$$U_{ij}(\boldsymbol{q}_{dij}) \coloneqq U_{ija}(q_{dija}) + U_{ijp}(\boldsymbol{q}_{dijp}) \tag{10}$$

The potential energy of the *i*-th finger is obtained by

$$U_i(\boldsymbol{q}_{di}) \coloneqq \sum_{j=1}^{n} U_{ij}(\boldsymbol{q}_{dij}) \tag{21}$$

where

$$\begin{aligned} \boldsymbol{q}_{di} &\coloneqq [\boldsymbol{q}_{di1}^T, \cdots, \boldsymbol{q}_{din}^T]^T \\ &= \left[q_{di1a}, \boldsymbol{q}_{di1p}^T, \cdots, q_{dina}, \boldsymbol{q}_{dinp}^T\right]^T \in \mathbb{R}^{4n} \end{aligned} \tag{32}$$

In the case that the number of fingers is m, the potential energy of the grasp system is obtained by

$$U(\boldsymbol{q}_d) \coloneqq \sum_{i=1}^{n} U_i(\boldsymbol{q}_{di}), \qquad \boldsymbol{q}_d \coloneqq [\boldsymbol{q}_{d1}^T, \cdots, \boldsymbol{q}_{dm}^T]^T \tag{13}$$

2.4. *Joint and finger surface elastics*

In the case that the *j*-th link surface contacts on the object, we obtain the following constraint:

$$\begin{gathered} {}^{b}A_{bo}\,{}^{bo}A_o(\boldsymbol{\varepsilon}_o)\,{}^{o}A_{Loij}\,{}^{Loij}A_{Coij}(\alpha_{oij}) \\ = {}^{b}A_{i0}\,{}^{i0}A_{i1}(q_{i1a}) \times \cdots \times {}^{i(j-1)}A_{ij}(q_{ija}) \\ \times {}^{ij}A_{Lfij}(\boldsymbol{q}_{ijp})\,{}^{Lfij}A_{Cfij}(\alpha_{fij})\,{}^{Cfij}A_{Coij} \\ (j = 1,2,\cdots,n) \end{gathered} \tag{14}$$

where parameter $\boldsymbol{\varepsilon}_o$ represents position and orientation of the object. Parameters α_{oij} and α_{fij} represent surface displacement of the contact points on the object and the link.

$$\begin{aligned}{}^{Loij}A_{Coij}\left(\alpha_{oij}\right) &:= {}^{Loij}A_{\kappa oij}A_r\left(\kappa_{oij}\alpha_{oij}\right){}^{Loij}A_{\kappa oij}^{-1}\\ {}^{Lfij}A_{Cfij}\left(\alpha_{fij}\right) &:= {}^{Lfij}A_{\kappa fij}A_r\left(\kappa_{fij}\alpha_{fij}\right){}^{Lfij}A_{\kappa fij}^{-1}\\ {}^{Loij}A_{\kappa oij} := A_t\left(-\kappa_{oij}^{-1}\boldsymbol{u}_1\right), &\quad {}^{Lfij}A_{\kappa fij} := A_t\left(-\kappa_{fij}^{-1}\boldsymbol{u}_1\right)\end{aligned} \tag{15}$$

Curvature κ is given as the following characteristics:

$$\begin{cases}\kappa > 0 & \text{convex surface}\\ \kappa = 0 & \text{flat surface}\\ \kappa < 0 & \text{concave surface}\end{cases} \tag{16}$$

Relation of the contact coordinate frame between the link and the object surfaces is set as the following condition:

$${}^{Cfij}A_{Coij} = A_r(\pi) \tag{17}$$

From (15), we have

$$\begin{aligned}A_{tr}\left(\boldsymbol{q}_{dijp}\right) &= {}^{ijp}A_{ij}{}^{i(j-1)}A_{ij}^{-1}\left(q_{ija}\right) \times \cdots\\ &\times {}^{i0}A_{i1}^{-1}(q_{i1a}){}^{i0}A_{bo}{}^{bo}A_o(\boldsymbol{\varepsilon}_o)\,{}^{o}A_{Loij}\\ &\times {}^{Loij}A_{Coij}\left(\alpha_{oij}\right){}^{Cfij}A_{Coij}^{-1}\,{}^{Lfij}A_{Cfij}^{-1}\left(\alpha_{fij}\right)\end{aligned} \tag{18}$$

The displacement on the link surface, $\boldsymbol{q}_{dijp}$, is given by a function of the object displacement $\boldsymbol{\varepsilon}_o$, joint position displacement $q_{di1a}, \cdots, q_{dija}$, contact position displacement $\boldsymbol{\alpha}_{ij}$.

$$\boldsymbol{q}_{dijp}\left(\boldsymbol{\varepsilon}_o, \boldsymbol{\alpha}_{ij}, q_{di1a}, \cdots, q_{dija}\right),\quad (j = 1,2,\cdots,n) \tag{19}$$

where

$$\boldsymbol{\alpha}_{ij} := \begin{bmatrix}\alpha_{oij}\\ \alpha_{fij}\end{bmatrix} \in \mathbb{R}^2 \tag{20}$$

Consequently, the potential energy of the i-th finger is represented as the following form:

$$\begin{aligned}U_{iq}(\boldsymbol{\varepsilon}_o, \boldsymbol{\alpha}_i, \boldsymbol{\beta}_i) &:= U_i\big(q_{di1a}, \boldsymbol{q}_{di1p}^T(\boldsymbol{\varepsilon}_o, \boldsymbol{\alpha}_{i1}, q_{di1a}), \cdots,\\ &\qquad q_{dina}, \boldsymbol{q}_{dinp}^T(\boldsymbol{\varepsilon}_o, \boldsymbol{\alpha}_{in}, q_{di1a}, \cdots, q_{dina})\big)\\ &= \sum\nolimits_{j=1}^{n}\Big\{U_{ija}\left(q_{dija}\right)\\ &\quad + U_{ijp}\left(\boldsymbol{q}_{dijp}\left(\boldsymbol{\varepsilon}_o, \boldsymbol{\alpha}_{ij}, q_{di1a}, \cdots, q_{dija}\right)\right)\Big\}\end{aligned} \tag{21}$$

where

$$\boldsymbol{\alpha}_i := \begin{bmatrix}\boldsymbol{\alpha}_{i1}\\ \vdots\\ \boldsymbol{\alpha}_{in}\end{bmatrix} \in \mathbb{R}^{2n}, \quad \boldsymbol{\beta}_i := \begin{bmatrix}q_{di1a}\\ \vdots\\ q_{dina}\end{bmatrix} \in \mathbb{R}^n \tag{22}$$

3. Partial derivative of potential energy

3.1. *The first derivative of the potential energy*

The first partial derivative of the potential energy is derived and its initial condition is considered, then we have the followings:

$$\begin{aligned}\left.\frac{\partial U_{iq}(\boldsymbol{\varepsilon}_o, \boldsymbol{\alpha}_i, \boldsymbol{\beta}_i)}{\partial \boldsymbol{\varepsilon}_o}\right|_0 &= \sum_{l=1}^{n}{}^{Lfil}B_o^T\boldsymbol{\tau}_{ilp}\\ \left.\frac{\partial U_{iq}(\boldsymbol{\varepsilon}_o, \boldsymbol{\alpha}_i, \boldsymbol{\beta}_i)}{\partial \boldsymbol{\alpha}_{ij}}\right|_0 &= K_{ij}^T\boldsymbol{\tau}_{ijp}, \quad K_{ij} := \begin{bmatrix}-\boldsymbol{u}_2 & -\boldsymbol{u}_2\\ \kappa_{oij} & -\kappa_{fij}\end{bmatrix}\\ \left.\frac{\partial U_{iq}(\boldsymbol{\varepsilon}_o, \boldsymbol{\alpha}_i, \boldsymbol{\beta}_i)}{\partial q_{dija}}\right|_0 &= \tau_{ija} - \boldsymbol{v}_\zeta^T\sum_{l=j}^{n}{}^{Lfil}B_{ij}^T\boldsymbol{\tau}_{ilp}\end{aligned} \tag{23}$$

3.2. *The second derivative of the potential energy*

The second partial derivative of the potential energy is derived and its initial condition is considered, then we have the followings:

$$\begin{aligned}&\frac{\partial^2 U_{iq}(\boldsymbol{\varepsilon}_o, \boldsymbol{\alpha}_i, \boldsymbol{\beta}_i)}{\partial \boldsymbol{\varepsilon}_o \partial \boldsymbol{\varepsilon}_o^T}\\ &= \sum_{l=1}^{n}\left\{{}^{Lfil}B_o^T S_{ilp}{}^{Lfil}B_o + \boldsymbol{v}_\zeta\left[\boldsymbol{\tau}_{ilp}^T I_{23}^T{}^{Lfil}\boldsymbol{p}_o\right]\boldsymbol{v}_\zeta^T\right\}\\ &\left.\frac{\partial^2 U_{iq}(\boldsymbol{\varepsilon}_o, \boldsymbol{\alpha}_i, \boldsymbol{\beta}_i)}{\partial \boldsymbol{\alpha}_{ij} \partial \boldsymbol{\varepsilon}_o^T}\right|_0 = K_{ij}^T S_{ijp}{}^{Lfij}B_o + \begin{bmatrix}1\\ 1\end{bmatrix}\left[\boldsymbol{\tau}_{ijp}^T\boldsymbol{v}_x\right]\boldsymbol{v}_\zeta^T\\ &\left.\frac{\partial^2 U_{iq}(\boldsymbol{\varepsilon}_o, \boldsymbol{\alpha}_i, \boldsymbol{\beta}_i)}{\partial q_{dija} \partial \boldsymbol{\varepsilon}_o^T}\right|_0\\ &\qquad = -\sum_{l=j}^{n}\left[\boldsymbol{v}_\zeta^T{}^{Lfil}B_{ij}^T S_{ilp}{}^{Lfil}B_o + \boldsymbol{\tau}_{ilp}^T I_{23}^T\Omega I_{23}{}^{Lfil}B_o\right]\\ &\left.\frac{\partial^2 U_{iq}(\boldsymbol{\varepsilon}_o, \boldsymbol{\alpha}_i, \boldsymbol{\beta}_i)}{\partial \boldsymbol{\alpha}_{ik} \partial \boldsymbol{\alpha}_{ij}^T}\right|_0\\ &= \begin{cases}K_{ik}^T S_{ijp}K_{ij} + \left[\boldsymbol{\tau}_{ijp}^T\boldsymbol{v}_x\right]\begin{bmatrix}\kappa_{oij} & \kappa_{oij}\\ \kappa_{oij} & -\kappa_{fij}\end{bmatrix} & (1 \le k = j \le n)\\ 0_{2\times 2} & (\text{otherwise})\end{cases}\\ &\left.\frac{\partial^2 U_{iq}(\boldsymbol{\varepsilon}_o, \boldsymbol{\alpha}_i, \boldsymbol{\beta}_i)}{\partial q_{dika} \partial \boldsymbol{\alpha}_{ij}^T}\right|_0\\ &= \begin{cases}-\boldsymbol{v}_\zeta^T{}^{Lfij}B_{ik}^T S_{ijp}K_{ij} - \boldsymbol{\tau}_{ijp}^T\boldsymbol{v}_x\begin{bmatrix}1 & 1\end{bmatrix} & (1 \le k \le j \le n)\\ 0_{1\times 2} & (\text{otherwise})\end{cases}\\ &\left.\frac{\partial^2 U_{iq}(\boldsymbol{\varepsilon}_o, \boldsymbol{\alpha}_i, \boldsymbol{\beta}_i)}{\partial q_{dika} \partial q_{dija}}\right|_0\\ &= \begin{cases}s_{ija} + \displaystyle\sum_{l=j}^{n}\left[\boldsymbol{v}_\zeta^T{}^{Lfil}B_{ik}^T S_{ilp}{}^{Lfil}B_{ij}\boldsymbol{v}_\zeta + \boldsymbol{\tau}_{ilp}^T I_{23}^T{}^{Lfil}\boldsymbol{p}_{ij}\right]\\ \qquad (1 \le l \le k = j \le n)\\ \displaystyle\sum_{l=j}^{n}\left[\boldsymbol{v}_\zeta^T{}^{Lfil}B_{ik}^T S_{ilp}{}^{Lfil}B_{ij}\boldsymbol{v}_\zeta + \boldsymbol{\tau}_{ilp}^T I_{23}^T{}^{Lfil}\boldsymbol{p}_{ij}\right]\\ \qquad (1 \le l \le k < j \le n)\end{cases}\end{aligned} \tag{24}$$

4. Frictionless enveloping grasp

4.1. *Constraints of frictionless contact*

In the case that the i-th finger surface contacts on the objects with frictionless condition, contact point displacement $\boldsymbol{\alpha}_i$ and joint position displacement $\boldsymbol{\beta}_i$ shift to the position locally minimizing the energy. Consequently, we have the following constraint:

$$\begin{bmatrix} \dfrac{\partial U_{iq}(\boldsymbol{\varepsilon}_o, \boldsymbol{\alpha}_i, \boldsymbol{\beta}_i)}{\partial \boldsymbol{\alpha}_i} \\ \dfrac{\partial U_{iq}(\boldsymbol{\varepsilon}_o, \boldsymbol{\alpha}_i, \boldsymbol{\beta}_i)}{\partial \boldsymbol{\beta}_i} \end{bmatrix} = \begin{bmatrix} 0_{2n} \\ 0_n \end{bmatrix} \tag{25}$$

From this condition, we have $3n$ constraints, then the freedom of the system is given only as the parameter $\boldsymbol{\varepsilon}_o$.

$$U_{iq}^{fs}(\boldsymbol{\varepsilon}_o) := U_{iq}\left(\boldsymbol{\varepsilon}_o, \boldsymbol{\alpha}_i(\boldsymbol{\varepsilon}_o), \boldsymbol{\beta}_i(\boldsymbol{\varepsilon}_o)\right) \tag{26}$$

The first partial derivative is derived, its initial condition is considered. Considering the constraint of Eq. (25), we have the following gradient:

$$G_i^{fs} := \left.\frac{\partial U_{iq}^{fs}(\boldsymbol{\varepsilon}_o)}{\partial \boldsymbol{\varepsilon}_o}\right|_0 = \sum_{l=1}^{n} {}^{Lfil}B_o^T \boldsymbol{\tau}_{ilp} = \sum_{l=1}^{n} {}^{o}W_{Lfil} {}^{Lfil}\boldsymbol{f} \tag{27}$$

where the symbol $\boldsymbol{\tau}_{ilp}$ is given by

$$\boldsymbol{\tau}_{ilp} = I_{23}^T {}^{Lfil}\boldsymbol{f}, \quad {}^{Lfil}\boldsymbol{f} = \begin{bmatrix} {}^{Lfil}f_x \\ {}^{Lfil}f_y \end{bmatrix}, \quad {}^{Lfil}f_x < 0, \quad {}^{Lfil}f_y = 0 \tag{28}$$

The second partial derivative is given as

$$H_i^{fs} := \left.\frac{\partial^2 U_{iq}^{fs}(\boldsymbol{\varepsilon}_o)}{\partial \boldsymbol{\varepsilon}_o \partial \boldsymbol{\varepsilon}_o^T}\right|_0 = U_{iq,\varepsilon\varepsilon} + Q_i^{fs} \begin{bmatrix} U_{iq,\varepsilon\alpha} \\ U_{iq,\varepsilon\beta} \end{bmatrix} \tag{29}$$

where the symbol Q_i^{fs} express

$$\begin{aligned} Q_i^{fs} &:= \begin{bmatrix} \left.\dfrac{\partial \boldsymbol{\alpha}_i^T}{\partial \boldsymbol{\varepsilon}_o}\right|_0 & \left.\dfrac{\partial \boldsymbol{\beta}_i^T}{\partial \boldsymbol{\varepsilon}_o}\right|_0 \end{bmatrix} \\ &= -\begin{bmatrix} U_{iq,\alpha\varepsilon} & U_{iq,\beta\varepsilon} \end{bmatrix} \begin{bmatrix} U_{iq,\alpha\alpha} & U_{iq,\beta\alpha} \\ U_{iq,\alpha\beta} & U_{iq,\beta\beta} \end{bmatrix}^{-1} \end{aligned} \tag{30}$$

For example, the symbol $U_{iq,\alpha\varepsilon}$ means the second partial derivative of $U_{iq}(\boldsymbol{\varepsilon}_o, \boldsymbol{\alpha}_i, \boldsymbol{\beta}_i)$ by $\boldsymbol{\alpha}_i$ and $\boldsymbol{\varepsilon}_o$. The potential energy of the grasp system is given by the following the formula:

$$U^{fs}(\boldsymbol{\varepsilon}_o) := \sum_{i=1}^{m} U_{iq}^{fs}(\boldsymbol{\varepsilon}_o) \tag{31}$$

Total grasp wrench (force and moment) of the grasp system and the grasp stiffness matrix are obtained by

$$\begin{aligned} G^{fs} &:= \left.\frac{\partial U^{fs}(\boldsymbol{\varepsilon}_o)}{\partial \boldsymbol{\varepsilon}_o}\right|_0 = \sum_{i=1}^{m} G_i^{fs} \\ H^{fs} &:= \left.\frac{\partial^2 U^{fs}(\boldsymbol{\varepsilon}_o)}{\partial \boldsymbol{\varepsilon}_o \partial \boldsymbol{\varepsilon}_o^T}\right|_0 = \sum_{i=1}^{m} H_i^{fs} \end{aligned} \tag{32}$$

The matrix H^{fs} depends on contact position, contact direction, contact force, local curvature at contact points.

5. Effect of local curvatures at contact points

5.1. *Partial derivative of the local curvatures*

The grasp stiffness matrix is partially differentiated by the local curvature, then we have the following formula:

$$\begin{aligned} \frac{\partial H_i^{fs}}{\partial \kappa_{oij}} &= \frac{\partial U_{iq,\varepsilon\varepsilon}}{\partial \kappa_{oij}} + \left\{ \frac{\partial}{\partial \kappa_{oij}} \begin{bmatrix} U_{iq,\varepsilon\alpha} \\ U_{iq,\varepsilon\beta} \end{bmatrix} \right\}^T \left[Q_i^{fs}\right]^T \\ &\quad + Q_i^{fs} \left\{ \frac{\partial}{\partial \kappa_{oij}} \begin{bmatrix} U_{iq,\varepsilon\alpha} \\ U_{iq,\varepsilon\beta} \end{bmatrix} \right\} \\ &\quad + Q_i^{fs} \left\{ \frac{\partial}{\partial \kappa_{oij}} \begin{bmatrix} U_{iq,\alpha\alpha} & U_{iq,\beta\alpha} \\ U_{iq,\alpha\beta} & U_{iq,\beta\beta} \end{bmatrix} \right\} \left[Q_i^{fs}\right]^T \\ &= \boldsymbol{a}\boldsymbol{b}_1{}^T + \boldsymbol{b}_1\boldsymbol{a}^T + \left({}^{Lfij}\boldsymbol{f}^T\boldsymbol{u}_1\right)\boldsymbol{b}_1\boldsymbol{b}_1{}^T \\ &= \left({}^{Lfij}\boldsymbol{f}^T\boldsymbol{u}_1\right)\boldsymbol{b}_1\boldsymbol{b}_1{}^T \preccurlyeq 0_{3\times 3} \\ \frac{\partial H_i^{fs}}{\partial \kappa_{fij}} &= -\boldsymbol{a}\boldsymbol{b}_2{}^T - \boldsymbol{b}_2\boldsymbol{a}^T + \left({}^{Lfij}\boldsymbol{f}^T\boldsymbol{u}_1\right)\boldsymbol{b}_2\boldsymbol{b}_2{}^T \\ &= \left({}^{Lfij}\boldsymbol{f}^T\boldsymbol{u}_1\right)\boldsymbol{b}_2\boldsymbol{b}_2{}^T \preccurlyeq 0_{3\times 3} \end{aligned} \tag{33}$$

The vectors $\boldsymbol{a}$, $\boldsymbol{b}_1$, $\boldsymbol{b}_2 \in \mathbb{R}^3$ are given by

$$\begin{aligned} &\boldsymbol{a} := {}^{Lfij}B_o^T S_{ijp} \boldsymbol{v}_\zeta \\ &+ Q_i^{fs} \begin{bmatrix} \begin{bmatrix} 0_{2(j-1)\times 1} \\ K_{ij}^T S_{ijp}\boldsymbol{v}_\zeta + \boldsymbol{u}_2\left[{}^{Lfij}\boldsymbol{f}^T\boldsymbol{u}_1\right] \\ 0_{2(n-j)\times 1} \end{bmatrix} \\ -\begin{bmatrix} \boldsymbol{v}_\zeta^T {}^{Lfij}B_{i1}^T S_{ijp}\boldsymbol{v}_\zeta \\ \vdots \\ \boldsymbol{v}_\zeta^T {}^{Lfij}B_{ij}^T S_{ijp}\boldsymbol{v}_\zeta \\ 0_{(n-j)\times 1} \end{bmatrix} \end{bmatrix} = \boldsymbol{0}, \\ &\boldsymbol{b}_1 := Q_i^{fs} \begin{bmatrix} 0_{2(j-1)\times 1} \\ \boldsymbol{u}_1 \\ 0_{2(n-j)\times 1} \\ 0_{n\times 1} \end{bmatrix}, \quad \boldsymbol{b}_2 := Q_i^{fs} \begin{bmatrix} 0_{2(j-1)\times 1} \\ \boldsymbol{u}_2 \\ 0_{2(n-j)\times 1} \\ 0_{n\times 1} \end{bmatrix} \end{aligned} \tag{34}$$

Because we have ${}^{Lfij}\boldsymbol{f}^T\boldsymbol{u}_1 < 0$, the derivatives $\frac{\partial H_i^{fs}}{\partial \kappa_{oij}}$ and $\frac{\partial H_i^{fs}}{\partial \kappa_{fij}}$ are negative semi definite. The value of the local curvature is smaller, the grasp stability is higher.

6. Conclusions

This paper analyzed frictionless planar enveloping grasps. Not only joint displacement but also link surface displacement is replaced with elastic feature. Contact constraints between the finger links and the grasped object were derived, and independent parameters were clarified. By using independent parameters, the potential energy of the grasp system was derived. In the case of frictionless contact, the wrench vectors were obtained by the first partial derivative of the energy, the grasp stiffness matrix was obtained by the second partial derivative. By using partial derivative of the stiffness matrix by local curvature at contact points, the effect of the curvature on the grasp stability was clarified.

References

[1] T. Yamada, R. Johansson, A. Robertsson, and H. Yamamoto, Static Stability Analysis of a Planar Object Grasped by Multifingers with Three Joints, Robotics, 4, 464-491, doi:10.3390/robotics4040464, 2015.

[2] T. Yamada, K. Niwa, H. Yamamoto, H. Kawasaki, and T. Mouri, Automatic generation of grasp positions using the partial differentiation of the grasp stiffness matrix, J. of Robotic and Mechatronic Systems, Vol. 2, No. 2, pp. 1-15, 2017.

[3] T. Yamada, Y. Kishi, J. Sato, K. Ito, H. Yamamoto, Quasistatic Stability of Frictionless Planar Enveloping Grasps, Proc. of JSME Robomech2019 (CD-ROM), 2P1-H05, 2019.

Appendix A

In two dimensions, a homogeneous transformation matrix of a frame Σ_b with respect to a frame Σ_a is represented by the following form:

$$ {}^{a}A_b := \begin{bmatrix} {}^{a}R_b & {}^{a}\boldsymbol{p}_b \\ 0_{1\times 2} & 1 \end{bmatrix} \in \mathbb{R}^{3\times 3} \tag{35} $$

where ${}^{a}\boldsymbol{p}_b \in \mathbb{R}^2$ is a position vector, ${}^{a}R_b \in \mathbb{R}^{2\times 2}$ is a rotation matrix expressing relative orientation. The following symbols are also defined in this paper.

$$ \begin{gathered} \boldsymbol{u}_1 := \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{u}_2 := \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \\ \boldsymbol{v}_x := \begin{bmatrix} \boldsymbol{u}_1 \\ 0 \end{bmatrix}, \quad \boldsymbol{v}_y := \begin{bmatrix} \boldsymbol{u}_2 \\ 0 \end{bmatrix}, \quad \boldsymbol{v}_\zeta := \begin{bmatrix} 0_{2\times 1} \\ 1 \end{bmatrix} \\ \mathrm{Rot}(\zeta) := \begin{bmatrix} \cos\zeta & -\sin\zeta \\ \sin\zeta & \cos\zeta \end{bmatrix}, \\ \Omega := \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \mathrm{Rot}\left(\frac{\pi}{2}\right) \\ A_t(\boldsymbol{x}) := \begin{bmatrix} I_2 & \boldsymbol{x} \\ 0_{1\times 2} & 1 \end{bmatrix}, \qquad A_r(\zeta) := \begin{bmatrix} \mathrm{Rot}(\zeta) & 0_{2\times 1} \\ 0_{1\times 2} & 1 \end{bmatrix} \\ A_{tr}(\boldsymbol{\varepsilon}) := A_t(\boldsymbol{x})A_r(\zeta), \qquad \boldsymbol{\varepsilon} := \begin{bmatrix} \boldsymbol{x} \\ \zeta \end{bmatrix}, \qquad \boldsymbol{x} := \begin{bmatrix} x \\ y \end{bmatrix} \\ {}^{a}B_b := \begin{bmatrix} {}^{a}R_b & -\Omega\,{}^{a}\boldsymbol{p}_b \\ 0_{1\times 2} & 1 \end{bmatrix}, \quad I_{23} = [I_2 \quad 0_{2\times 1}] \in \mathbb{R}^{2\times 3} \\ {}^{a}W_b := \begin{bmatrix} {}^{a}R_b \\ {}^{a}\boldsymbol{p}_b \times {}^{a}R_b \end{bmatrix} = \left[I_{23}\,{}^{b}B_a\right]^T \end{gathered} \tag{36} $$

Authors Introduction

Dr. Takayoshi Yamada

He received the B.E., M.E, Ph.D. degrees in Mechanical Engineering from Nagoya Institute of Technology, Japan, in 1991, 1993, 1995, respectively. He is a Professor with the Dept. of Mechanical Engineering, Gifu University. His research interests include grasping, manipulation, sensing, and automation systems.

Dr. Junya Sato

He received the B.E., M.E., and Ph.D. degrees in Eng. from Iwate University in Japan, in 2012, 2014, and 2017, respectively. He is an Assistant Professor with the Dept. of Mech. Eng., Faculty of Eng., Gifu Univ., Japan. His research interests include applications of evolutionary computations to computer vision.

Dr. Kazuaki Ito

He received the B.S., M.S. and Ph.D. degrees in electrical and computer engineering from Nagoya Institute of Technology, Japan, in 1998, 2000, and 2003, respectively. He is a Professor with the Dept. of Mech. Eng., Gifu University. His current research interests include applications of motion control theory and soft computing techniques for mechatronic systems and real-world haptics.

Dr. Hidehiko Ymamoto

He received the M.E. and Ph. D. degrees in Mech. Eng. from Nagoya Institute of Technology, in 1980, 1991. He was a professor with the Dept. of Mech. Eng., Gifu University. He is currently a specially appointed professor, Gifu University.

My Tally -A Personal Book Keeping Mobile Application

Li Zhihan
Institute of Computer Science and Innovation, UCSI University, Kuala Lumpur, Malaysia

Abdul Samad Shibghatullah
Institute of Computer Science and Innovation, UCSI University, Kuala Lumpur, Malaysia

Nur Hazirah Hamdan
Faculty Administration Science and Policy Studies, Universiti Teknologi MARA,UiTM, Pahang Branch, Raub Campus Malaysia

E-mail: 1001852923@ucsiuniversity.edu.my, abdulsamad@ucsiuniversity.edu.my, nurhazirahhamdan@gmail.com

Abstract

Bookkeeping has a long history and is charming and colorful. Throughout the ages, people have used the experience of abacus bookkeeping; Later, we used pens and books to keep accounts, supplemented by calculation results; In the information age of the 21st century, the popularization of computer software makes us turn to excel or other desktop software with powerful analysis function for computer bookkeeping. However, all the above methods are limited by time and space. For example, they can't record anytime and anywhere, so they will delay the best recording time, or they need to spend a lot of time on account statistical analysis in the later stage. In this way, people's bookkeeping habits cannot continue, nor can they better carry out their own financial analysis. More importantly, they cannot achieve the purpose of financial planning. In modern economic society, people's economic activities are frequent and there are many kinds of consumption. In this case, there is an urgent need to study an application dedicated to daily consumption bookkeeping. The personal bookkeeping application designed in this paper not only reflects the convenience of recording income and expenditure information, but also analyses the financial situation, so that people can record information faster and in real time, and provide users with the ability to record and manage daily financial events.

Keywords: Financial management; Bookkeeping; Android application

1. Introduction

With the continuous development of society and science and technology, people's material living standard and consumption level have also been greatly improved. Especially some young people, they prefer new consumption patterns, but at the same time, such habits also bring some disadvantages. For example, some people's financial chaos has put their lives in trouble. In this case, different bookkeeping tools have emerged to meet people's needs today. So, this is the macro background of choosing this topic. Today, mobile phone has become the most important terminal in people's daily life. In the past, perhaps our mobile phones could only provide the most basic communication functions, but with the birth of Android system, our life has ushered in great changes no less than those brought about by the "industrial revolution" [1]. The ability of mobile phones based on Android system to process all kinds of information has been qualitatively improved, and the open source operating system Android gives everyone who is willing to use it the right to enjoy this high-quality service. It's like a piece of white paper on which we can draw anything we want. It is important to give freedom when creating new things. According to [2], the global mobile application market has reached US $1540.5 billion in 2019, and its compound annual growth rate (CAGR) is expected to be 11.5% from 2020 to 2027. In the past decade, the Internet has become the main medium for communication through a variety of devices. In the past few years, due to the expansion of e-commerce industry and product classification, the number of mobile application buyers has increased sharply. The availability of low-cost data plans from telecom operators reduces the cost of mobile Internet, so it attracts more online users [3], which also accelerates the download of cross platform mobile applications. Among them, developing countries such as India, China and Brazil are the main factors driving the growth of APP market [2]. After the outbreak of coronavirus (covid-19), applications based on social media, games and entertainment have the highest downloads compared with other applications. In addition, there has been strong growth in demand for e-commerce, healthcare and educational applications [4]. This shows that the Android based application market is very considerable.

In addition, Android studio is the official Android application development integrated development environment (IDE) based on IntelliJ idea. It provides many functions that can improve the efficiency of Android application construction, such as: flexible construction system based on gradle; Lint tools that capture performance, ease of use, version compatibility, and other issues; C + + and NDK support; Built in support for Google cloud platform, which can easily integrate Google cloud messaging and App Engine, etc. [5].

Compared with the traditional development software Eclipse, Android studio has many advantages. For example, colors and pictures can be previewed in real time in layout and code; String can be previewed in real time; Multi screen preview and screenshot are provided with equipment frame, which can record simulator video at any time; You can directly open the location of the file; Move, search and jump across projects; Automatically save without Ctrl + s; etc. [6][7].

As you can see, Android applications are very popular in today's social life. In this case, developing a personal bookkeeping application based on Android can take advantage of Android to get the likes of many users, so as to further download and install. At the same time, because the mobile phone is easy to carry, it can greatly improve the value of the software, so that people can not only complete bookkeeping anytime and anywhere, but also better help people improve their habit of unwilling to record and help them achieve effective financial management [8]. The personal bookkeeping application of this project is based on Android platform. It has the advantages of simple design, convenient use, and management function and can help people manage money effectively. Its emergence can help those who can't manage money at present.

2. Systematic review

Based on the comparison of current 5 major bookkeeping applications, Xero, Personal Capital, FreshBooks, QuickBooks and Spendee, we can make the following summary. These five kinds of software only support mobile applications do not support web applications, if you want to experience complete services, customers must download the corresponding software. They all have free trial services. For example, Xero can have a free trial for 31 days after binding its own identity information on the official website. Xero's subsequent payment projects are mainly divided into three payment plans, namely the early plan of $12 per month for self-employed operators and the growth plan of $34 per month for small and medium-sized enterprises, A maturity plan of $65 per month for established enterprises; Personal Capital has free basic services, but other services need to be paid; FreshBooks and QuickBooks also need to bind the customer's relevant information before they can get a free trial time of one month, if the time limit arrives, the customer needs to pay formally before using them; When you decide to use the Spendee (that is, decide to pay the subscription fee), it will give customers an additional week of free trial time, but other services need to be subscribed and paid. According to three different situations, Spendee has prepared three payment types.

As financial management software, they all have secure login or registration pages, and have basic convenience. Whether Xero, personal Capital or FreshBooks provide users with a clean and comfortable interface, making it easy to learn and use, which also lays a foundation for absorbing users. In terms of basic functions, all software have the functions of recording expenditure and income, commenting information, analyzing information, etc., which is the basis for users to get rid of the traditional bookkeeping method. Xero, Personal Capital and

Spendee can bind accounts with banks, which will promote the accuracy of tracking asset flows in one way. However, due to geographical reasons, some bank cards are not easy to add, which may lead to functional redundancy. Some customers also worry about this function because of security considerations.

It can also be seen from comparison that Xero, Personal Capital and QuickBooks lack feedback function, which may lack the motivation for software improvement. At the same time, according to the survey, it can be found that different financial software will focus on different aspects. For example, some software will be more targeted at enterprises, which will complicate their function settings, which is not conducive to people who do not know financial knowledge to better master.

In this research, we found that most bookkeeping applications are not tailored to individual needs. Therefore, designing a suitable personal application will be a good suggestion.

3. Methodology

The system development will use the software development life cycle (SDLC) for design, development and testing. SDLC is a process of software projects within a software organization. It includes detailed plans describing how to develop, maintain, replace, and change or enhance specific software [9][10]. The life cycle defines a way to improve software quality and the overall development process [11]. The SDLC method is mainly divided into six stages that are problem definition and planning; demand analysis; software design; program coding; software testing; and operation and maintenance [12][13].

Problem definition and planning: This stage mainly determines the development goals and feasibility of the software. The project plan will contain the project background, problem statement, goals, research questions, research goals and project plan [14][15].

Demand analysis: When it is determined that the software development is feasible, a detailed analysis of the functions that the software needs to achieve. Requirements will include systematic review, reason, scope and resources [16][17].

Software design: In this stage, the bookkeeping application based on Android is designed based on the results of the demand analysis. The design of functions, UI and database will eventually be implemented on the mobile application. Good software design will lay a good foundation for software programming.

Program coding: This stage is to convert the results of bookkeeping application design into computer executable program code. It is necessary to formulate uniform and standard compilation specifications in program coding.

Software testing: After the bookkeeping application is coded, it must undergo rigorous testing to find and correct the problems in the entire design and coding process of the software.

Operation and maintenance: After the bookkeeping application is completed and running for a period of time, it needs to be thoroughly evaluated and maintained at all times.

4. Results

The page is simple and clean, and the colour is mainly light tone, so that the page can bring affinity to users. When entering the app home page, you can first see the user's login page as shown in Figure 1, and the user can enter the ID and password to enter the home page; if the user has no registration information, you can click the registration button below to enter the registration page a shown in Figure 2. The registration page needs to fill in ID, gender, e-mail, country, telephone number and password. After completing the input as prompted, click the register button to complete the registration and return to the login page. The login interface also has the function of saving the account. If you click tick, you will directly enter the page when logging in the software next time. Users can choose according to their preferences.

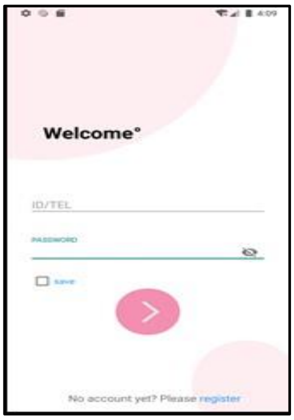
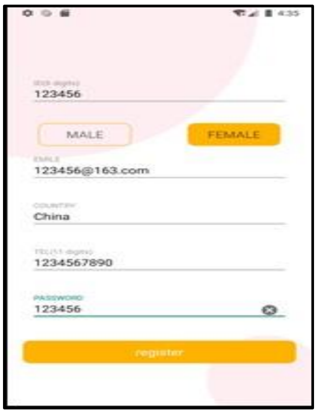


Fig. 1 Login page



123456

EMAIL
123456@163.com

COUNTRY
China

1234567890

PASSWORD
123456

Fig. 2 Registration page

Click the settings page as shown in Figure 3, there will be "log out", "feedback", "about" and "my data" functions. Select the logout button to return to the app login home page. And the button of "my data" will show the information when the user registered at that time. Click “feedback”, and users can put forward their own improvement suggestions. Click “about”, and users can see the information about My Tally application.

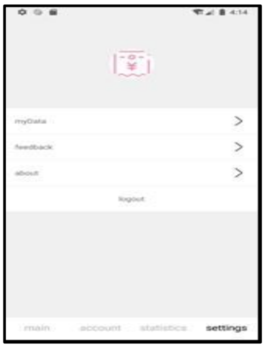

myData
feedback
about
logout
main
account
statistics
settings

Fig. 3 Setting page

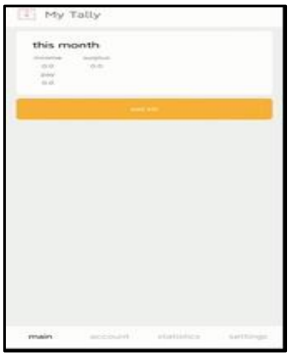
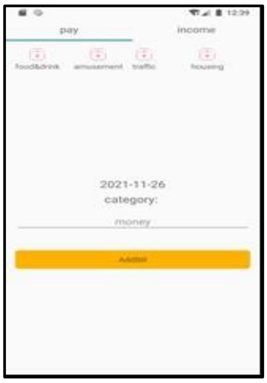
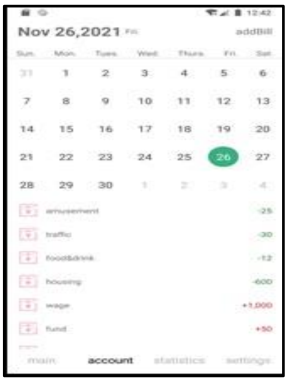

My Tally
this month
income
surplus
pay
add bill
main
account
statistics
settings

Fig. 4 Main page


pay
income
food&drink
amusement
traffic
housing
money
AddBill

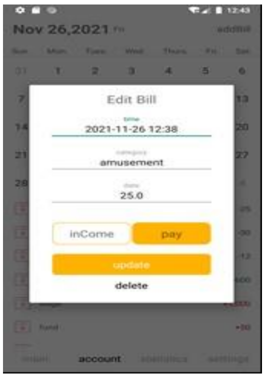
2021-11-26
category:

Fig. 5 Add expenses


Nov 26,2021
addBill
Sun. Mon. Tues. Wed. Thurs. Fri. Sat.
main
account
statistics
settings

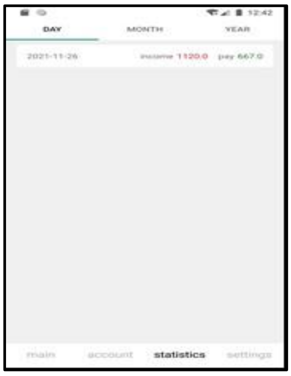
amusement -25
traffic -30
food&drink -12
housing -600
wage +1,000
fund +50

Fig. 6 Account page


Nov 26,2021
addBill
inCome
pay
update
delete
main
account
statistics
settings

Edit Bill
time 2021-11-26 12:38
category amusement
price 25.0

Fig. 7 Edit amount


DAY
MONTH
YEAR
main
account
statistics
settings

2021-11-26 income 1120.0 pay 667.0

Fig. 8 Statistics page

After successful login, the user will first enter the main page as shown in Figure 4. At the bottom of the interface are "main page", "account", "statistics" and "setting". At the top of the main page, you can see the user's overall expenditure, overall income and surplus status of the current month. Moreover, the main page can directly add accounts for bookkeeping as shown in Figure 5. Click the yellow "add bill" to enter the bookkeeping page, which is mainly divided into two categories: expenditure and income. “Record expenditure” and “Record income”, both of them can note the category of account. For example, expenditure includes food and beverage, entertainment, transportation, housing, etc., and income includes wage income, fund income, prize income, etc. Click the account page as shown in Figure 6, and the users can see the detailed revenue and expenditure information of each day according to the query calendar; at the same time, users can also change or delete existing accounts as shown in Figure 7. Click the statistics page as shown in Figure 8, the users can see the overall revenue and expenditure of each day, month or year, which can help users better manage their financial management.

5. Conclusion

"Bookkeeping" is considered to be another powerful entrance to Internet Finance after "payment". In the past, most bookkeeping tools on the market were manual bookkeeping, with cumbersome procedures and user stickiness, so it was difficult to maintain. Based on this situation, this paper designs an bookkeeping application based on Android platform. My Tally focuses on the construction of Android environment, the application of Android related knowledge and the implementation of computer algorithms. This app has functionally designed many aspects involved in people's daily consumption. It is a very suitable bookkeeping type app at present. My Tally combines the characteristics of the mobile client Android platform with the content needed by people in daily consumption, and then plans the solution, which has a strong representative significance in the current industry. After system testing, the personal bookkeeping management app designed in this paper has the characteristics of simple and clear, targeted function design and convenient and rapid use. With the development of technology, this app will constantly update its functions and expand its scope of application to meet the needs of more users.

Acknowledgment

The authors would like to thank Institute of Computer Science & Digital Innovation, UCSI University Kuala Lumpur, Malaysia for helping this project.

References

[1] 1 Techterms.com, "Smartphone", 2021. [Online]. Available: http://www.techterms.com/definition/smartphone [Accessed 15 October 2021].

[2] P. Borasi, and S. Baul, "Mobile Application Market Size & Share Report", 2020. [online] Available: https://www.grandviewresearch.com/industry-analysis/mobile-application-market/segmentation [Accessed 15 October 2021]

[3] Rajendran, S.D. and Wahab, S.N., "Investigating Last Mile Delivery Options on Online Shoppers Experience and Repurchase Intention", International Journal of Electronic Marketing and Retailing, 13(2), 224-241. 2022.

[4] Ismail, S., Nair, R.K., Sham, R., and Wahab, S.N., "Impacts of Online Social Media on Investment Decision in Malaysia", Indian Journal of Public Health Research & Development, 9(11), 241-1246. 2018.

[5] Du, J., "Design of Image Verification Code Based on Android SDK", Computer Science and Application, 10(03), 583-589. 2020.

[6] Malgaonkar, S., Sumeet, S., Radia, Y. and Philip, N., "A Review and Basic Guidelines on Developing Android Applications", International Journal of Computer Applications, 132(3), 42-49. 2015.

[7] Siti Azirah Asmai , Z. Zainal Abidin , Zuraida Abal Abas , Ahmad Fadzli Nizam Abdul Rahman and Muhammad Hafizi Mohd Ali, "Aedes Mosquito Larvae Recognition with A Mobile App" International Journal of Advanced Trends in Computer Science and Engineering 9(4):5059-5065, 2020

[8] Wahab, S.N., "Transforming Malaysian On-Demand Delivery Apps for Efficient Last-Mile Delivery: A SWOT Analysis", 2021 International Conference on Data Analytics for Business and Industry (ICDABI), 396-399, 25-26 October, Sakheer, Bahrain. doi:10.1109/ICDABI53623.2021.9655884. 2021.

[9] Sivananda, Dharshen, Siti Azirah Asmai, Zaheera Zainal Abidin, Zuraida Abal Abas, and Sabrina Ahmad. "Sentiment Analysis Dashboard Using Twitter Data in Power BI." Manuscript Editor 2021,61. 2021

[10] Sham., R., Wahab, S.N., and Hussin, A.A.A., "Smart Trolley Apps: A Solution to Reduce Picking Error", International Journal of Supply Chain Management, 7(5), 294-302. 2018.

[11] Ahmad, Sabrina, and Siti Azirah Asmai. "Measuring software requirements quality following negotiation through empirical study." International Journal of Applied Engineering Research 11.6 (2016): 4190-4196.

[12] Bahar, N., Wahab, S.N. and Ahmad, N.D., "Understanding Challenges Faced in Online Teaching and Learning Among Malaysian Universities' Instructors During COVID-19 Pandemic", Proceedings of the International Conference on e-Learning, ICEL, Vol. 2020-December, pp. 154-157, 6th International Conference on e-Learning, ICEL 2020, 6-7 December 2020, Sakheer, Bahrain. doi:10.1109/econf51404.2020.9385474. 2020.

[13] Asmai, S., Zukhairin, M.N.D.M., Jaya, A., Rahman, A.F.N.A. and Abas, Z., "Mosquito larvae detection using deep learning". International Journal of Innovative Technology and Exploring Engineering (IJITEE), 8(12),.804-809.2019

[14] Keoy Kay Hooi, Koh Yung Jing, Chit Su Mon, Hafeez Khalid, Lee Luqman, Thong Chee Ling (2022). "Impact of Assisted Technological Enablement on Entrepreneurial Success among Malaysian Higher Education Students". 1st International Conference on AI in Cybersecurity (ICAIC) 24 May-26 May 2022. University of Houston-Victoria, USA.

[15] Keoy Kay Hooi, Koh Yung Jing, Genaro Japos, Chit Su Mon, Lee Luqman, Ho Meng Chuan. (2022). "Examining Model of Entrepreneurial Success Factors Among Undergraduate Students: A Quantitative Study". The International Academic Symposium of Social Science 2022 (IASSC 2022) 3rd of July 2022 from Kelantan, Malaysia.

[16] Sulaiman, Rabatul Aduni, Dayang Norhayati Abang Jawawi, and Shahliza Abdul Halim. "Classification Trends Taxonomy of Model-based Testing for Software Product Line: A Systematic Literature Review." KSII Transactions on Internet and Information Systems (TIIS) 16.5 (2022): 1561-1583.

[17] Sulaiman, Rabatul Aduni, Dayang NA Jawawi, and Shahliza Abdul Halim. "Cost-effective test case generation with the hyper-heuristic for software product line testing." Advances in Engineering Software 175 (2023): 103335.

Authors Introduction

Li Zhihan

She completed her Bachelor of Science in Business Information System Computing at UCSI University, Cheras, Kuala Lumpur on 20th August 2022. Now she is preparing to continue her Master degree. Her research interest are in Information System, Business and Finance.

Abdul Samad 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 Associate Professor at Institute of Computer Science & Digital Innovation, UCSI University, Kuala Lumpur, Malaysia. His current research interests include optimization, modelling and scheduling

Nur Hazirah Hamdan

She received her bachelor degree in finance from Universiti Teknologi MARA, Melaka, Malaysia in 2011, Master Business Administration from Universiti Teknologi MARA, Shah Alam, Malaysia in 2014, and Ph.D. holder in Islamic Banking and Finance from International Islamic University Malaysia in 2022. She is currently Senior Lecturer at Universiti Teknologi MARA, Pahang Branch, Raub Campus, Malaysia. Her research interests include Islamic finance and banking, corporate finance, entreprenuership and technology acceptance and adoption.

A Survey on Charitable Acts, Challenges and Using Charitable Mobile Application

Nan Pepin
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

Nur Hazirah Hamdan
Faculty Administration Science and Policy Studies,
Universiti Teknologi MARA,UiTM, Pahang Branch, Raub Campus, Malaysia

E-mail: 1001953764@ucsiuniversity.edu.my, abdulsamad@ucsiuniversity.edu.my,
kasthurisuba@ucsiuniversity.edu.my, nurhazirahhamdan@gmail.com

Abstract

A donation is a gift that is given voluntarily or sincerely without the expectation of receiving something in return. Donating in contemporary terminology is better known by the term philanthropy. Today, the emergence of charity platforms resulted in a modern society increasingly making donations in cyberspace, which are referred to as online charity. Although the use of digital technology allows for philanthropic activities to be carried out, there are still shortcomings in this online system. This paper aims to study mobile charitable applications among users. The survey has been conducted via Goggle form and distributed to 110 participants through the internet. The findings show that there are not many applications of this sort being produced by other developers and there is a need to develop a mobile application that can help donators and receivers. The proposed mobile applications should be faster, easy to manage and transparent.

Keywords: Charity, Application, Donation, Survey

1. Introduction

The increasing use of the Internet and easy access has also made the Internet the main choice in the daily activities of Malaysians, including in charitable activities. According to data reports released by from Department of Statistics Malaysia (DOSM), an individual who use the internet climbed by 5.4 percentage points from 84.2 percent in 2019 to 89.6 percent in 2020. This growth is due to daily activities such as working, studying, and buying to be done digitally. Using the internet as a means of ethical funding is a viable option. Currently, all network access between two people is connected in a single network. This is because the Internet is the most effective way to connect all of its users without regard to distance or time [1]. Because of the wide usage and accessibility of the Internet, it has become the preferred method of conducting charitable activities.

1.1 *Charity*

As human beings, being socially frequently encountering obstacles in the process of their existence, resulting in human beings never being completely independent of the additional support of other humans. In turn, it stimulates the impulse to assist one another and assist one another since, at their core, human beings are willing to cooperate in order to achieve happiness in their life. Humans who live in communities are classified into groups based on the strata of society and rank. Using a system of classification that ranges from the amount of education to money, humans are divided into two groups, the wealthy and the less wealthy. However, the natural urge of human beings to help one another and to help one another has always led people who seem to be self-sufficient to be willing and able to help one another, one of which is through charitable contributions [2]

Contributions to charity are also linked to the development of the technological revolution, which has resulted in a modern society increasingly making donations in cyberspace, which are referred to as online charity. According to study [3] [4] [5] one of the factors of charitable giving in the context of online donations is merely the act of seeking monetary donations. There are many different sorts of online philanthropy platforms accessible, including websites, social media, and mobile applications for smartphones [6]. Although the use of digital technology allows for philanthropic activities to be carried out, there are still shortcomings in this online system [7] [8] [9] and unclear incentives is one of the shortcomings in online charity. The benefits of making a donation are unclear, whether from the fundraiser's or the donor's perspective. Fundraisers must use the fundraising platform to raise donations by publishing information about the campaign and the quantity amount of money they require. It will take quite a long time for a fundraising campaign to succeed before users can withdraw their funds, especially on crowdfunding platforms. In addition, the dependence on monetary means that donors have no choice but to give donations in the form of money. Most current platforms only focus on money fundraising. This limitation causes the donor or receiver to be hindered because not all parties need help or to give contributions other than money. Some people need help such as daily necessities, time or manpower. Scholars used the disparities between time and money donations to illustrate that time donations require more human and social resources than money gifts [6] [10].

This paper intends to study on user charitable acts, challenges and using mobile charitable application. Result from this study will be used to the following processes to be carried out for next stage in developing for the seekaandhelp application.

2. Methodology

The survey will be carried out in the form of an online questionnaire for data collection because it has a low return rate and people are more willing to reply to questions [11][12][13].

2.1 *Survey question types*

Single response questions, multiple-choice questions, Likert scale questions, and open-ended survey questions will all be used in this questionnaire. A single answer question is a survey question that uses the radio button format of the circular button that represents the alternatives in the list to allow respondents to choose only one answer from a list of options. In contrast to single-answer questions, multiple-choice questions are often denoted by a rectangular checkbox. Respondents were given the option of checking all of the alternative answers supplied. Likert scale questions are surveys that use a selection of answers ranging from strongly disagree to strongly agree to measure respondents' opinions and feelings. A Likert scale was used to assess respondents' opinions or attitudes toward the topics under consideration. Furthermore, the other type of survey question that will be used in this survey question is an open-ended survey question. Respondents will be required to enter their responses into a textbox for this sort of survey question.

2.2 *Survey Tools*

The Google form platform will be used to distribute questionnaires to respondents. Google Forms is the preferred method of disseminating this survey due to its online nature and ability to reach participants in urban and rural areas. The survey will be given to persons living in urban or suburban areas for this study. This questionnaire is distributed at random via internet platforms such as CN, social media, WhatsApp, Facebook, and others. This questionnaire survey intends to collect replies from at least 100 people [14][15][16].

Section 1 is a questionnaire regarding the respondents' demographics. Four questions will be asked in this section: gender, age group, level of education, and employment status. Demographic questions are an important part of every survey. Demographic questions are intended to aid in the identification and quantification of characteristics that may influence responders.

Section 2 of the questionnaire will inquire about the respondent's involvement in charitable activities. Participants will be asked 11 questions in this section to determine if they are donors or receivers of donations. The inquiries are on the regularity of their charitable activities, how, where, what items are donated or requested, what problems they have faced, and etc.

The third section of the questionnaire will inquire about participants' experiences with the charity mobile app. In this section, participants will be asked five questions about whether they have ever used the charity app on their smartphone and what they thought of their experience with it. Two of the questions in this part are short answer questions, while the other three are Scala-like answers [17][18][19][20].

3. Result and Discussion

After the survey has been completed and the responds from the 110 respondents have been collected, the data will be analyzed based on the responses received from the respondents.

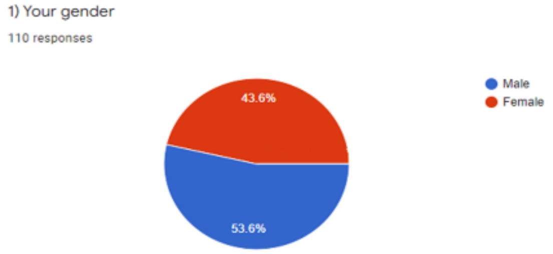


Fig.1.

Figure 1 illustrates a pie chart indicating the number of respondents based on the gender of those who responded to the survey questions. The result shown male respondents represented 53.6 per cent of all respondents. Moreover, a total of 43.6 percent of the respondents indicated that they were female, as a result, on average, out of 110 respondents gathered, 61 were male and 48 were female has participate in this survey.

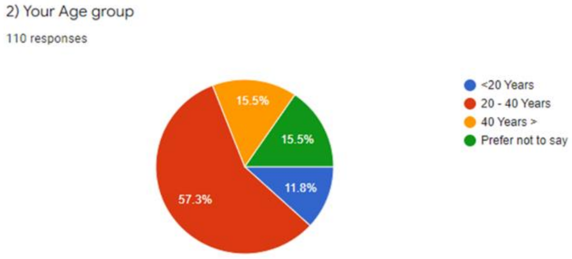


Fig.2.

Figure 2 illustrates the overall age group of those who answered the survey questions. The results of pie chart show that a total of 57.3 percent, or a total of 63 respondents, are between the ages of 20 and 40 years. Thereafter, the next 15.5 per cent, or a total of 17 responders, comes from the group of people between the ages of 40 and above, as well as from the group of people who prefer not to share their age. While 11.8 percent, or a total of 13 respondents, are under the age of 20, the majority of respondents are beyond the age of 20.

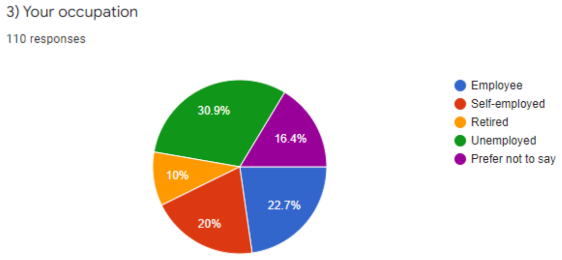


Fig.3.

Result of respondents are visualized in Figure 3 there are 34 respondents who are unemployed, as indicated by the percentage of 30.9. The number of employed respondents is 22.7 percent, representing a total of 25 respondents, while the number of self-employed respondents is 20 percent, representing 22 respondents, and the number of respondents who choose not to prefer their occupation is 16.4 percent, representing a total of 18 respondents while the number of retired respondents is 10 percent, representing 11 respondents. It can be concluded that the vast majority of those who took part in the survey are unemployed. The survey that was done focusing on charitable in general, and the participants in this study came from a variety of different employment backgrounds as well. As a result, measures about employment are generally asked to determine the level of balance or unemployment among those who take part in this survey.

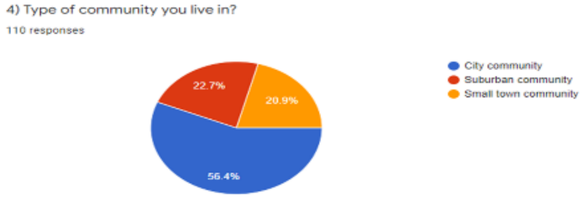


Fig.4.

© The 2023 International Conference on Artificial Life and Robotics (ICAROB2023), Feb. 9 to 12, on line, Oita, Japan

Figure 4 show the results of the survey on the type of community in which respondents live. The pie chart illustrates that the biggest percentage, 56.4%, comes from a total of 62 respondents who indicated they live in urban communities. The next highest percentage, 22.7 percent, represents a total of 25 respondents who live in suburban communities, while the lowest percentage, 20.9 percent, represents a total of 23 respondents who reside in small-town communities. In short, respondents from urban areas are more likely to take part in this survey than those from sub-urban and small-town areas.

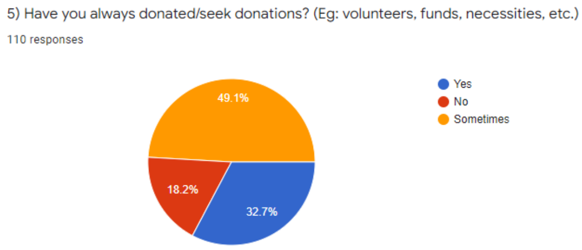


Fig.5.

Figure 5 presents the total number of 110 respondents who have donated and who are looking for donations. The results of the pie chart illustrate that the highest number of respondents which is 49.1 per cent, or 54 of the total respondents stated that they sometimes do or seek donations. Followed by 32.7 per cent representing a total of 36 respondents who always donate or seek donations, and a further 18.2 per cent representing a total of 20 respondents who have never donated or requested donations. According to the results of this survey, the average number of respondents who make or seek donations is higher than the average number of respondents who have never made or sought donations. This indicates that charitable giving and seeking donations are activities that also respondents generally participate in it.

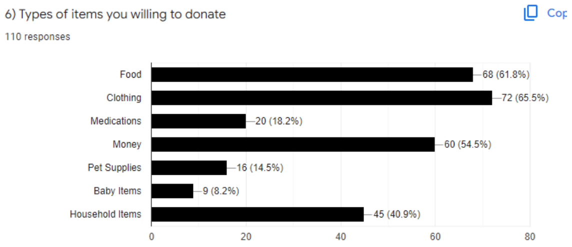


Fig.6.

Figure 6 illustrates a bar chart indicating the number of different types of items that respondents are willing to donate in relation to the total number of items. The bar chart shows that clothing is the most common type of item to be donated by respondents, amounting to 65.5 per cent of all items donated or 72 respondents in total. Followed by 68.1 per cent of respondents who indicated a willingness to donate food items, totaling 68 respondents. The third category includes up to 54 per cent of 60 total respondents who are willing to offer items in the form of money. The fourth item that people are willing to donate is a household item, which scored 40.9 per cent of the survey, reflecting a total of 45 respondents. Medication ranked fifth, with 18.2 per cent of the total of 20 respondents. Followed by pet supplies at 14.5 per cent, representing a total of 16 respondents, and baby items at 8.2 per cent, indicating a total of 9 respondents willing to donate their items. According to the data, the most common types of commodities given to persons in need were clothing, food, and money. This also indicates that respondents were involved in giving charity in accordance with the types of items they were able to provide.

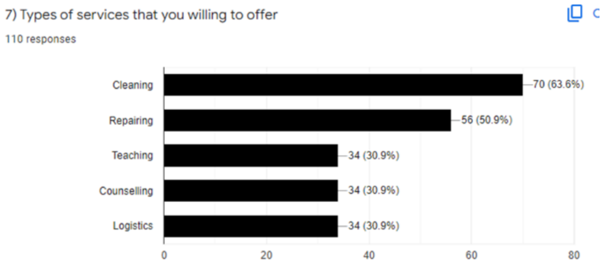


Fig.7.

Figure 7 illustrates the results of the survey regarding the types of services that respondents would like to provide to those in need. The bar chart depicts a total of 63.6 per cent, or a total of 70 respondents, who are willing to provide cleaning services. Following that, service repair at 50.9 per cent represented a total of 56 respondents, and teaching, counselling, and logistics at 30.9 percent each represented a total of 34 respondents. The result shows the type of services offered by respondents in the form of cleaning and repairing is the most common of the services which are provided. This also implies that respondents offer services in accordance with the types of services they can perform.

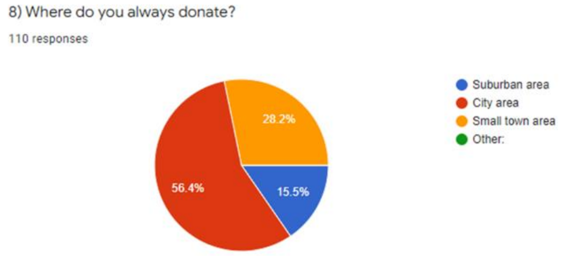


Fig.8.

Figure 8 show a survey of the respondents donated preferred location. The pie chart illustrates a result of 56.4 per cent, which indicates a total of 62 respondents who made charitable contributions in the city. Following that, 28.2 per cent of the 31 responses were by respondents who donated in small-town areas, while 15.5 per cent, or 17 responses, were by respondents who donated in suburban areas. Result show that out of 110 respondents, respondents who donated in the city area were the most frequent when compared to respondents who donated in the city area and suburban area. This indicates that the majority of respondents to this survey donated in accordance with their preferred donation location.

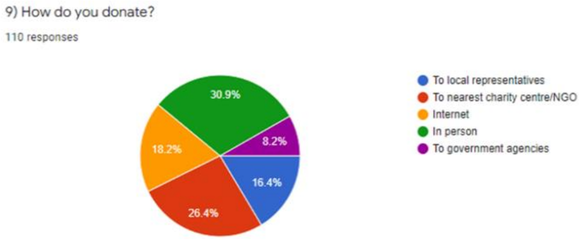


Fig.9.

Figure 9 show the survey results of participants about their charitable donations practices. The pie chart illustrates an overall percentage of 30.9 per cent, which represents a total of 34 respondents who give in person to make their donation. Representing 26.4 per cent a total of 27 respondents donated to a local charity centre and also non-governmental organization (NGO). It was discovered that 20 respondents which made donations through the internet, accounted for 18.2 per cent of the result. Those respondents who reported that they made donations to their local representatives or government agencies accounted for 16.4 per cent of the total, and those who stated that they made donations to their local representatives or government agencies made up 8.2 per cent. According to the data collected from this survey, the most common approaches of donating are in-person, through a local charity center, and also through the internet. The findings of this study indicate how respondents preferred approach when making a donation

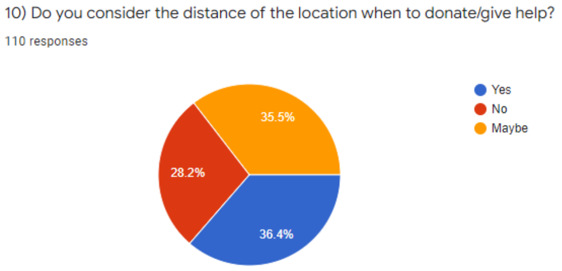


Fig.10.

Figure 10 shows the survey of respondents regarding the distance and location of their donations. On the pie chart, the data result of 36.6 per cent, or a total of 40 respondents, is from respondents who agreed to take into consideration the distance and location of donors while making charity donations. While a total of 35.5 per cent, or a total of 37 respondents, indicated that it is likely to take into account the distance and location when making a donation. Account at 28.2 per cent which represents a total of 30 respondents, indicated that they do not consider distance or location when making a donation. In summary, out of the 110 data points obtained, the majority of respondents stated that their distance and location should be taken into consideration. This survey is to evaluate respondents according to their willingness to make donations regardless of the distance and location.

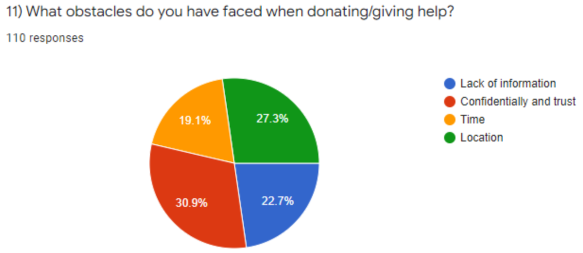


Fig.11.

Figure 11 show survey of what respondents' encounters with donation obstacles were presented. According from the pie chart, a total of 30.9 per cent, or a total of 34 respondents, answered that the most challenging part of donating is preserving confidentiality and trust in the charity. A total of 30 respondents stated that location was a challenge when making a donation which resulted in 27.3 per cent of the total number of respondents. Further, 22.7 per cent, or a total of 24 respondents, stated that a lack of information is their major challenge while 19.1 per cent, or a total of 21 respondents, answer that a lack of time is their main obstacle in making donations. To conclude, based on 110 responses, the most common challenges faced by respondents, are a lack of confidence and trust, a lack of location, and a lack of information. Apart from determining what motivates respondents to donate, it is necessary to consider the obstacles they may experience.

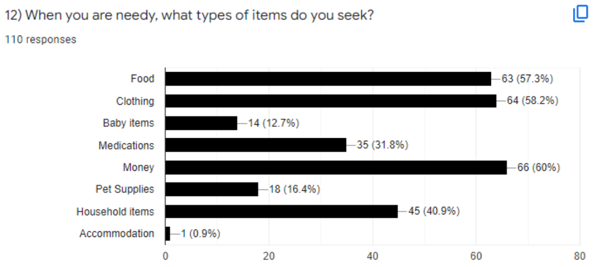


Fig.12.

Figure 12 shows the results of a survey conducted with the respondent regarding the respondents needy. According to the results of the bar chart, 60 per cent of respondents, or 66 in total, indicated that money is a necessity that they seek when in need. Clothing is the most sought-after item, according to a total of 58.2 per cent of respondents representing a total of 66 respondents, followed by food, which received 57.3 per cent of responses representing a total of 63 respondents. Furthermore, 40.9 per cent indicated the need for household items, representing a total of 45 respondents, while 31.8 per cent said they need medicine items, representing a total of 35 respondents. While pet supplies accounted for 16.4 per cent of 18 respondents and baby supplies accounted for 12.7 per cent of 14 respondents.

When it comes to searching for types of items when are in need, according to the data collected, respondents most frequently look for money, clothes, and food. This suggests that the respondents to this survey sought help based on the type of their preferred item that met their needs.

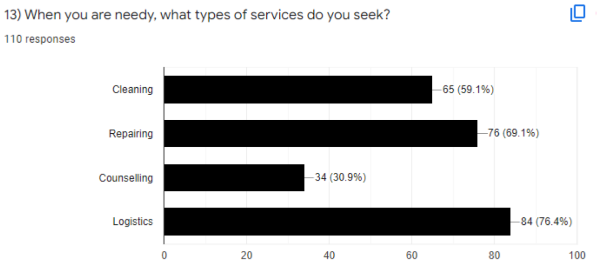


Fig.13.

Figure 13 shows the results of a survey conducted among respondents to identify the services that they would require if they were in need. According to the bar chart, 76.4 per cent of the respondents, or a total of 86 respondents, indicated that they were looking for logistical services. Total of 69.1 per cent, which represented 76 respondents are seeking repair services. Furthermore, cleaning and counselling services were sought after, with 59.1 per cent and 30.9 per cent, respectively, representing a total of 65 and 34 respondents.

According to the data collected in summary, the types of services such as logistics, repair, and cleaning are the most frequently sought by respondents when they are in need. This implies that the respondents who took part in this survey looking for service support based on the type of preferred services that met their needs when they were in need.

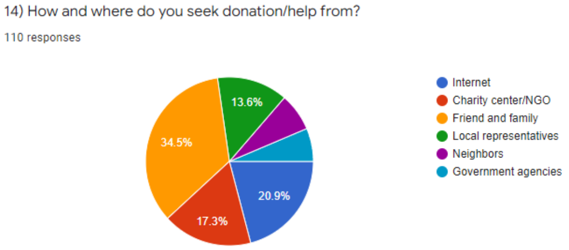


Fig.14.

Figure 14 show respondents were asked where they looked for donations when they were in need. It was discovered that 34.5 per cent of respondents, or a total of 38 respondents, stated that they seek help from their friends and family. There were a total of 23 respondents who seek help through the internet, resulting in 20.19 per cent of respondents. Furthermore, 17.3 per cent of a total of 19 respondents, seek help from nearby charitable organizations or non-governmental organizations (NGOs). On the other hand, 13.6 per cent, representing a total of 7 respondents who looked for help through local representatives, and 6.4 per cent represented a total of 5 respondents who looked for help from government agencies.

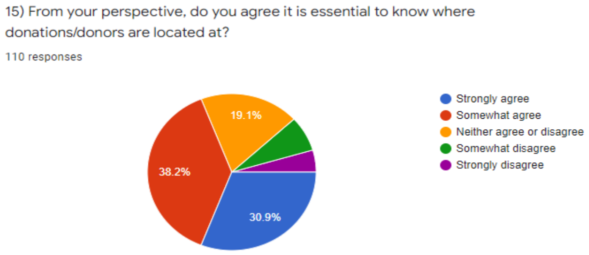


Fig.15.

The results of a survey from respondents' perspectives on the location of the donor or the donation are visualized in Figure 15. Based on the result, a total of 38.2 per cent, which represented 42 respondents who are somewhat in agree. Followed by 30.9 per cent, representing a total of 34 respondents who strongly agree. A total of 19.1 per cent, or 21 respondents, indicated that they were neither agreeing nor disagreeing. While a total of 7.3 per cent representing 8 respondents indicated somewhat disagree and lastly a total of 4.5 per cent representing 5 respondents indicated that they strongly disagreed to know the location of donations and donors.

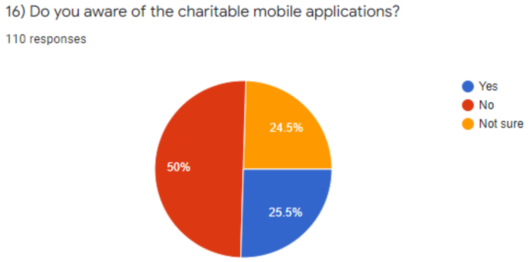


Fig.16.

The results on awareness in charitable mobile application from respondents are visualized in Figure 16. From the illustrate pie chart, a total of 50 per cent of respondents indicated that they were not aware of the donation application, which amounted to 55 respondents of the total. Furthermore, 25.5 per cent of respondents representing a total of 28 respondents indicated that they are aware of charitable applications, while 24.5 per cent of respondents representing a total of 27 respondents stated that they are not sure to be aware of charitable mobile applications.

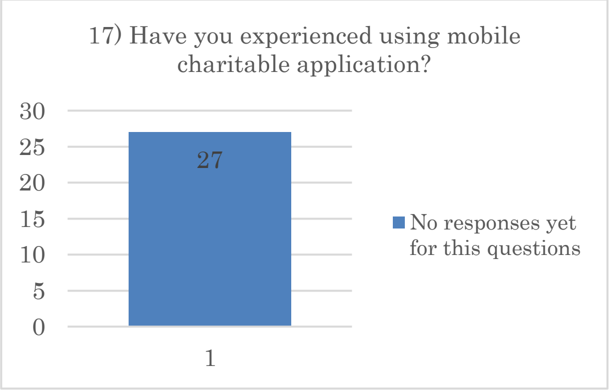


Fig.17.

Figure 17 shows the results of a survey of participants on their experiences with charity applications. This question is a type of short answer which need participate to fill up before submitting it. The result for this data collected, none of the 27 respondents who responded had previous experience with charity applications. This indicates that the majority of respondents have never used the charity's mobile application.

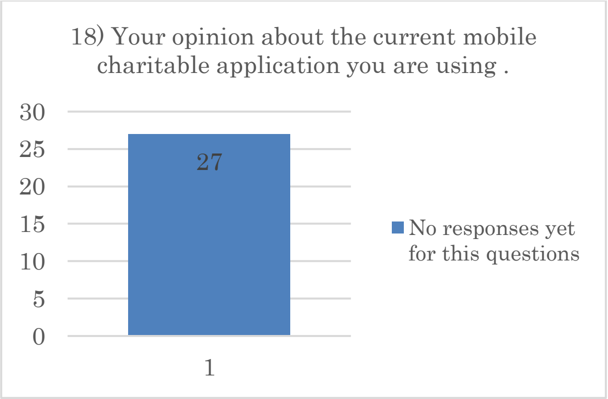


Fig.18.

Figure 18 is a short answer question type. The question in this survey to collect of respondents' opinions about their experience using the charitable application existing. Based on the results obtained, no answers were given by the respondents who had experience using the donation application. Therefore, this indicates that most respondents have never used charitable mobile applications on their smartphones.

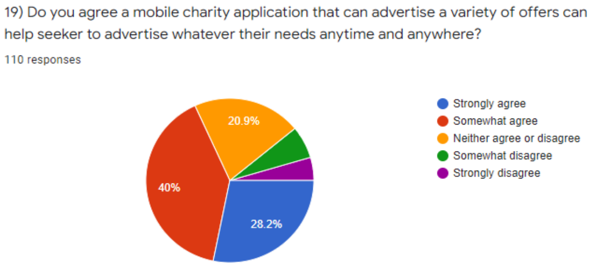


Fig.19.

Figure 19 illustrates participants' responses to a survey on whether applications that advertise various types of offers can help them in addressing their needs. The pie chart shows that up to 40 percent of respondents somewhat agreed, which represents 44 respondents. Following with, 28.2 per cent represent 31 of responders indicated that they strongly agree. 20.9 per cent of respondents of 23 people, indicated that they were neither agreeing nor disagreeing. While 6.4 per cent, or 7 respondents, indicated they somewhat disagree, and lastly 4.5 per cent or 5 respondents, stated they strongly disagree.

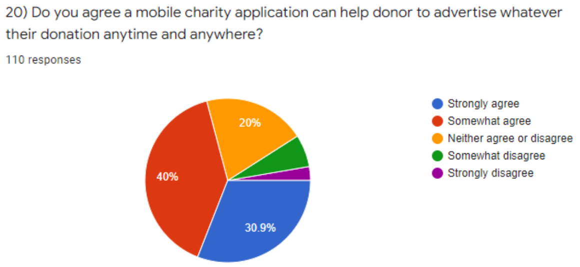


Fig.20

Figure 20 illustrates the results of a survey of respondents who were asked whether a charity application that can advertise any donation will be able to help donors to advertise their donations anywhere they are located. The pie chart shows that up to 40 per cent of respondents somewhat agreed which equals to 44 respondents. Following by 28.2 per cent of responders 31 in total said that they strongly agree. The result of 20.9 per cent represents 23 of the respondents who stated that they were neither agreeing nor disagreeing. While 6.4 per cent or seven respondents claimed they somewhat disagree and 4.5 per cent or five respondents stated they strongly disagree. In summary, the average number of respondents who agree is greater than the average number of respondents who disagree. This indicates how the availability of mobile charity applications able of advertising any donation allows donors to advertise their donations practically anywhere.

5. Conclusion and future work

The results reported in this paper are limited by a number of factors. There is a need to conduct a more detailed analysis of decisions and crucial analysis features. Further study is predicted to be able to pay attention to other factors or causes of interest in using mobile charity application donating. Using more than one independent variable is expected to be able to give better results than the outcomes of this investigation. Moreover, constraints both from the sources of the journal and the theory utilized.

Acknowledgment

The authors would like to thank Institute of Computer Science & Digital Innovation, UCSI University Kuala Lumpur, Malaysia for helping this research.

References

1. B. M. Leiner *et al.*, “A brief history of the internet,” *ACM SIGCOMM Comput. Commun. Rev.*, vol. 39, no. 5, pp. 22–31, 2009, doi: 10.1145/1629607.1629613.
2. S.N. Wahab, Y.M. Loo, and C.S. Say, “Antecedents of Blockchain Technology Application among Malaysian Warehouse Industry”, *International Journal of Logistics Systems and Management*, 37(3), 427-444. 2020.
3. N. Bahar, S.N. Wahab, N.D. Ahmad, “Understanding challenges faced in online teaching and learning among Malaysian universities’ instructors during COVID-19 pandemic”, Proceedings of the International Conference on e-Learning, ICEL, pp. 154-157, 6th International Conference on e-Learning, ICEL 2020, 6-7 December 2020, Sakheer, Bahrain. doi:10.1109/econf51404.2020.9385474. 2020.
4. A. Kumar and S. Chakrabarti, “Charity Donor Behavior: A Systematic Literature Review and Research Agenda,” *J. Nonprofit Public Sect. Mark.*, vol. 00, no. 00, pp. 1–46, 2021, doi: 10.1080/10495142.2021.1905134.
5. S. A. Bin-Nashwan and M. Al-Daihani, “Fundraising campaigns via social media platforms for mitigating the impacts of the COVID-19 epidemic,” *J. Islam. Mark.*, vol. 12, no. 3, pp. 576–597, 2020, doi: 10.1108/JIMA-07-2020-0200.
6. Y. Li and L. Yu, “Factors influencing social media users’ continued intent to donate,” *Sustain.*, vol. 12, no. 3, 2020, doi: 10.3390/su12030879.
7. S. Sura, J. Ahn, and O. Lee, “Factors influencing intention to donate via social network site (SNS): From Asian’s perspective,” *Telemat. Informatics*, vol. 34, no. 1, pp. 164–176, 2017, doi: 10.1016/j.tele.2016.04.007.
8. Y. Chen, R. Dai, J. Yao, and Y. Li, “Donate time or money? The determinants of donation intention in online

crowdfunding," *Sustain.*, vol. 11, no. 16, 2019, doi: 10.3390/su11164269.

9. Lely Priska D. Tampubolon, "Contribution of the Social Media and Transparency Needed on Crowd funding Website," *Int. J. Progress. Sci. Technol.*, vol. 9, no. 1, pp. 91–100, 2018.
10. A. Tomczak and A. Brem, "A conceptualized investment model of crowdfunding," *Ventur. Cap.*, vol. 15, no. 4, pp. 335–359, 2013, doi: 10.1080/13691066.2013.847614.
11. J. Hemer, "A snapshot on crowdfunding," *Enconstor*, p. 39, 2011.
12. S. Ismail, R.K. Nair, R. Sham, and S.N. Wahab, "Impacts of Online Social Media on Investment Decision in Malaysia", *Indian Journal of Public Health Research & Development*, 9(11), 1241-1246. 2018.
13. Bryant, W.K., Jeon-Slaughter, H., Kang, H. et al. Participation in Philanthropic Activities: Donating Money and Time. Journal of Consumer Policy 26, 43–73 (2003). https://doi.org/10.1023/A:1022626529603
14. N. Bahar, S.N. Wahab, and M. Rahman, "Impact of knowledge management capability on supply chain management and organizational practices in logistics industry", VINE Journal of Information and Knowledge Management Systems, 51(5), 677-692. 2021.
15. S.N. Wahab, M.I. Hamzah, N.M. Sayuti, W.W. Lee, and S.Y. Tan, "Big Data Analytics Adoption: An Empirical Study in Malaysia Warehousing Sector", *International Journal of Logistics Systems and Management*, 40(1), 121-144. 2021.
16. Keoy, K. H., Thong, C. L., Cherukuri, A. K., Koh, Y. J., Chit, S. M., Lee, L., Genaro, J. & Kwek, C. L. (2022). An Investigation on the Impact of Technological Enablement on the Success of Entrepreneurial Adoption Among Higher Education Students: A Comparative Study. Journal of Information & Knowledge Management, 21(2), 23.
17. K. Hafeez; K. H. A. Keoy; M. Zairi; R. Hanneman; S.C. Lenny Koh (2010) E-supply chain operational and behavioural perspectives: an empirical study of Malaysian SMEs International Journal of Production Research, 1366-588X, Volume 48, Issue 2, Pages 525 546.
18. Hafeez, K., Keoy, K.H. and R. Hanneman (2006), "E-business capabilities model: Validation and comparison between adopter and non-adopter of e-business companies in UK", Journal of Manufacturing Technology Management. A Special Issue on E technology and Manufacturing Enterprise Competitiveness. ISBN: 101-84663-080-013. Vol. 17, No. 6, pp. 806-828.
19. Sulaiman, R. Aduni, Dayang NA Jawawi, and Shahliza Abdul Halim. "A Dissimilarity with Dice-Jaro-Winkler Test Case Prioritization Approach for Model-Based Testing in Software Product Line." KSII Transactions on Internet and Information Systems (TIIS) 15.3 (2021): 932-951.
20. Aduni Sulaiman, R., Dayang NA Jawawi, and Shahliza Abd Halim. "Derivation of test cases for model-based testing of software product line with hybrid heuristic approach." International Conference of Reliable Information and Communication Technology. Springer, Cham, 2019.

Authors Introduction

Nan Pepin

She completed her Bachelor of Science in Business Information System at UCSI University, Cheras, Kuala Lumpur on 31st December 2022. Now she is preparing to continue her Master degree. Her research interest are in Information System, Business and Finance.

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 Associate Professor at Institute of Computer Science & Digital Innovation, UCSI University, Kuala Lumpur, Malaysia. His current research interests include optimization, modelling and scheduling

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. She obtained her doctoral degree from Malaysia University of Science & Technology. She has supervised many undergraduate students as main supervisors and co-supervisors. She has publications in Scopus-Indexed Journals and Web of Science. She was also a co-researcher in Pioneer Science Incentive Fund (PSIF) in the area of augmented reality. Her research interests include human-computer interaction, human personality types, augmented reality, e-learning, mobile commerce and e-commerce.

Nur Hazirah Binti Hamdan

She received her bachelor degree in finance from Universiti Teknologi MARA, Melaka, Malaysia in 2011, Master Business Administration from Universiti Teknologi MARA, Shah Alam, Malaysia in 2014, and Ph.D. holder in Islamic Banking and Finance from International Islamic University Malaysia in 2022. She is currently Senior Lecturer at Universiti Teknologi MARA, Pahang Branch, Raub Campus, Malaysia. Her research interests include Islamic finance and banking, corporate finance, entreprenuership and technology acceptance and adoption.

Online Parcel Management System (PMS) for Small and Medium Company

Chew Cheng Jin
Institute of Computer Science and Innovation, UCSI University, Kuala Lumpur, Malaysia

Kasthuri Subaramaniam
Institute of Computer Science and Innovation, UCSI University, Kuala Lumpur, Malaysia

Abdul Samad Shibghatullah
Institute of Computer Science and Innovation, UCSI University, Kuala Lumpur, Malaysia

Salini Devi Rajendran
School of Food Studies & Gastronomy, Faculty of Social Sciences & Leisure Management, Taylor's University Lakeside Campus,Malaysia

Nur Hazirah Hamdan
Faculty Administration Science and Policy Studies, Universiti Teknologi MARA,UiTM, Pahang Branch, Raub Campus Malaysia

E-mail: 1001540205@ucsiuniversity.edu.my, kasthurisuba@ucsiuniversity.edu.my, abdulsamad@ucsiuniversity.edu.my, salini.rajendran@taylors.edu.my, nurhazirahhamdan@gmail.com

Abstract

The rising trend of e-commerce has caused the traditional shopping slowly being replaced by online shopping. Nowadays, online shopping has been very popular and slowly being accepted by. Due to the demand of the people in the field of online shopping increasing, which lead to the growing of the number of parcel day by day. The proposed system is to solve all the limitation of the current parcel management system and to improve the performance of the current system which will reduce the workload of the staff and to let the recipient of the parcel using the lesser time to collect their parcels. This research is going to gather more information of the recipient viewpoint on how Parcel Management System can help current system to reduce the workload and manpower to handle all these parcels, increase efficiency and save time and effort which make the staff of the reception counter easier to work and no need to use back the old method to record the details of the parcel.

Keywords: Parcel management system, Online shopping, E-commerce, Courier services.

1. Introduction

In this new era, delivery services are assuming a significant duty in community because of the rising trend of the E-commerce. The increasing trend of E-commerce prompt to a changing consumer purchasing behavior to online shopping stores instead of physical outlets thus increase the income of goods and products delivery by courier companies. As e-commerce division is growing faster in Malaysia, the last mile delivery is an important aspect to provide online shoppers in Malaysia a very good online shopping experience. The importance of last mile deliveries to customers has increased. The last mile issues occur during the arrangement of distribution service from a transportation hub to the final delivery destination of the consumer's house or workplace [1][2]. However, there are issues occurring in this service, and these issues significantly result in customer dissatisfaction and low efficiency. Convenience and time saving were among the most complained by online shoppers. This has been highlighted in recent study by Morganosky and Cude [5] that online shoppers cannot achieve convenience and time saving advantages of online shopping. Furthermore, due to delays in delivery or the issues of failed delivery, some online shoppers even feel online shopping takes longer than traditional shopping. An average of 38% of home customers were required to collect missed delivered things from a post office or other warehouse in the year 2016 [3][4][5]. Added by Charlton (2018), 63% concerned on delivery speed when shopping online. According to statistics by Statista on year 2017, customers demand delivery flexibility the most from their last mile delivery services, which occupies 65%. Following by speed of delivery (61%), real time visibility (51%), delivery options (45%), and specific delivery slot selection (41%). With 67% of Malaysians online, making Malaysia the most astounding penetration of online shoppers followed by 57% contributed by Thailand and 52% by Singapore. It is also stated that by looking at the development rate of our region in general, Malaysia are one of the speediest developing markets, keeping pace with China at 25% development rate [6][7].

In line with rapid growth in e-commerce, delivery services are currently encountering a fast development to satisfy the expanding demand. With the aims of being quick, safety, controllable and traceable. Several courier companies have built up a very extraordinary logistics network and system in their logistics process. The courier companies are playing an increasingly significant role in the supply chain of manufacturing companies. This is because of the amount of deliveries is getting larger and larger than before because of E-commerce. Innovation is required to deal with the amount of the delivery services markets. Most of the supplier in the retail area aims to create a maintainable supply chain, the job of courier companies is becoming more significant than before due to E-commerce. The effect of E-commerce in the supply chain is significant, not only in shipping fee aspect with their future way to conveyance and impacts on delivery [8][9][10].

The proposed system that the researcher proposed to small and medium company is Parcel Management System (PMS) for a small and medium company. Parcel Management System (PMS) is a computerized system for reception counter to handle all the parcels that have send to reception counter. The staff of the reception counter can use it to insert the details information of the parcel in the system, search and allocate where is location of the parcel they put, update the status of the parcel which allow the staff to know which parcel haven't collect by the recipient of the parcel, and also can delete the information of the parcel if they key in the wrong information. After the staff have enter the information of the parcel to the system and put the parcel to the locker that specially to put all the parcels, and the system will allocate the empty locker for the parcel based on the date, type of parcel and the courier company. If the recipient of the parcel wants to collect the parcel, they just must show the QR code to the staff, and the staff will scan for it. The system will verify the QR code information, if the QR code is valid so the system will tell where the location of the parcel locates at. After the recipient of the parcel have collect the parcel, the system will auto update the status of the parcel from to "Haven't Collect" to "Collected" and the empty locker will be used for the next parcel. The aim of this system is to reduce the workload and manpower to handle all these parcels, increase efficiency and save time and effort which make the staff of the reception counter easier to work and no need to use back the old method to record the details of the parcel [3][4][5][6][7].

2. Previous work

In bringing changing in customer purchasing behaviour, innovation in technology is one of the important reasons that must be considered too. Although an outstanding technology has been achieved through the introduction of Internet, consumer convenience can also be affected by the innovation in Information Technology (IT). The sign of improvement can be seen by the products being sold on the web managed to attract many customers. Other than that, the analytics of the web shops and the guidance on web shops on which the supplier is becoming more acquainted and familiar with the consumer purchasing behaviour of its consumers had also successfully induce improvement in technology. By having this innovation, the suppliers are now getting to know the frequent possible terms which are always used by the consumers to search for at the online shops, nevertheless the suppliers can get to perceive those histories which consumer already have with through online purchase. That information gathered from the consumer tells quite lots regarding the consumer purchasing behaviour thus this can be utilized to promote some of the products particularly for the consumer. With all the data they have successfully collected, every kind of advertising technique can be utilized. Moreover, the data are also used to promote a specific sort of item for selling to the consumer. Through emphasizing the advertisement on those items which more popular and gained more interest from consumers, this will have a big impact on consumer [6][7][8][9][10].

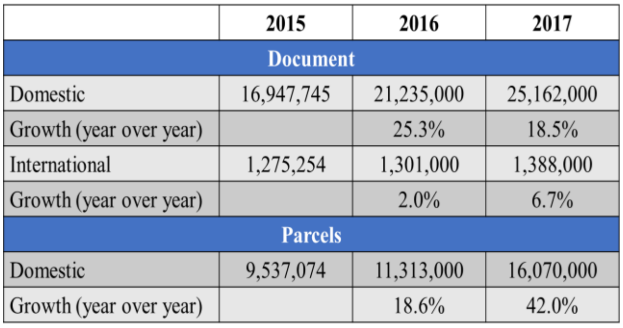
Another technique like loyalty programmers are used to acquire and keep up the consumers that are purchasing items on the web shops. Although many different kinds of factors are impacting the conductivity of the web shoppers as well as their repeated purchasing behaviour, however the utilitarian value are still caused by the most critical ones in repeated purchasing behaviour, that is the convenience factor. This value is made for consumer thanks to the good quality and high service level and consumer convenience that consumer get at the web stores. Customer loyalty is built when a better offering of prices, high service levels as well as more information are provided as what is in the physical shops. When customer loyalty is established, this will subsequently lead to repeated purchasing from these group of customers. Due to having a larger amount of consumer convenience, web shop is said to have more utilitarian value when compared to those physical shops and this bring a positive effect, which guarantees a more often online purchasing by the loyal customers [6][7][8][9][10].

In creating consumer convenience, service and fast delivery are the significant components that must be considered. Nowadays it is very usual and possible for the most common to order something on the web shop and the items are delivered straight to the consumer's place on the following day or even the same day if the delivery service is fast enough. The network of distribution must be proficient and systematically organised in order to have an express delivery of the items, leads additional consumer convenience. The increasing of demand on web shopping will cause the rising of deliveries which would largely influence the freight traffic as well as the road network. When there is changing caused, which is the shifting of bulk quantities of the physical stores into a delivery system, this may cause the changing in the distribution by warehouse too because all the orders are being packed and distributed piece by piece. Delivery to home becomes a service that is highly appreciated and applied by the consumers, but a great deal of logistical problems is arisen due to the engagement of different parties in the logistic process.

Courier company is becoming more significant because courier company can help the seller or the manufactory to deliver the goods and services to consumer. Courier companies have few methods to deliver the parcel to the consumer which is standard delivery services or express delivery services. If the consumer really needs to use that parcel emergency, then they will choose express delivery services, because they can receive the parcel in shorter time than the standard delivery services, but they must pay extra for that services. The courier company will emphasize the parcel with express delivery services first than standard delivery services. Consumer only need to pay for some delivery fee then they will get their parcel in a few days, so they will no need to go to the physical store to buy something, which is suitable and bring convenience to those busy and lazy people. Courier companies also provide a track and trace services for consumer to track down where is the location of their parcel. When the sender passes the parcel to courier company, the courier company will give the sender a tracking number. The recipient can use the track and trace services to track down the parcel anytime and anywhere, and also estimate what is the time they can receive the parcel so that they can be ready to stay at home to receive the parcel and to prevent the in charge person from courier company need to come twice to deliver the parcel if the recipient cannot receive the parcel at that time [6][7][8][9][10].

In this era, Malaysia's E-commerce transactions are charting a new height and growing at a healthy pace. As the statistic chart in Table 1 shows that the number of domestic parcels handled by courier service companies has grew by 42% to 16 million in year 2017 from 18.6% in year 2016, which supports the view that e-commerce in Malaysia is growing rapidly. In this research, we found that most bookkeeping applications are not tailored to individual needs. Therefore, designing a suitable personal application will be a good suggestion.

Table 1: Statistic chart of parcels and documents handled by courier in Malaysia (Mun, 2018)



	2015	2016	2017
Document			
Domestic	16,947,745	21,235,000	25,162,000
Growth (year over year)		25.3%	18.5%
International	1,275,254	1,301,000	1,388,000
Growth (year over year)		2.0%	6.7%
Parcels			
Domestic	9,537,074	11,313,000	16,070,000
Growth (year over year)		18.6%	42.0%

The courier company is the one who is benefit from online shopping which they involved in the back-end work of delivering the parcels from the seller of e-commerce to the consumer. The rise of e-commerce is the motive force of the burgeoning express delivery services. E-commerce contribute a lot revenue to the courier company, and it is still growing rapidly as deliver the parcel to consumer is also one of the reasons of successful of e-commerce [6][7][8][9][10].

3. Methodology

The system development approach that the researcher used for this system is Rapid Application Development (RAD) Model. The reason that the researcher selects this methodology is because there is lack of time and money to do this project because of the researcher is still a student so can't spend much money to buy the scanner to make the system more perfect but the researcher still able to use laptop's built-in webcam as QR Code scanner to scan the QR code that send to the recipient of the parcel. Since the objective of Rapid Application Development (RAD) methodologies is high quality systems, fast development and delivery and low costs. (P Beynon-Davies, 1999) Since the researcher don't have much time for this project, so the researcher needs to develop a system in a short time and without spending a lot of money. Because of the requirement analysis during planning stage is reduced, so there will have more time to be spared as the time to develop a system can be reduced. Rapid Application Development (RAD) methodologies is good for small project like Parcel Management System (PMS) since this project is an independent project, so there will be easier during decision making and required less resources when develop the project. One of the benefits that use Rapid Application Development (RAD) methodologies is prototyping, the system can be tested by using prototyping to find out where is the limitation of the system and will try to solve the limitation and improve it.

There is four of the stages in Rapid Application Development (RAD) methodology life cycle, which is requirement planning phase, user design phase, construction phase and transition (cutover) phase [11][12][13][14].

The first stage of Rapid Application Development (RAD) methodology is requirement planning stage. The stage is the most important stage in this methodology because this stage will define the objective of the system which is what the system will do and achieve when developing the system. During this stage researching the current problem also is an important activity, through doing the research the researcher will have the opportunity to take other research as reference and this helps to compare and learn something new from other people's works. Thus, other than doing researching also need to be defining the requirement of the project which is the aim, objective and the scope of the system. when all the requirement is met, the researcher will start to be finalizing the requirements by writing the proposal to be send to the stakeholders for approval. Once the proposal is being accepted, it can be avoiding miscommunications between stakeholder and developer and all the works can be started and may proceed to the other stage [15][16][17][18].

The second stage of the Rapid Application Development (RAD) methodology is user design stage. Prototype of the system will be developed in this stage. Firstly, a survey will be created based on the feature of the system, the survey will be included all the question about what the user thinks about the limitation of the current system and their review about the current system and the recommendation how to improve it in the proposed system. Secondly, the survey will be distributed to the user which is the staff of the small and medium company. Thirdly, design the diagram such as use case diagram, activity diagram, class diagram to represent the function and feature of the system. After this, can start to design the prototype interface design, then finalized the system design. At last, demonstrate the prototype interface design. After that can proceed to the next stage which is rapid construction stage.

The third stage of the Rapid Application Development (RAD) methodology is rapid construction stage. In this stage, it will be starting to develop the system and start coding based on the design at the previous stage. Preparing the data that needed to develop the system, coding, testing and implementation of the system will be conducted in this stage. In this stage will based on the interface design that created at last stage and start to coding to it and make it functionable. Testing is important and will keep take place to make sure the requirement of the system is met and find out what is the bug and try to solve it.

The last stage of the Rapid Application Development (RAD) methodology is cutover (transition) phase. In this stage, every function of the system can be used and ready to be used by the user. The system developed in this stage become operational, the system is considering a success when all the requirement of the system is met and with zero bug which is no error when using it [19][20].

4. Results

Figure 1 shows the login page of PMS which require the staff to enter the correct staff ID and Password before entering to the homepage. If the staff key in the invalid Staff ID and Password, then the system will show an error message.

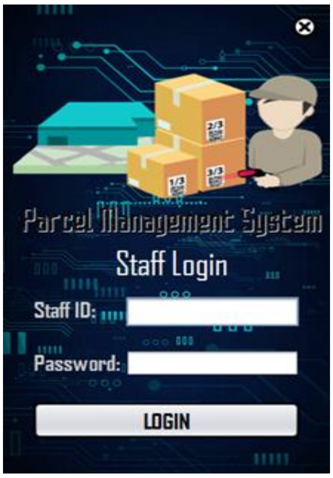

Parcel Management System
Staff Login
Staff ID:
Password:
LOGIN

Fig. 1 Login page of PMS

Figure 2 is the PMS's homepage after the staff login to the system, the staff can choose to add the parcel information by using add parcel button or empty locker button, search parcel, update parcel and delete parcel.

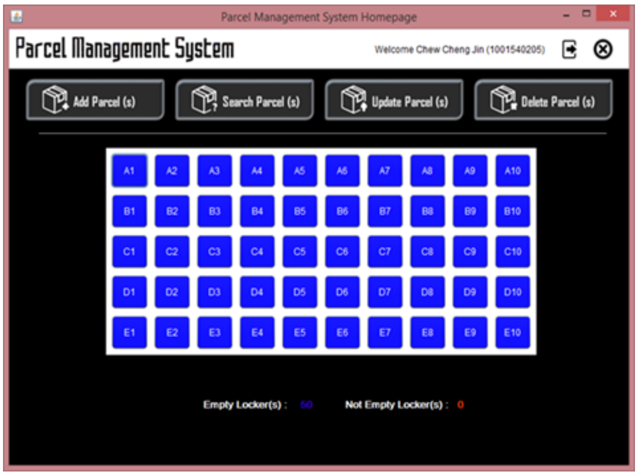
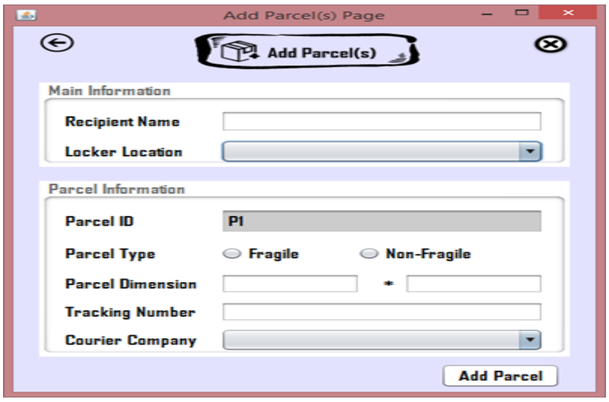

Parcel Management System Homepage
Parcel Management System
Welcome Chew Cheng Jin (1001540205)
Add Parcel (s)
Search Parcel (s)
Update Parcel (s)
Delete Parcel (s)
A1 A2 A3 A4 A5 A6 A7 A8 A9 A10
B1 B2 B3 B4 B5 B6 B7 B8 B9 B10
C1 C2 C3 C4 C5 C6 C7 C8 C9 C10
D1 D2 D3 D4 D5 D6 D7 D8 D9 D10
E1 E2 E3 E4 E5 E6 E7 E8 E9 E10
Empty Locker(s) : 50
Not Empty Locker(s) : 0

Fig. 2 Homepage of PMS

Figure 3 shows the "Add Parcel Page" after the staff press the "Add Parcel" button from homepage.


Add Parcel(s) Page
Add Parcel(s)
Main Information
Recipient Name
Locker Location
Parcel Information
Parcel ID
Parcel Type
○ Fragile
○ Non-Fragile
Parcel Dimension
*
Tracking Number
Courier Company
Add Parcel

Parcel ID: P1

Fig. 3 Add parcel (1)

Figure 4 shows the option of locker location, the combo box will only show the empty locker, if the locker number is not empty then the combo box will not have the locker number.

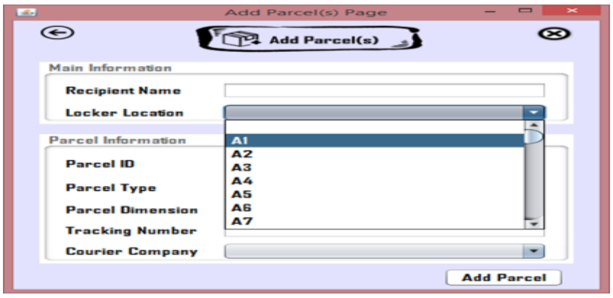
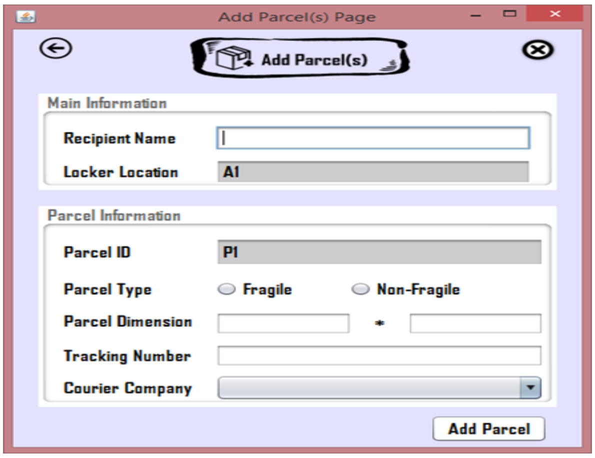

Add Parcel(s) Page
Add Parcel(s)
Main Information
Recipient Name
Locker Location
Parcel Information
Parcel ID
Parcel Type
Parcel Dimension
Tracking Number
Courier Company
A1
A2
A3
A4
A5
A6
A7
Add Parcel

Fig. 4 Add parcel (2)

Figure 5 is “Add Parcel Page (Empty Locker Button)” after the staff press the “Empty Locker” button which is blue colour button from homepage. The locker location will show the locker number that the staff just choose, and the system will lock the textbox to prevent changes.


Add Parcel(s) Page
Add Parcel(s)
Main Information
Recipient Name
Parcel Information
Parcel Type
Fragile
Non-Fragile
Parcel Dimension
Tracking Number
Courier Company
Add Parcel

Locker Location: A1

Parcel ID: P1

Fig. 5 Add Parcel Page (Empty Locker Button)

Figure 6 is “Search Parcel Page” after the staff press the “Search Parcel” button from homepage. In this page, the staff can search the location of the parcel by using QR code, Recipient ID, combo box which need the staff to select the parcel information. The staff only can check the uncollected parcel and check today’s parcel.

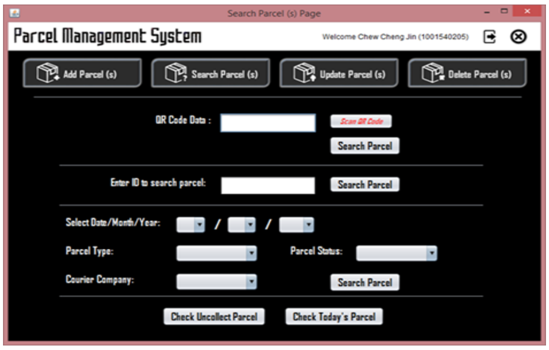

Search Parcel (s) Page
Parcel Management System
Welcome Chew Cheng Jin (1001540205)
Add Parcel (s)
Search Parcel (s)
Update Parcel (s)
Delete Parcel (s)
QR Code Data :
Scan QR Code
Search Parcel
Enter ID to search parcel:
Search Parcel
Select Date/Month/Year:
Parcel Type:
Parcel Status:
Courier Company:
Search Parcel
Check Uncollect Parcel
Check Today's Parcel

Fig. 6 Search Parcel Page

Figure 7 is “Webcam QR Code Scanner” after the staff press the “Scan QR Code” button from Search Parcel Page. In this page, the staff require to take the QR Code from recipient of the parcel and show the QR Code in front of the scanner and have to make sure that the environment and the phone brightness is bright enough, if not the scanner was not able to get the QR Code data.

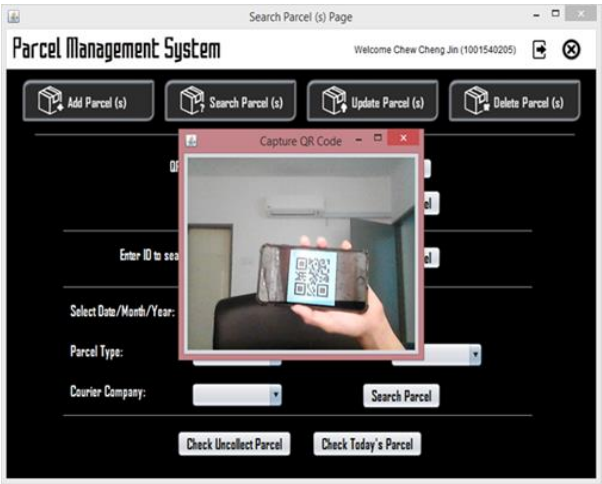
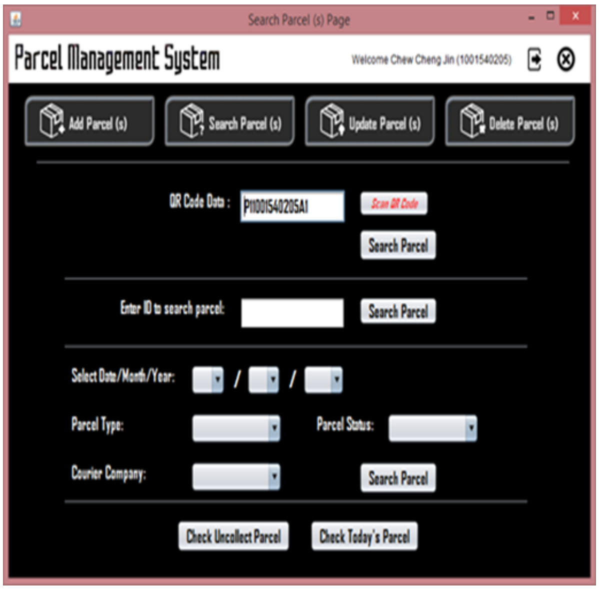

Search Parcel (s) Page
Parcel Management System
Welcome Chew Cheng Jin (1001540205)
Add Parcel (s)
Search Parcel (s)
Update Parcel (s)
Delete Parcel (s)
Capture QR Code
Select Date/Month/Year:
Parcel Type:
Courier Company:
Search Parcel
Check Uncollect Parcel
Check Today's Parcel

Fig. 7 Webcam QR code scanner

Figure 8 is “Search Parcel Button”. After the staff scan the QR Code, must select the “Select Parcel Button” to start to search the location of the parcel and it will redirect to “Search Result Page” if the QR Code is valid. The error message will show up if the text field is empty or the QR Code data is invalid.


Search Parcel (s) Page
Parcel Management System
Welcome Chew Cheng Jin (1001540205)
Add Parcel (s)
Search Parcel (s)
Update Parcel (s)
Delete Parcel (s)
Scan QR Code
Search Parcel
Enter ID to search parcel:
Search Parcel
Select Date/Month/Year:
Parcel Type:
Parcel Status:
Courier Company:
Search Parcel
Check Uncollect Parcel
Check Today's Parcel

QR Code Data : P11001540205A1

Fig. 8 QR Code Searching

5. Conclusion

In conclusion, the research project is studied and go through again and again to see whether the aim and objective that have write during proposal stage it is met or not. After the development and all the testing of the system have done, it can be said that all the requirement and goals that were proposed in the proposal during the proposal stage have been achieved. The output of this research project is the development of the Parcel Management System (PMS). The Parcel Management System (PMS) is programmed by using java programming language and use the Oracle Database. This system can help the staff of company reception counter which is responsible to manages the parcels that send to the reception counter. This system also can send the email with QR Code to the recipient of the parcel. This system only allows staff of the company to use it, any outsider's parcel will not be accepted. This system has improved the current parcel management system which is still using the traditional method to record the parcel information. With the help of the new system, the staff able to increase and speed up the overall parcel management process. Most importantly, the new system can increase the efficiency and productivity of the parcel managing and save a lot of time and effort of staff when they are trying to search and allocate the parcel.

Acknowledgment

The authors would like to thank Institute of Computer Science & Digital Innovation, UCSI University Kuala Lumpur, Malaysia for helping this project.

References

1. Asle Fagerstrømab, N. E. (2017). What's the "Thing" in Internet of Things in Grocery Shopping? A Customer Approach. ScienceDirect, 384-388.
2. C.K. Sunitha, D. M. (2010). Online Shopping-an overview. Holy Cross College, 1-10.
3. Chang, J. H. (2014). An introduction to using QR codes in scholarly journals. science editing, 113-117.
4. Foo, Y. N. (28 September, 2015). The Edge Markets. Retrieved from E-commerce boom fuels growth of courier firms:https://www.theedgemarkets.com/article/e-commerce-boom-fuels-growth-courier-firms
5. Hazlina Haron, S. Z. (2017). Critical Success Factors on Delivery Preferences of E-Commerce in Malaysia. International Journal of Academic Research in Business and Social Sciences, 1383-1393.
6. Khan, R. N. (2015). Rapid Applications Development Techniques: A Critical Review. International Journal of Software Engineering and Its Applications, 163-176.
7. Mun, N. W. (27 January, 2018). PWC. Retrieved from Malaysia's e-commerce growing despite doubts (Focus Malaysia): https://www.pwc.com/my/en/assets/press/180122-focusmy-msia-ecommerce-growing-despite-doubts.pdf?fbclid=IwAR2l49bHux9AbwsaKf61rAwuWAGsh9BFTBfWkfcCdwkzdq5kxHWqg574TtM
8. Nurul Izzah, D. R. (2016). Relationship-Courier Partner Logistics and E-Commerce Enterprises in Malaysia: A Review. Indian Journal of Science and Technology, 1-10.
9. P Beynon-Davies, C. C. (1999). Rapid application development (RAD): an empirical review. European Journal of Information Systems, 211-223.
10. Pei-Ju Wu, K.-C. L. (2017). Unstructured big data analytics for retrieving e-commerce logistics knowledge. ScienceDirect, 237-244.
11. Poort, C. (2017). E-commerce on the supply chain. Erasmus School of Economics, 1-51.
12. Varun Gupta, V. K. (2018). Analysis of shopping trends employing E-Commerce Application: A Comparative Case Study. ScienceDirect, 1728-1738.
13. Hafeez, K., Keoy, K.H. and R. Hanneman (2006), "E-business capabilities model: Validation and comparison between adopter and non-adopter of e-business companies in UK", Journal of Manufacturing Technology Management. A Special Issue on E technology and Manufacturing Enterprise Competitiveness. ISBN: 101-84663-080-013. Vol. 17, No. 6, pp. 806-828.
14. Mazumdar Asnan Tajul Islam., Islam Shayla, Thong Chee Ling & Keoy Kay Hooi. (2022). "NFC-based Mobile Application for Student Attendance in Institution of Higher Learning". 1st International Conference on AI in Cybersecurity (ICAIC) 24 May-26 May 2022. University of Houston-Victoria, USA.
15. Sulaiman, Raduni, D. N. A. Jawawi, and Shahliza Abdul Halim. "Features and Behaviours Mapping In Model-based Testing in Software Product Line." IOP Conference Series: Materials Science and Engineering. Vol. 884. No. 1. IOP Publishing, 2020.
16. Sulaiman, Rabatul Aduni, Dayang NA Jawawi, and Shahliza Abdul Halim. "Cost-effective test case generation with the hyper-heuristic for software product line testing." Advances in Engineering Software 175 (2023): 103335.
17. S. N. Wahab, Y. M. Loo, and C. S. Say, "Antecedents of blockchain technology application among Malaysian warehouse industry," Int. J. Logist. Syst. Manag., vol. 37, no. 3, pp. 427–444, 2020, doi: 10.1504/IJLSM.2020.111414.
18. S. N. Wahab, "Transforming Malaysian On-Demand Delivery Apps for Efficient Last-Mile Delivery: A SWOT Analysis," in 2021 International Conference on Data Analytics for Business and Industry (ICDABI), 2021, pp. 396–399.
19. Ammar Ashraf Narul Akhla, Thong Chee Ling, Abdul Samad Shibghatullah, Chit Su Mon, Aswani Kumar Cherukuri, Chaw Lee Yen and Lee Chiw Yi,2022, "Impact of Real-Time Information for Travellers: A Systematic Review", Journal of Information & Knowledge Management, 21(4),2250065-1-2250065-21.
20. Gan Shu Qian, Ammar Ashraf Bin Narul Akhla, Thong Chee Ling, Abdul Samad Shibghatullah, Chit Su Mon, Chaw Lee Yen, and Lee Chiw Yi, 2022, "A Comparative Study of Navigation API ETA Accuracy for Shuttle Bus Tracking", Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 13337 LNCS, 446–461.

Authors Introduction

Chew Cheng Jin

He completed his Bachelor of Science in Computing at UCSI University, Cheras, Kuala Lumpur on 15th August 2020. Now he is working and preparing to continue his Master degree. His research interest are in Information System, Computing and Business.

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. She obtained her doctoral degree from Malaysia University of Science & Technology. She has supervised many undergraduate students as main supervisors and co-supervisors. She has publications in Scopus-Indexed Journals and Web of Science. She was also a co-researcher in Pioneer Science Incentive Fund (PSIF) in the area of augmented reality. Her research interests include human-computer interaction, human personality types, augmented reality, e-learning, mobile commerce and e-commerce.

Abdul Samad 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 Associate Professor at Institute of Computer Science & Digital Innovation, UCSI University, Kuala Lumpur, Malaysia. His current research interests include optimization, modelling and scheduling

Salini Devi Rajendran

She is a senior lecturer at the School of Food Studies and Gastronomy, Faculty of Social Sciences and Leisure Management, Taylor's University, Kuala Lumpur. She received her Ph.D. in Halal Product Management from Halal Products Research Institute, Universiti Putra Malaysia (UPM). She was also awarded as a professional technologist by the Malaysia Board of Technologists (Ts) and a chartered member of the Chartered Institute of Logistics and Transport (CMILT).

Nur Hazirah Hamdan

She received her bachelor degree in finance from Universiti Teknologi MARA, Melaka, Malaysia in 2011, Master Business Administration from Universiti Teknologi MARA, Shah Alam, Malaysia in 2014, and Ph.D. holder in Islamic Banking and Finance from International Islamic University Malaysia in 2022. She is currently Senior Lecturer at Universiti Teknologi MARA, Pahang Branch, Raub Campus, Malaysia. Her research interests include Islamic finance and banking, corporate finance, entrepreneurship and technology acceptance and adoption.

Automatic Classification Method for Plastic Bottles and Caps Using Multi Attention Eff-UNet

Shunsuke Moritsuka, Tohru Kamiya
Kyushu Institute of Technology, 1-1 Sensui-cho, Tobata-ku, Kitakyushu-shi, Fukuoka 804-8550, Japan
E-mail: kamiya@cntl.kyutech.ac.jp

Abstract

In Japan, increasing amounts of waste are becoming a social problem. One of approaches to solve the problem is recycle of the plastic bottles. However, they are thrown away with their caps still attached, and it should be removed by hand. To solve this problem, we developed a method for automatic identification of plastic bottles and caps using deep learning technique. In this paper, we propose a method that combines different numbers of Efficient blocks and adds an attention structure and verify its usefulness through experiments.

Keywords: Plastic bottle recycling, Convolutional neural network, Segmentation, Eff-UNet, Attention

1. Introduction

In Japan, increasing amounts of general waste are becoming a social problem. It is increasing year by year and is expected to be as large as 41.67 million tons in 2020[1], which will fill up landfills soon. Therefore, it is necessary to reduce the amount of general waste as much1as possible and to devise ways to prevent landfills from becoming full. One solution to reduce the amount of general waste is to reuse plastic bottles. For this reason, some communities have special bags for disposing of plastic bottles. However, many bottles are thrown away with the caps still attached to bags, which are manually sorted at the landfill after checking with the human eye to see if the caps are still attached. This involves standing for long periods of time doing the same work repeatedly, which causes fatigue and oversight.

In recent years, digital cameras and other devices have become more precise and sensitive, and AI (Artificial Intelligence) technology has been developed in analysis software. Therefore, it was thought that by effectively utilizing deep learning, automatic detection of plastic bottles and caps could be realized, automatic sorting could be performed, and the problem could be solved.

Based on an above mentioned, this paper describes a development of a new model that combines Eff-UNet structure, which specializes in semantic segmentation, with an attention structure that removes irrelevant regions such as the background, among Convolutional Neural Networks (CNN), one of deep learning methods, using images of plastic bottles obtained from the Internet and the evaluation of its performance.

2. Method

This section describes an image analysis technique that automatically extracts regions of a plastic bottle and its cap from an input image. Although various methods exist for image extraction, semantic segmentation is employed in this study because it is sufficient to identify regions of the plastic bottle and cap. The proposed method is described below.

2.1. *U-Net*

U-Net[2] is a model which is developed for biomedicine by Olaf et al. and consists of an encoder and a decoder (Fig. 1). In the encoder section, an input image is convolved multiple times to extract image features. In the decoder section, up-sampling is performed to obtain a probability map of the same size as an input image based on the features extracted by the encoder. Since the feature values are lost in this process, a skip connection structure is used to convey the information of the large feature map on the encoder side to the decoder side, facilitating the acquisition of object location information during it.

2.2. *Eff-UNet*

Eff-UNet[3] is a model proposed by Baheti, in which the encoder part of U-Net is changed to EfficientNet[4] as shown in Fig. 2. It is a model created by combining squeeze and excitation-optimized mobile inverted bottleneck convolution (MBConv)[5]. Fig.3 shows EfficientNet and Fig. 4 shows MBConv respectively. It is characterized by its ability to balance three parameters of model depth, breadth, and resolution. Therefore, an improvement of accuracy by scaling up the model has already been implemented. It also saves time compared to conventional models because the model is trained with optimal parameters.

Based on an above, it is a model that is expected to improve the overall performance of the algorithm by incorporating a strong CNN like EfficientNet as an encoder and U-Net as a decoder with a skip-connection structure that can minimize information, and This method was also considered to be effective in this study.

2.3. *Attention Gate (AG)*

In this paper, we adopt the Attention Gate (AG) used in Attention U-Net[6] devised by Oktay et al. Fig. 5 shows the AG architecture. It consists of a flow where features obtained by up-sampling are passed through AG and multiplied with the original output, and a skip-connection structure. by using Sigmoid function as last activation function of AG, original features are Since the value of 0 to 1 is multiplied, the value of the region of interest remains unchanged, while the values of other regions become close to 0, enabling learning that is specialized to the region of interest. Therefore, while background information unrelated to recognition may have a negative impact in normal image recognition, the use of attention makes it possible to focus attention on relevant regions.

2.4. *Proposed Method (MAE-UNet)*

We propose MAE-UNet (Multi Attention Eff-UNet). The architecture of the proposed method is shown in Fig. 6. There are two key points. First, AG is used in the decoder part of Eff-UNet like Attention U-net to remove regions not related to recognition and reduce over-detected regions. Second, inspired by the DDU-Net[7] proposed by Tang et al., we combine Eff-UNet up to block6, which is dedicated to global segmentation, and Eff-UNet up to block4, which is dedicated to local segmentation. Conventional Eff-UNet has a skip connection structure specialized for global segmentation, which causes a problem of information loss during learning. Therefore, by using Eff-UNet with a small number of MBConv blocks, local segmentation can be performed, and the loss of information can be minimized, leading to improved accuracy.

3. Experiment

This section describes data set, evaluation method, and its experimental results. Specifically, one plastic bottle is placed in the center of each image.

3.1. *Data Set*

The data set used in this study consists of 210 images, 200 obtained from the Internet and 10 images taken by us. During training, data augmentation was performed by randomly flipping horizontally, scaling, smoothing, and normalizing. An annotation work was done by our own hands and divided into three classes: plastic bottles, caps, and backgrounds.

3.2. *Evaluation Method*

We evaluate the accuracy by 5 cross validations. It means that all data is randomly divided into five sets. The method is to use one set of them for testing (42 images) and the remaining four sets (168 images) for learning, evaluate them, and determine the average value for the five sets to obtain many results. The accuracy of the segmentation is determined by the Intersection over Union (IoU), which is the ratio of the product set of predictions and grand tours computed for each class to the sum set, and the average Intersection over Union (mIoU), which is obtained by averaging the IoU for each class. The equation for IoU is shown in Eq. (1). The Intersection region in the equation is the product set of the annotation and output results, and the Union region is the union set.

$$IoU = \frac{area\ of\ Intersection}{are\ of\ Union} \tag{1}$$

3.3. *Experimental Result*

The experimental results are shown in Table. 1 for comparison with three methods: U-Net, Eff-UNet, and MAEU-Net proposed in this paper.

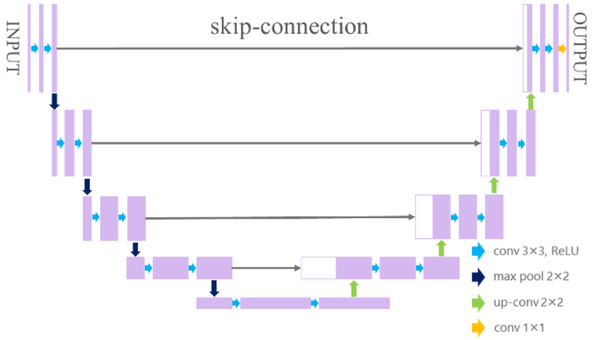

INPUT
skip-connection
OUTPUT
conv 3×3, ReLU
max pool 2×2
up-conv 2×2
conv 1×1

Fig.1. U-Net Architecture

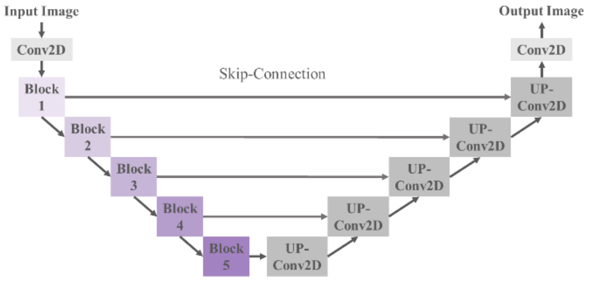

Input Image
Conv2D
Block 1
Skip-Connection
Block 2
Block 3
Block 4
Block 5
UP-Conv2D
Output Image

Fig.2. Eff-UNet Architecture

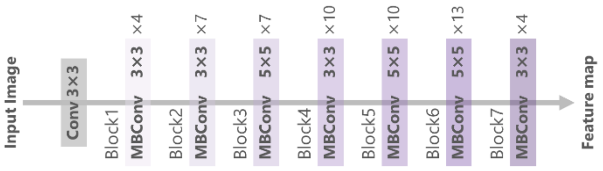

Input Image
Conv 3×3
Block1 MBConv 3×3 ×4
Block2 MBConv 3×3 ×7
Block3 MBConv 5×5 ×7
Block4 MBConv 3×3 ×10
Block5 MBConv 5×5 ×10
Block6 MBConv 5×5 ×13
Block7 MBConv 3×3 ×4
Feature map

Fig.3. EfficientNet Architecture

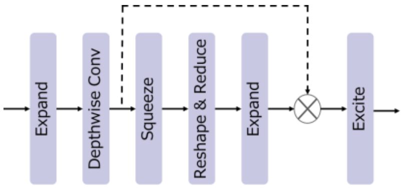

Expand
Depthwise Conv
Squeeze
Reshape & Reduce
Expand
Excite

Fig.4. MBConv Architecture

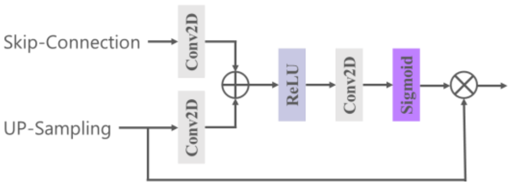

Skip-Connection
UP-Sampling
Conv2D
Conv2D
ReLU
Conv2D
Sigmoid

Fig.5. Attention Gate (AG) Architecture

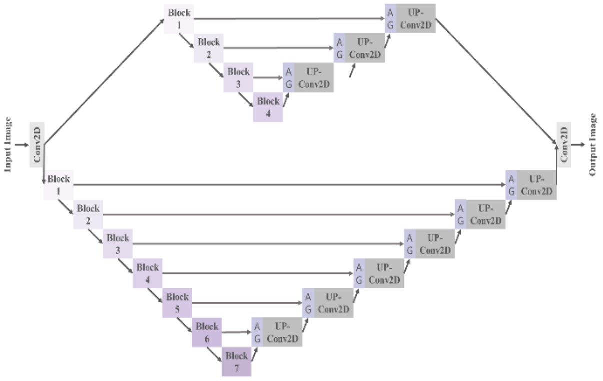
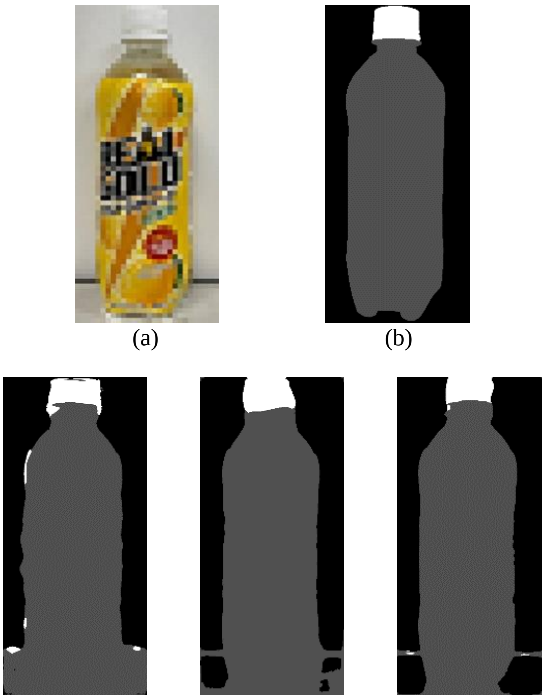

Input Image
Conv2D
Block 1
Block 2
Block 3
Block 4
Block 5
Block 6
Block 7
A G
UP-Conv2D
Output Image

Fig.6. Propose Method Architecture



(a) (b)

(C) (d) (e)

Fig. 7. Experimental Result (Gray:Body, White:Cap); (a)Input Image (Blur image) (b)CorrectLlabel (c)Segmentation Result (U-Net) (d)Segmentation Result (Eff-UNet) (e)Segmentation Result (MAE-UNet)

Table 1. Experiment Result

Model	Bottle	Cap	mIoU
U-Net	0.974	0.817	0.895
Eff-UNet	0.976	0.837	0.906
MAE-UNet	0.980	0.878	0.923

4. Discussion

In this paper, experiments were conducted using images obtained from an Internet as well as images of plastic bottles photographed. Table 1. shows that the proposed method gave the best results, improving accuracy by 2.8% over U-Net and 1.7% over Eff-UNet. The reason for improved accuracy can be attributed to the fact that the attention structure suppressed an over detection of background regions. Experimental result shown in Fig. 7 also indicate that a lower over detection is suppressed. In addition, the use of a model that combines different numbers of Eff-UNet blocks is considered to have improved accuracy compared to the conventional method. From Fig. 7, the roundness of a cap can be sensed when the proposed method is used. However, since the accuracy of the caps does not exceed 90%, further improvement in accuracy is required to achieve practical application. As a solution, it was considered that the area of the cap was so small for the plastic bottle that there might be a loss of features. Therefore, it was thought that the loss of information could be reduced by complicating the skip connection structure as in Unet++[8]. Also, since the number of images is very small, increasing the number of images is also effective.

5. Conclusion

In this paper, to improve the efficiency of plastic bottle sorting, we constructed a model combining different numbers of blocks of Eff-UNet, which incorporates the Attention structure in the encoder part, and extracted the regions of PET bottle bodies and caps. As a result, we succeeded in improving the accuracy of the model compared to the other two methods (U-Net and Eff-UNet). In the future, we plan to improve the accuracy by improving the partitioning network and increasing the number of data sets.

References

1. Ministry of the Environment, “Waste Management in Japan (2020 Edition)”.
2. O. Ronneberger et al., “U-Net: Convolutional Networks for Biomedical Image Segmentation”, International Conference on Medical Image Computing and Computer-assisted Intervention, pp.234-241, 2015.
3. B. Baheti. et al., “Eff-UNet: A Novel Architecture for Semantic Segmentation in Unstructured Environment”, Proceedings 2020 Institute of Electrical and Electronics Engineers/CVF Conference on Computer Vision and Pattern Recognition Workshops, pp.358-359, 2020.
4. M. Tan. et al., “Effcientnet: Rethinking Model Scaling for Convolutional Neural Networks”, International Conference on Machine Learning, pp.6105-6114, 2019.
5. Y. Lyu. et al., “Factorizing and Reconstituting Large-kernel MBConv for Lightweight Face Recognition”, Proceedings of the Institute of Electrical and Electronics Engineers/CVF International Conference on Computer Vision Workshops, 2019.
6. O. Oktay. et al., “Attention U-Net: Learning Where to Look for the Pancreas”, Accepted to Publisher in Medical Imaging with Deep Learning 2018, arXiv: 1804.03999, 2018. (Access 2022/11/09)
7. M. Jiang. et al., “A Novel Deep Learning Model DDU-net Using Edge Features to Enhance Brain Tumor Sgmentation on MR Images”, Artificial Intelligence in Medicine 121, 2021.
8. Z. Zhou. et al., “UNet++: ANested U-Net Architecture for Medical Image Segmentation”, 4th Deep Learning in Medical Image Analysis (DLMIA) Workshop, pp.3-11, 2018.

Authors Introduction

Mr. Shunsuke Moritsuka

Shunsuke Moritsuka graduated from the Department of Mechanical and Control Engineering, Faculty of Engineering, Kyushu Institute of Technology, Japan in 2021. He is currently a Master course student in Kyushu Institute of Technology, Japan.

Dr. Tohru Kamiya

Tohru Kamiya 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 control engineering at Kyushu Institute of Technology. His research interests are focused on medical application of image analysis. He is currently working on automatic segmentation of multi-organ of abdominal CT image, and temporal subtraction of thoracic MDCT images

Research on the structure of consciousness of people who maintain and manage parks

Jun Sanbuichi
Interdisciplinary Graduate School of Agriculture and Engineering, University of Miyazaki,
1-1 Gakuen Kibanadai-nishi, Miyazaki 889-2192, Japan

Minoru Kumano
Interdisciplinary Graduate School of Agriculture and Engineering, University of Miyazaki,
1-1 Gakuen Kibanadai-nishi, Miyazaki 889-2192, Japan

Toru Hiraoka
Faculty of Information Systems, University of Nagasaki,
123 Kawashimo-cho, Sasebo-shi, Nagasaki 858-8580, Japan
E-mail: na21003@student.miyazaki-u.ac.jp, kumano@cc.miyazaki-u.ac.jp,
hiraoka@sun.ac.jp
www.miyazaki-u.ac.jp

Abstract

Due to population decline, unused vacant land is expected to increase in the region in the future. However, consciousness indicators of how to encourage maintenance and management are not clear. Therefore, the purpose of this study is to clarify the consciousness structure of people who maintain and manage park. As a result, it was proved that six consciousnesses (Interest in greening etc.) are related to the comprehensive consciousness index of park maintenance behavior.

Keywords: Consciousness structure, Maintaining managed, Covariance structure analysis

1. Introduction

In the future, population is expected to decline in Japan[1]. As a result, unutilized vacant land is expected to increase in each region. So, it is considered important to have local residents maintain and manage local spaces. National and local governments have also begun to encourage such behavior[2]. But consciousness indicators of how to encourage maintenance and management are not clear. In a previous study that addresses this issue, there is a study that focused on the nawabari theory. Nawabari theory is the belief that humans (like animals) exhibit territorialized consciousness and behavior. The research of theory is a direction that encourages maintenance and management. Kobayashi studied territorialized formation mainly from the hardware and provided directions[3],[4]. Fujitani et al. conducted a follow-up study and confirmed that human territorial behavior is also not easily influenced by the times. And software of consciousness was also found to have a significant impact on territorialized formation[5]. Sanbuichi et al. focuses on the software of consciousness and explores the relationship between territorialized behavior and consciousness[6],[7]. However, there is a lack of analysis of spaces that are used distantly and publicly. As a result, we have not been able to demonstrate a consciousness index that would encourage the maintenance and management of the space.

Therefore, the purpose of this study is to clarify the consciousness structure of people who maintain and manage park.

2. Research methods

This study follows the flow chart as shown in Fig. 1. First, select consciousness which seems to be associated with maintaining park. Given that a mix of consciousnesses influences the consciousness associated with maintaining park according to the existing papers[3],[4],[5],[6],[7](Include reference papers for that paper.), consciousnesses are preset for the purpose of this study(Fig. 2). Based on the preset consciousnesses, questions (evaluation items) are developed and the survey is conducted to people who maintain park. Factor and covariance structure analyses will be performed on all obtained data to demonstrate the existence of the preset consciousness and to quantitatively measure the importance of observed variables.

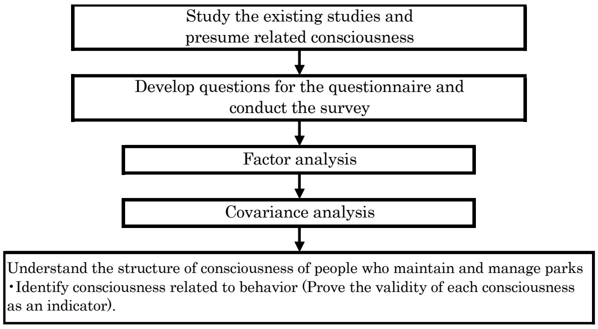

Study the existing studies and presume related consciousness
Develop questions for the questionnaire and conduct the survey
Factor analysis
Covariance analysis
Understand the structure of consciousness of people who maintain and manage parks
・Identify consciousness related to behavior (Prove the validity of each consciousness as an indicator).

Fig.1 Research flow

3. Research results

The survey was administered directly to the target population and direct responses were collected (June 14, 2008). The target audience is the people who maintain the park. (They are mainly volunteers who belong to “HANA MIDORINOKAI.”) The obtained individual attributes of the respondents are shown in Table. 1. The next section will elicit conclusions from total 150 collected response sheets. The consciousness associated with maintaining park appear to be associated with consciousness of interest in greening, secure feeling in the social field, consciousness in city planning, expectation for human effects, resistance to the other people and resistance to tasks according to the reference materials and the transforming process of consciousness (desires to appeal and pride, keep holding, expand the territory, promotion factors, etc).

Table.1 Individual Attributes of Respondents

Attributes	%	Attributes	%
Gender:		Years of residence:	
Male	41.9	3 years or shorter	1.6
Female	56.5	3 to less than 5 years	3.1
Non-responder	1.6	5 to less than 10 years	3.1
Age:		10 to less than 15 years	17.2
20s	1.6	15 to less than 20 years	57.8
30s	3.1	20 years or longer	17.2
40s	7.8		
50s	14.3	Distributed questionnaire sheets: 64	
60s	31.0	Effective collected response sheets: 64	
70s or greater	42.2	Effective response rate: 100%	

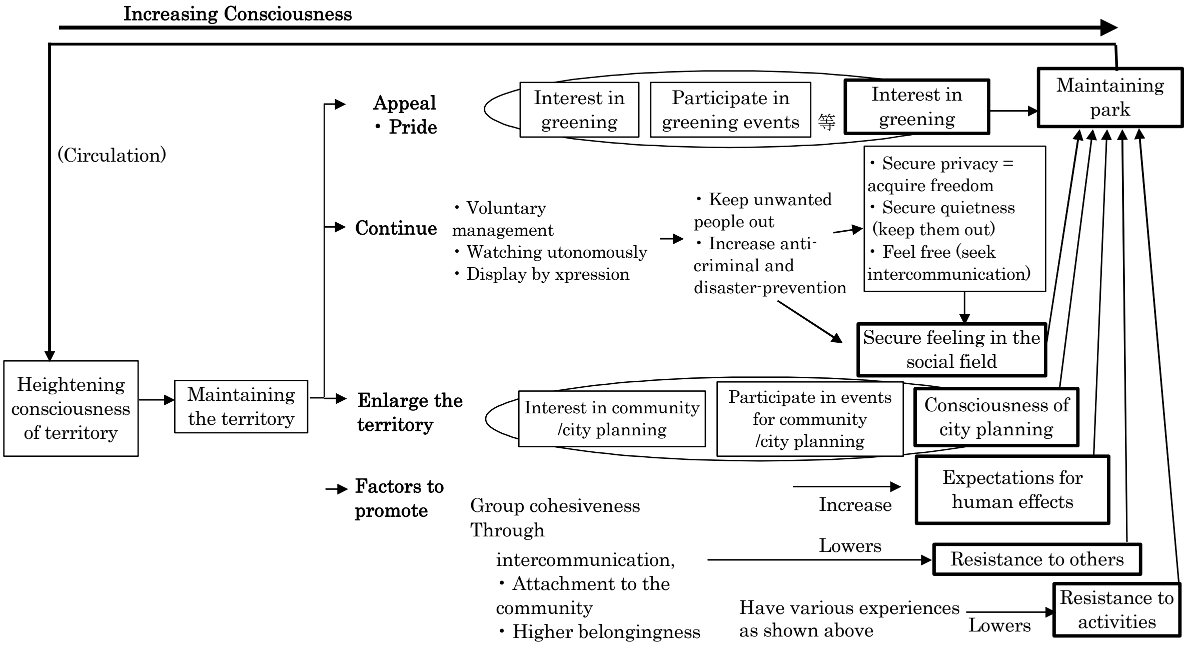

Increasing Consciousness
(Circulation)
Appeal
・Pride
Interest in greening
Participate in greening events
等
Interest in greening
Maintaining park
Continue
・Voluntary management
・Watching utonomously
・Display by xpression
・Keep unwanted people out
・Increase anti-criminal and disaster-prevention
・Secure privacy = acquire freedom
・Secure quietness (keep them out)
・Feel free (seek intercommunication)
Secure feeling in the social field
Heightening consciousness of territory
Maintaining the territory
Enlarge the territory
Interest in community /city planning
Participate in events for community /city planning
Consciousness of city planning
Factors to promote
Group cohesiveness Through intercommunication,
・Attachment to the community
・Higher belongingness
Increase
Expectations for human effects
Lowers
Resistance to others
Have various experiences as shown above
Lowers
Resistance to activities

Fig.2 Consciousness assumed to be related to maintaining park (process of transforming consciousness)

The authors set up questions related to the above (Table. 2). The respondents were asked to answer the questions in 5-point scale: 1. No, I do not (not applicable), 2. Not so much, 3. I do not know, 4. Maybe, and 5. Yes, I do (applicable). In the obtained responses, those to the questions highly correlated each other will be compiled, and be subjected to the factor analysis to highlight the residents' sub-consciousness and establish the latent variables in the covariance structural analysis. Table 3 shows the definitions of factor loading, eigenvalue, cumulative contribution rate and factors in the factor analysis (after the rotation by the varimax method). Since Factor 1 is high in the factor loading with D1:" Maintenance of park helps greening and beautification of the community," D2:" The streets and the community will look better by maintaining park," and D3:" Maintenance of park activates community activities in the area," it is defined as the expectations for human effects. Similarly, Factor 2 is defined as the secure feeling in the social field, Factor 3 as interest in greening, Factor 4 as Resistance to activities, Factor 5 as consciousness of city planning and Factor 6 as resistance to other people (the number of factors is determined as six according to the transforming process of consciousness and the factor loading values). The results of the factor analysis are defined as the latent consciousness of the managing of park, and they are used as the latent variables in the structural model of the resident consciousness for the purpose of this paper. The upper variables are given to (the consciousness associated with) maintaining private garden, and the latent variables obtained in the factor analysis between the observed variables and the upper variables to build the structural model of consciousness (path chart) by taking the preset consciousnesses into account (see Fig.3). Table 4 shows the results of the covariance structural analysis in the structural model of consciousness. For the information, Amos5.0 (SPSS) is used for the verification of the parameters. The covariance structural analysis resulted in that the parameters meet a significant level of 1% or 5%. Furthermore, the values of GFI (Goodness of Fit Index: 0.825) and AGFI(Adjusted Goodness of Fit Index: 0.757) are relatively good. As a result, it was proved that six consciousness (Interest in greening, secure feeling in the social field, consciousness of city planning, Expectations for human effects, Resistance to other people and Resistance to activities) are related to the comprehensive consciousness index of park maintenance behavior.

Table. 2 Evaluation items for assumed consideration

No.	Code	Evaluation items
1	A1	I like have flowers and greens in my house.
2	A2	I like glowing flowers.
3	A3	I like seeing flowers and greens of others or other houses.
4	B1	I have many acquaintances in the area where I live.
5	B2	I am very familiar to the area where I live.
6	C1	I want to take up my opinion to the city planning.
7	C2	I want to attend a symposium regarding the city planning.
8	C3	I want to participate in the festival or events in my community.
9	D1	Maintenance of park helps greening and beautification of the community.
10	D2	The streets and the community will look better by maintaining park.
11	D3	Maintenance of park activates community activities in the area.
12	E1	I am not concerned about other people's behaviors and what they say.
13	E2	I like have chatting with people.
14	F1	I do not mind if I have to spend money more to maintain park.
15	F2	I do not mind if tasks to maintain park increase.
16	F3	I do not mind spending time to maintain park.

Table. 3 .Results of factor analyses of consciousnesses

Name of variables	factor1	factor2	factor3	factor4	factor5	factor6
D 1	0.865	0.075	0.067	0.181	0.228	0.076
D 2	0.861	0.146	-0.024	0.240	0.032	0.122
D 3	0.733	0.108	0.104	0.065	0.155	0.179
B 1	0.060	0.853	0.104	0.125	0.123	0.124
B 2	0.100	0.839	0.146	0.119	0.164	0.018
A 1	-0.015	-0.060	0.706	0.194	-0.111	0.143
A 2	-0.043	0.084	0.639	0.110	-0.044	0.032
A 3	0.159	0.135	0.536	0.056	0.054	-0.048
F 1	0.198	0.118	0.199	0.682	0.039	0.036
F 2	0.378	0.081	0.269	0.679	0.082	0.261
F 3	0.026	0.208	0.204	0.440	0.224	0.498
C 1	0.033	0.169	-0.037	0.065	0.683	0.182
C 2	0.248	0.198	-0.064	0.063	0.595	-0.149
C 3	0.202	0.495	-0.036	-0.004	0.355	0.136
E 1	0.383	-0.007	-0.042	0.090	-0.106	0.454
E 2	0.282	0.200	0.086	0.082	0.127	0.416
Eigen value	2.581	1.915	1.412	1.320	1.167	0.864
Culumulative contribution rate	16.1%	28.1%	36.9%	45.2%	52.5%	57.9%
Definitions of factors	Expectations for human effects	Secure feeling in the social field	Interest in greening	Resistance to activities	Consciousness of city planning	Resistance to other people

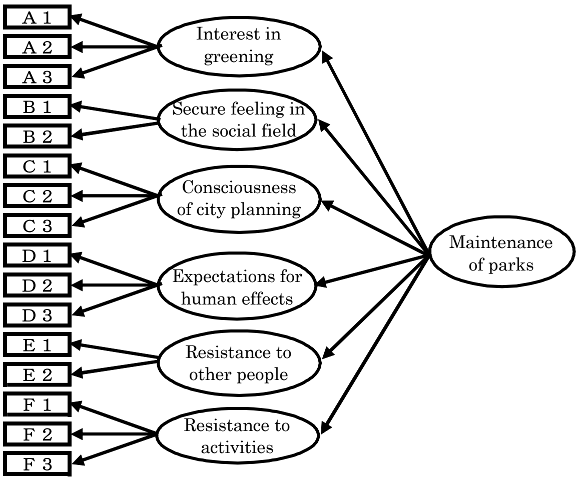

A 1
A 2
A 3
B 1
B 2
C 1
C 2
C 3
D 1
D 2
D 3
E 1
E 2
F 1
F 2
F 3
Interest in greening
Secure feeling in the social field
Consciousness of city planning
Expectations for human effects
Resistance to other people
Resistance to activities
Maintenance of parks

Fig.3 Structural Model of Consciousness of park (Path Chart)

4. Conclusion

This study is intended to draw indexes of the consciousnesses in the process to maintaining park, i.e., the process to create a social field, in the process through the survey of consciousness of the residents. The study hypothesized a transformation process of consciousness and conducted a questionnaire survey, factor analysis, and covariance structure analysis. The results revealed that actions to maintain the park are influenced by six consciousnesses (interest in greening, secure feeling in the social field, consciousness of city planning, Expectations for human effects, Resistance to other people and Resistance to activities).

The survey of residents' consciousness identified part of the process of creating a social field. However, clarifying the territorialization process of some spaces leaves many mysteries unsolved. It is important to conduct a comprehensive survey and discussion of attitudes in each space in the future, including those who maintain and manage close public spaces and those who maintain and manage remote private spaces.

Table.4 Results of covariance structural analysis of consciousnesses

Relation of each variable			Path coefficient	t-value	確率
Interest in greening	<---	Maintenance of parks	0.292	1.996	*
Secure feeling in the social field	<---	Maintenance of parks	0.684	2.491	*
Consciousness of city planning	<---	Maintenance of parks	0.464	2.448	*
Expectations for human effects	<---	Maintenance of parks	0.995	4.06	**
Resistance to other people	<---	Maintenance of parks	0.725	3.115	**
Resistance to activities	<---	Maintenance of parks	1		
A 1	<---	Interest in greening	1		
A 2	<---	Interest in greening	0.638	3.166	**
A 3	<---	Interest in greening	0.496	2.901	**
B 1	<---	Secure feeling in the social field	1		
B 2	<---	Secure feeling in the social field	0.978	4.237	**
C 1	<---	Consciousness of city planning	1		
C 2	<---	Consciousness of city planning	0.992	3.167	**
C 3	<---	Consciousness of city planning	0.802	2.995	**
D 1	<---	Expectations for human effects	0.967	11.287	**
D 2	<---	Expectations for human effects	1		
D 3	<---	Expectations for human effects	0.785	7.601	**
E 1	<---	Resistance to other people	1		
E 2	<---	Resistance to other people	0.856	2.612	**
F 1	<---	Resistance to activities	0.736	5.205	**
F 2	<---	Resistance to activities	1		
F 3	<---	Resistance to activities	0.657	4.662	**
GFI			0.825		
AGFI			0.757		

**:1% Significance level *:5% Significance level

References

[1] Statistics bureau of Japan: Population estimation, Ministry of internal affairs and communications, 2022

[2] Ministry of land, infrastructure, transport and tourism (Land policy division of real estate and construction economics bureau): New utilization of vacant lots(etc). -Advanced initiatives related to the utilization of vacant lots(etc), Ministry of land, infrastructure, transport and tourism, 2021.

[3] Kobayashi H., Suzuki S: A study on "territory of grouped inhabitants" in small housing groups -Part 1 process of territialization-, Journal of Architecture and Planning (Transaction of AIJ), vol.307, pp102-111, 1981.

[4] Kobayashi H: "Shuju no nawabararigaku" Territoriality in hosing, shokokusha, 1992.

[5] Fujitani H., Jung J. and Kobayashi H: Long-term changes of group territory in urban low-raise housing- A study on long-term changes of territoriality part1-, Journal of Architecture and Planning (Transaction of AIJ), vol.77, No.672, pp283-289, 2012.

[6] Sanbuichi J., Yuzawa A. and Kumano M: An examination about structure of consciousness of open gardener, Journal of the Japanese Institute of Landscape Architecture, Vol.70, No.5, pp.391-396, 2007.

[7] Sanbuichi J. and Yuzawa A: A Study of Maintaining Private Garden and Disparities in Consciousnesses Associated with Opening or Closing Private garden, Journal of Architecture and Planning (Transaction of AIJ), Vol. 75, No. 647, pp129-138, 2010.

Authors Introduction

Mr. Jun Sanbuichi

He received B.S degrees in environmental construction engineering in 2005 from National Institute of Technology, Tokuyama College. He received M.S degrees in Construction Engineering in 2007 from Maebashi Institute of Technology. He is acquiring the PhD in University of Miyazaki.

Dr. Minoru Kumano

He is a professor at Miyazaki University in Japan. 1981 Graduated from the Department of Construction Engineering, Toyohashi University of Technology.

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.

An Accuracy Evaluation of Multibody Dynamics for the Knee Support Exoskeleton Model with Respect to Implicit Methods for Numerical Integration

Shintaro Kasai
Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku, Kitakyushu, 808-0196, Japan

Hiroaki Wagatsuma
Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku, Kitakyushu, 808-0196, Japan
Email: kasai.shintaro660@mail.kyutech.jp, waga@brain.kyutech.ac.jp

Abstract

Numerical integration takes an important role to analyze models described by ordinary differential equations and it largely contributes to an assurance of the accuracy in displacement analyses of multibody dynamics if the model consists of several bodies especially with dynamic components such as springs and dampers. Exoskeletal assistive devices require flexible materials for absorbing reaction forces from human joints, which implies an inevitable necessity of an accurate evaluation of elastic effects in the model. In the present study, we introduced implicit methods for numerical integration to analyze the model under the formulation of multibody dynamics and computer experiments demonstrated results from explicit and implicit methods for the numerical integration as a comparison analysis. It can improve a degree of accuracy in inverse dynamics even in the theory of flexible Multibody Dynamics (fMBD).

Keywords: Numerical integration, multibody dynamics (MBD), analysis error, ordinary differential equation

1. Introduction

For the sake of analyses of kinematics and kinetics that consist of multiple bodies known as the multibody systems, multibody dynamics (MBD) analysis is getting to be the standard method [1],[2],[3],[4],[5]. Thus, the description of ordinary differential equations accompanied with an appropriate numerical integration acts an important role in ensuring accuracy, especially in the displacement analysis of systems, which are incorporating dynamic elements such as spring and damper components [6]. Indeed, flexible materials are of interest for consideration of bodies absorbing and utilizing reaction force. In consideration of human joint models, those factors are inevitable not only for the detailed analysis of the joint mechanisms but also development of exoskeleton type support device which assists human movement. Therefore, it is highly important to reconstruct an accurate computer experiment even with elastic effects in the model. An increase in the number of bodies and elastic components provides a high complexity in MBD analysis, which may cause the numerical error in the computer experiment.

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.

In the present study, we introduced the implicit numerical integration methods under the MBD formulation [7] and evaluate the effect in comparative analysis between results of explicit and implicit solutions of numerical integrations [7],[8],[9]. In this sense, the advantages to utilize the implicit methods can be revealed in the computer experiment, which realizes the necessity and appropriate selection of the numerical method depending on the target system. It will contribute to an actual numerical solution not only for rigid-body mechanics but also for flexible multibody dynamics (fMBD), which embeds the finite element method (FEM) into the original description of MBD.

2. Methodology

2.1. Numerical error under MBD analysis

In MBD-based dynamic analysis, the coordinates and rotation angles of each mechanical element in the analytical system are obtained by solving the second-order ordinary differential equations based on the following differential algebraic equations for the generalized acceleration matrix $\ddot{q}$. The meanings of the characters in Eq. (1) are as Table 1.

$$\begin{bmatrix} M & \Phi_q^T \\ \Phi_q & 0 \end{bmatrix} \begin{bmatrix} \ddot{q} \\ \lambda \end{bmatrix} = \begin{bmatrix} Q^A \\ \gamma \end{bmatrix} \tag{1}$$

Table 1. The planning and control components.

M	Mass matrix
Φ_q	Jacobian matrix differentiated from constraint equation in generalized coordinates
$\ddot{q}$	Generalized acceleration matrix
λ	Lagrange multiplier
Q^A	Generalized force
γ	Acceleration equation

In computer analysis, this second-order ordinary differential equation can be solved by numerical integration methods, however errors accumulate at each analysis step. These numerical errors appear in the deviations of nodal coordinates connecting each element in the kinematic analysis, and the accuracy of the numerical integration methods can be evaluated by the distance of the nodal coordinates (Fig. 1).

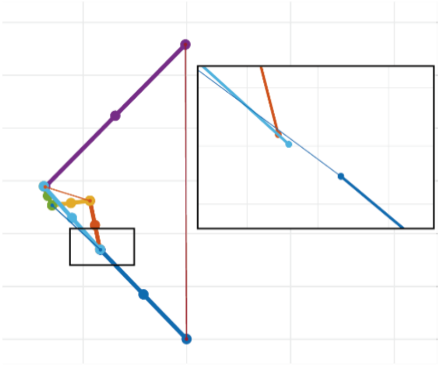
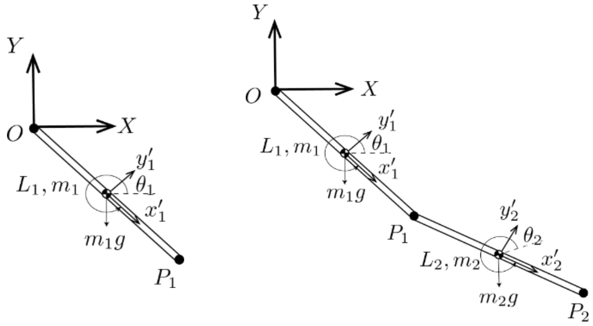


Fig.1 Displacements of the nodal coordinates caused by errors of numerical integration in knee linkage model [6].

2.2. *Numerical integration methods*

In order to compare and verify the analytical errors in the MBD analysis, we applied the numerical integration methods to the dynamic analysis with MBD of single and double pendulums (Fig. 2) to simply verify the errors.



a. Single pendulum b. Double pendulum

Fig. 2. Generalized coordinate systems of pendulums for MBD analysis.

For comparing errors caused by numerical integration methods, we applied Runge-Kutta Gill's method [10], the two-stage fourth order and the three-stage sixth order implicit Runge-Kutta (IRK) method. Gill's method is explicit fourth order Runge-Kutta method. The s-stage IRK method is described as the following equation [7],[8]. The coefficients *a*, *b* and *c* in Eq. (2) are due to the Butcher array (Table 2).

$$\begin{cases} k_i = f(x_0 + c_i h, y_0 + h \sum_{j=1}^{s} a_{ij} k_j) \\ y_{(x_0+h)} = y_0 + h \sum_{i=1}^{s} b_i k_i \end{cases} \tag{2}$$

Table 2. The Butcher array for IRK

c_1	a_{11}	$\cdots$	a_{1s}
$\vdots$	$\vdots$	$\ddots$	$\vdots$
c_s	a_{s1}	$\cdots$	a_{ss}
	b_1	$\cdots$	b_s

In MBD analysis, since the generalized acceleration matrix $\ddot{q}$ is obtained by solving a second-order ordinary differential equation as Eq. (3), the equations for obtaining the generalized coordinates matrix q and the generalized velocity matrix $\dot{q}$ with IRK are as Eq. (4) and Eq. (5).

$$\ddot{q} = f(t, q, \dot{q}) \tag{3}$$

$$\begin{cases} k_i = f(t_n + c_i h, q_n + h \sum_{j=1}^{s} a_{ij} l_j, \dot{q}_n + h \sum_{j=1}^{s} a_{ij} k_j) \\ l_i = \dot{q}_n + h \sum_{j=1}^{s} a_{ij} k_j \end{cases} \tag{4}$$

$$(i = 1, \cdots, s)$$

$$\begin{cases} y'_{n+1} = p_{n+1} = P_n + h \sum_{i=1}^{s} b_i k_i \\ y_{n+1} = y_n + h \sum_{i=1}^{s} b_i l_i \end{cases} \tag{5}$$

Within each integral computation step, the Newton-Raphson method is inserted to find solutions of the simultaneous equations to derive all coefficients k_i and l_i in Eq. (4), then q and $\dot{q}$ are obtained by Eq. (5). The Butcher arrays for 2stage 4order IRK (Table 3) and 3stage 6order IRK (Table 4) are shown as follows [8].

Table 3. The Butcher array for 2-4 IRK

$\frac{1}{2} - \frac{\sqrt{3}}{6}$	$\frac{1}{4}$	$\frac{1}{4} - \frac{\sqrt{3}}{6}$
$\frac{1}{2} + \frac{\sqrt{3}}{6}$	$\frac{1}{4} + \frac{\sqrt{3}}{6}$	$\frac{1}{4}$
	$\frac{1}{2}$	$\frac{1}{2}$

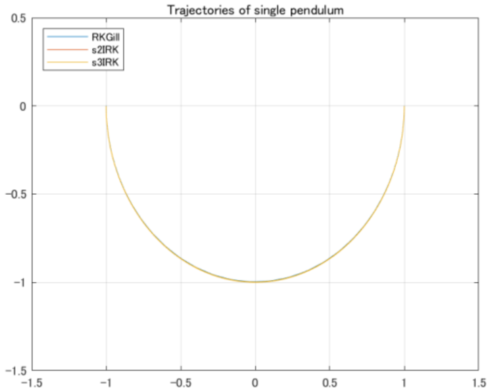
Table 4. The Butcher array for 3-6 IRK

$\frac{1}{2} - \frac{\sqrt{15}}{10}$	$\frac{5}{36}$	$\frac{2}{9} + \frac{\sqrt{15}}{15}$	$\frac{5}{36} + \frac{\sqrt{15}}{30}$
$\frac{1}{2}$	$\frac{5}{36} + \frac{\sqrt{15}}{24}$	$\frac{2}{9}$	$\frac{5}{36} - \frac{\sqrt{15}}{24}$
$\frac{1}{2} + \frac{\sqrt{15}}{10}$	$\frac{5}{36} + \frac{\sqrt{15}}{30}$	$\frac{2}{9} + \frac{\sqrt{15}}{15}$	$\frac{5}{36}$
	$\frac{5}{18}$	$\frac{4}{19}$	$\frac{5}{18}$

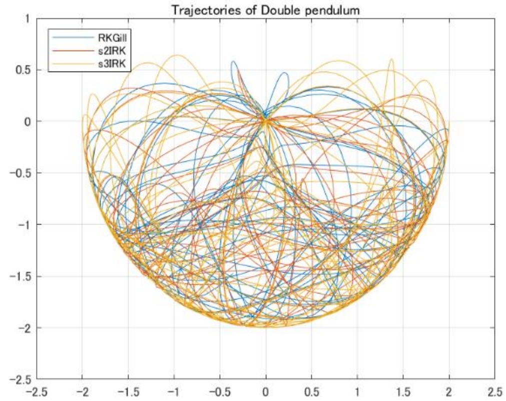
3. Results and Discussion

3.1. *Accuracy of numerical integration*

Each numerical integration method was implemented in the numerical computation of MBD analysis of single and double pendulums (time increment $h = 0.1$), and the calculation errors were verified. The differences in the trajectory by numerical integrations are shown in Fig. 3.

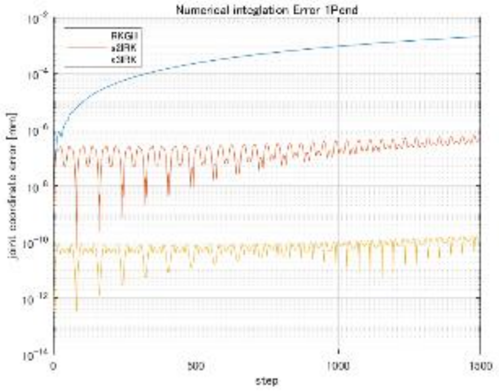


a. Single pendulum

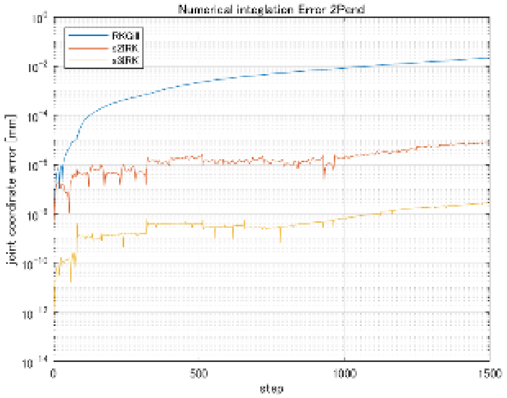


b. Double pendulum
Fig.3 Trajectories of pendulums

Fig. 4 shows the step-by-step transition of the joint coordinate error computed by numerical integrations.



a. Single pendulum



b. Double pendulum
Fig. 4. Comparison of coordinate errors by numerical integration method

According to these figures, the errors in the single and double pendulum are smaller for the implicit solution than for the explicit solution. It is also clear that the analytical error becomes smaller as the order of the numerical integration method increases. Therefore, in the analysis of a system consisting of multiple elements based on MBD, it is possible to evaluate the accuracy of numerical integral calculations from the errors in the joint coordinates of the mechanical elements constituting the system to be analyzed.

3.2. *Performance of the numerical integration on MBD analysis*

In numerical computation, the analytical time is important as analytical accuracy. Fig. 5 shows the changes in numerical computation time due to the expansion of the order of the implicit numerical integration and the number of variables in the MBD dynamics analysis of the pendulums. The number of computation steps is 1500.

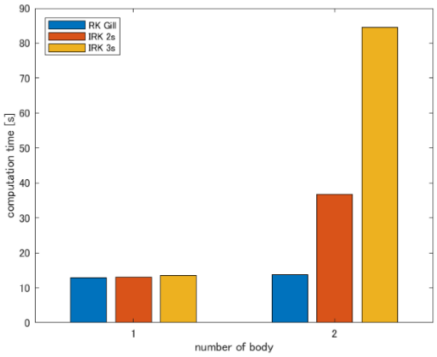


Fig. 5. The analysis times of numerical integrations

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.

4. Conclusion

The validation of the numerical integration in the pendulum model analysis with MBD shows methods to suppress the error due to the numerical integration of the second-order ordinary differential equations in the kinetic and dynamic analysis with MBD analysis.

In further analysis, these numerical integration methods can be adopted to the dynamics/inverse dynamics of knee joint mechanism and human walking analysis based on fMBD analysis, and comparisons of their analytical accuracy can be discussed. However, there are issues in implementing the higher order IRK to multielement mechanism analysis, such as improving the performance by streamlining the substitution process for symbolic matrix formulas in MATLAB.

Acknowledgements

This work was supported in part by JSPS KAKENHI (16H01616, 17H06383), Project on Regional Revitalization Through Advanced Robotics (Kyushu Institute of Technology/Kitakyushu city, Japan) and Kitakyushu Foundation for the Advancement of Industry, Science and Technology (FAIS). The authors gratefully acknowledge ARIZONO orthopedic supplies Co., Ltd. for assistance with the formulation of exoskeleton-type assistive devices in their products.

References

1. 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.
2. P. E. Nikravesh, Planar Multibody Dynamics: Formulation, Programming with MATLAB, and Applications, 2nd edn., CRC Press, Boca Raton, 2018.
3. 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.
4. 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.
5. 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.
6. S. Kasai, P. Dachkinov, K. Tanaka and H. Wagatsuma, A MBD-based Knee Link Model to Analyze Singular Postures to Find Necessary Restrictions Toward Personalized Knee Support Exoskeleton, Journal of Robotics, Networking and Artificial Life, Vol. 9(2); September (2022), pp. 128–135.
7. M. Iwamura, Introduction to Multibody Dynamics, Morikita Publishing, 2018.
8. E. Hairer, S. P. Norsett and G. Wanner, Solving Ordinary Differential Equations Ⅰ, Springer, 1993.
9. E. Hairer and G. Wanner, Solving Ordinary Differential Equations Ⅱ, Springer, 1996.
10. Wolfram Research, Inc., Runge-Kutta Gill's method, https://mathworld.wolfram.com/GillsMethod.html

Authors Introduction

Mr. Shintaro Kasai

He received his Bachelor's degree in Engineering in 2021 from the Faculty of Engineering, Kyushu Institute of technology in Japan. He is currently a master student in Kyushu Institute of Technology, Japan

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 IEEE.

A Basic Concept of the Nonlinear Oscillator-Based Hough Transform Implementation to Improve the Voting Procedure in the Scheme of Continuous Dual Spaces

Amarbold Purev
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu 808-0196, Japan

Hiroaki Wagatsuma
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu 808-0196, Japan
Email: purev.amarbold118@mail.kyutech.jp, waga@brain.kyutech.ac.jp

Abstract

Hough transform is a well-known algorithm to detect arbitrary lines in the image. The algorithm consists of two steps, mapping image edges to parameter space and voting to find a solution. The first step is executed in the continuous space, while the second relies on counting votes in discrete mesh, which increases susceptibility to noise. In the present study, we propose attractor dynamics in coupled nonlinear oscillators instead of discrete voting. In our computer experiment, image pixels were mapped to the parameter spaces and attracted to the crossing point of lines in the space. This approach may contribute to line detections in the image and other shape detections using a consistent mathematical formulation.

Keywords: Hough transform, line detection, nonlinear dynamics, coupled nonlinear oscillators, Kuramoto model

1. Introduction

The line detection from image data has the greatest need in computer vision [1][2][3], as the primitive component to form objects in the image [4]. The Hough Transform (HT) is the most frequent method to solve the problem, and the idea of the method is mathematically reasonable for the detection to project multiple line elements to the parametric space for generating a unique concentrating point. The projection of image data is treated between continuous spaces as Cartesian and polar coordinate systems; however, the finding of the concentration point is executed in a voting procedure in the discrete grids. In the comparison with the former process, the latter voting procedure causes a trade-off problem as the resolution-dependent noise removal effect. Apparently, high-resolution grids require a high computational cost. In the drawback of the method, another solution is expected to replace the latter procedure by using a mathematical treatment in the continuous space dynamics. Interestingly, convergence dynamics is known in the nonlinear dynamics, especially in attractor dynamics. According to the description of ordinary differential equations, those dynamics can be reconstructed in the computer experiment based on the numerical integration method, for instance, Euler method and Runge-Kutta method. If the dual space projection in HT and a natural convergence of moving points to represent a single line based on the nonlinear dynamics can be integrated, it provides a new method to solve the drawback of the original HT.

In the present paper, we focused on the mathematical solution of the straight-line detection, which can be derived from nonlinear oscillator dynamics providing synchronization phenomena in a limit cycle as well as the projection space of HT. If it is possible, moving points are concentrating automatically to detect the crossing points of lines in the projection space. Thus, the voting procedure is no longer necessary. In the section 2, the projection mechanism from Cartesian coordinate system to the polar coordinate system and its algorithm for computer experiments were reviewed, and a representative nonlinear oscillator dynamic was introduced for the replacement of the voting procedure in the section 3. In the section 4, the proposed system was formulated as the noise-robust line detection method, followed by results of the computer experiments in the section 5 and conclusion in the section 6.

2. Hough Transform Method

The Hough Transform (HT) is a method used in image analysis to detect shapes in an image [5]. The technique was introduced by Paul Hough in 1962 and has since been used in a variety of applications [6][7][8]. HT works by transforming the coordinates of edge points in an image from Cartesian coordinates to a more compact representation in parameter space [9].

$$\rho = x\cos(\theta) + y\sin(\theta) \tag{1}$$

Eq. (1) takes the coordinates of points in XY plane and projects into parameter space. Thus, every point in XY plane becomes a sine wave. Intersection of these sine waves give length value ρ, angle value θ, which is a line perpendicular to the line through edge points. This transformation is shown in Figure 1.

To find intersection of edge points in Figure 1(b) HT method creates an accumulation array with $R \times T$ size consisting of zero values. R is the number of rows for the ρ value, and T is the number of columns for θ value. For each value θ, the ρ value for each point is calculated, and the closest ρ value in the accumulation array is incremented by 1. This process is called the voting procedure of HT. After passing through all columns of the accumulation array, the cell with the highest point determines the characteristics of the perpendicular line. This transformation allows for the detection of shapes to be reduced to a peak detection problem in parameter space.

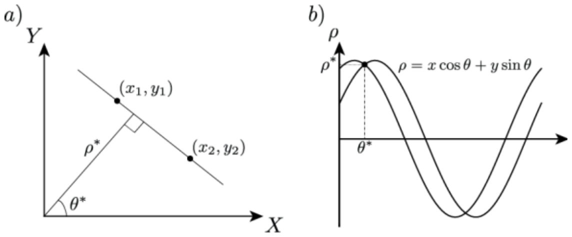


Fig. 1 Hough transformation. a) Representation of points $(\boldsymbol{x_1}, \boldsymbol{y_1})$ and $(\boldsymbol{x_2}, \boldsymbol{y_2})$, a line perpendicular to them with length $\boldsymbol{\rho}^*$ and angle $\boldsymbol{\theta}^*$ in image space. b) Transformed points as sine waves in parameter space.

One of the key advantages of the HT method is its simplicity, which allows for parallel processing of information. This makes the method robust to noise and allows for the detection of partially occluded or deformed shapes. Additionally, the method can be configured to detect a specific number of shapes in an image.

However, the simplicity of the HT method becomes a disadvantage when trying to detect shapes other than straight lines. The number of calculations increases as the number of parameters needed to describe the shape increases [10][11]. For example, detecting a circle in an image would require a three-dimensional accumulation array for the center coordinates of the circle (x, y) and its radius r. Researchers have attempted to improve the efficiency of the HT method through software implementations [12][13][14][15], hardware acceleration [16], and the use of neural networks [17].

The use of an accumulation array in HT also presents challenges, as the size of the array needs to be fine-tuned for accurate detection. If the array is too small, the method loses accuracy, and if it is too large, the computation becomes expensive. Algorithm 1 describes the computer program algorithm for traditional HT.

Algorithm 1. Classic HT for Line Detection

Data: (x, y) locations of edges in image space
Result: (ρ, θ) values of a cell with maximum vote
Initialization. Given number of rows R for ρ and number of columns T for θ, create accumulation array with zero values for N number of edges in the image space.

```
for $i = 1$ to $i_N$ do
    for $t = 1$ to $T$ do
        calculate $\rho_i = x_i\cos(\theta) + y_i\sin(\theta)$
        find closest $R$ row to $\rho_i$
        increment cell value by 1 with location $(R, t)$
    endfor
endfor
```

As shown in Algorithm 1, HT only detects lines that are infinite in length. This is because the HT represents lines using their parameters, which do not include information about the length of the line. Thus, the HT cannot detect lines of finite length.

3. Nonlinear Oscillator Dynamics

3.1. *Kuramoto model*

The Kuramoto model (KM) is a mathematical model used to study synchronization in oscillators [18]. It is commonly used in fields such as physics [19][20], biology [21], and chemistry, and has been applied in various contexts, including neurophysiology [22][23], distributed power generation [24][25], and secure communication [26][27]. In this paper, we propose to use the modified model of KM to classify edges in digital images based on their collinearity.

The general form of the KM is defined by the following equation

$$\dot{\varphi}_i = \omega_i + \frac{K}{N}\sum_{j=1}^{N}\sin(\theta_j - \theta_i). \tag{2}$$

Eq. (2) describes a system of N limit-cycle oscillators with natural frequencies ω_i, phases θ, and coupling constant K. The coupling between oscillators is based on the sine of their phase differences, and the time variation of phases follows first-order dynamics.

The use of the KM in image processing is rare because oscillators are dynamic objects, whereas edges in images are static. However, researchers in [28] have used it in color image segmentation, where they generated three oscillating curves corresponding to the pixel values of red, green, and blue for color images by a coupled network to produce superposition of oscillation for image pixels.

Nevertheless, we propose to use the attractor dynamics of coupled nonlinear oscillators in the KM to transform edge locations from discrete to continuous space. This allows for the classification of edges in parameter space using the KM, that will eliminate the need for an accumulation array in the HT method.

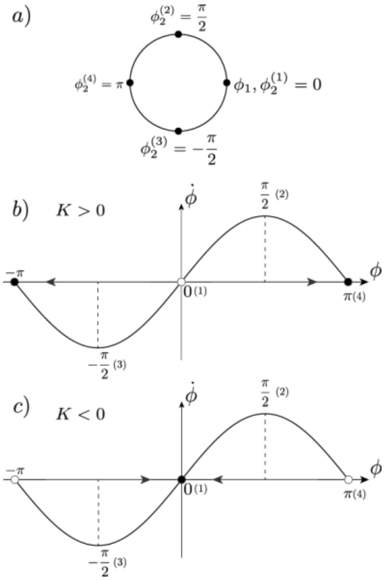


Fig. 2 Interaction between two nonlinear oscillators $\boldsymbol{\varphi_1}$and $\boldsymbol{\varphi_2}$. a) Initial positions. Behavior of the second oscillator b) $\boldsymbol{K > 0}$ c) $\boldsymbol{K < 0}$. Black dot represents stable point, white dot represents unstable point.

In order to understand how the KM can be used to detect lines in images, it is necessary to first analyze the behavior of simple systems. By studying their behavior, it will be possible to develop a model that can accurately classify edges in images based on their collinearity.

3.2. *Stability Analysis of two coupled Kuramoto oscillators*

Suppose that a system with two oscillators, $\dot{\varphi}_1$ and $\dot{\varphi}_2$, is given and behaves according to

$$\begin{cases}\dot{\varphi}_1 = 0\\ \dot{\varphi}_2 = K\sin(\varphi_2 - \varphi_1)\end{cases} \tag{3}$$

and subject to initial conditions $\varphi_i(0) = \varphi_i^0$, where φ is the phase of the ith oscillator, K is the coupling strength.

For simplification, we assumed $\dot{\varphi}_1 = 0$ in all cases, while concentrating on the dynamics of the second oscillator with initial values $\varphi_2^{(n)} = [0, \pi/2, -\pi/2, \pi], n = 1 \dots 4$, as shown in Figure 2.

We now give the definition of phase synchronization.

Definition 1. Let $\{\boldsymbol{\varphi_i(t)}\}_{\boldsymbol{i=1}}^{\boldsymbol{2}}$ be solution of our system. We say the oscillators converge to phase synchronization if

$$\lim_{t\to\infty}\left|\varphi_i(t) - \varphi_j(t)\right| = 0, \quad \text{for } i \neq j. \tag{4}$$

Let's see the behavior of our system given by Eq. (3) with different initial conditions

Example 1. $\varphi_1(0), \varphi_2(0) = 0$

It is easy to see that the corresponding solution is $\varphi_1(t) = 0$ and $\varphi_2(t) = 0$. Hence there is phase synchronization.

Example 2. $\varphi_1(0) = 0, \varphi_2(0) = \frac{\pi}{2}, K > 0$

$$\begin{aligned}\varphi_2(t) &= K\sin\left(\frac{\pi}{2} - 0\right)\\ &= K\sin\left(\frac{\pi}{2}\right)\\ &= K\end{aligned} \tag{5}$$

Since $K > 0, \varphi_2(t) \to \pi$ the second oscillator converges to the anti-phase.

Example 3. $\varphi_1(0) = 0, \varphi_2(0) = -\frac{\pi}{2}, K > 0$

$$\begin{aligned}\varphi_2(t) &= K\sin\left(-\frac{\pi}{2} - 0\right)\\ &= K\sin\left(-\frac{\pi}{2}\right)\\ &= -K.\end{aligned} \tag{6}$$

Since $\dot{\varphi}_2(t) < 0, \dot{\varphi}_1(t) \to -\pi$ the second oscillator converges to the anti-phase as well.

Example 4. $\varphi_1(0) = 0, \varphi_2(0) = \pi, K > 0$

$$\begin{aligned} \varphi_2(t) &= K\sin(\pi - 0) \\ &= K\sin(\pi) \\ &= 0 \end{aligned} \tag{7}$$

With the initial position for $\varphi_2(0) = \pi$ and $\varphi_2(t) = 0$, the second oscillator started at the anti-phase stable point of π and do not move as $t \to \infty$.

Even though two oscillators do converge in Example 1, the 0 point is unstable as shown in Examples 2-4. Therefore, the system given by Eq. (3) converges to anti-phase. However, if the coupling strength has negative value, after doing same calculations as the above examples, we can see that the two oscillators do converge.

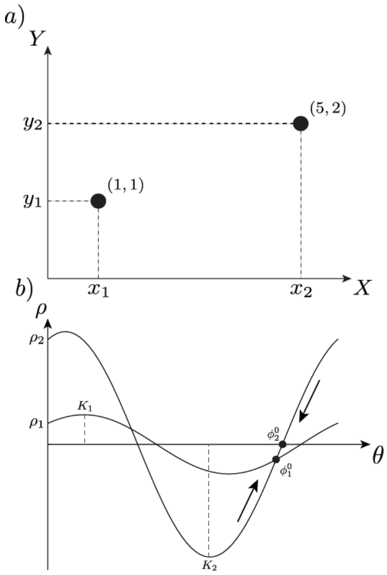


Fig. 3 Example of points (1,1) and (5,2). a) Image space. b) Hough space.

4. Proposed System for Noise-Robust Line Detection

4.1. *Hough Transform line detection for two points using coupled nonlinear oscillators*

The traditional HT method is described in Section 1.1 Instead of the traditional HT voting scheme, we propose replacing it with a nonlinear system of equations. Each point in the image space is treated as an individual nonlinear oscillator and coupled together with the following system of equations

$$\begin{cases} \dot{\varphi}_1 = 0 \\ \dot{\varphi}_2 = \dfrac{\rho_1 - \rho_2}{K^2} + (\varphi_1 - \varphi_2) \end{cases} \tag{8}$$

and subject to initial conditions $\varphi_i(0) = \varphi_i^0$, where φ is the phase of the ith oscillator, ρ is the Hough amplitude at the φ phase and K is the coupling strength based on the maximum amplitude of the ρ value.

To find the intersection of two points in the Hough space, ρ and θ values of each oscillator should converge. Similar to Eq. (3), we will set $\dot{\varphi}_1 = 0$ and concentrate on converging the second point to the first one. Which means, the oscillator $\dot{\varphi}_1$ will attract $\dot{\varphi}_2$ to itself.

Therefore, setting the initial phase of the first oscillator is crucial and we find its location analytically, where two sinusoidal lines intersect, which we can derive from the following assumption

$$\begin{gathered} \rho_1 = \rho_2 \\ x_1\cos\theta + y_1\sin\theta = x_2\cos\theta + y_2\sin\theta \\ \theta = \tan^{-1}(-(x_1 - x_2)/(y_1 - y_2). \end{gathered} \tag{9}$$

Accordingly, we set the initial phase of the first oscillator to $\varphi_1^0 = \tan^{-1}(-(x_1 - x_2)/(y_1 - y_2))$. For the initial phase of the second point, we are setting it at the fixed point of the Hough line, which means $\rho_2 = 0$, and with the same calculation as Eq. (9) the $\theta = \tan^{-1}(-x_2/y_2)$, and the starting position of the second oscillator $\varphi_2^0 = \tan^{-1}(-x_2/y_2)$. Coupling strength K is the maximum of the ρ value for each Hough line and equals to $K_i = x_i^2 + y_i^2$.

However, to avoid the anti-phase synchronization situation same as in Section 2.1 the sign changing rule must be applied.

Example 5 shows the verification of the above method.

Example 5. Two points, $p_1 = (x_1, y_1) = (1,1)$ and $p_2 = (x_2, y_2) = (5,2)$ are given in the image space, as shown in Figure 3. Initial values should be calculated to synchronize two oscillators for the coupling the same as Eq. (9). $\varphi_1^0 = \tan^{-1}(-(x_1 - x_2)/(y_1 - y_2)) = 4.9574$ rad, $\varphi_2^0 = \tan^{-1}(-x_2/y_2) = 5.0929$ rad, $K_2 = x_2^2 + y_2^2 = 29$.

From Figure 3, it is clear that φ_i^0 has 2 roots for each oscillator in 2π period. Therefore, depending on the location of the initial phase of the oscillator our proposed model should have the ability to change the sign. To first determine which direction the oscillator should move, the slope of the second Hough line should be calculated by the first derivative of ρ function

$$\frac{d\rho}{d\theta} = y\cos\theta - x\cos\theta. \tag{10}$$

In total there are four different cases are possible

(i)	$\varphi_1^0 > \varphi_2^0,\ \rho_2'(\varphi_2^0) > 0$
(ii)	$\varphi_1^0 < \varphi_2^0,\ \rho_2'(\varphi_2^0) < 0$
(iii)	$\varphi_1^0 > \varphi_2^0,\ \rho_2'(\varphi_2^0) < 0$
(iv)	$\varphi_1^0 < \varphi_2^0,\ \rho_2'(\varphi_2^0) > 0$

After analyzing the possibilities, the four different possibilities can be classified into two groups (i, ii) and (iii, iv), which results in the same behavior. Then Eq. (8) transforms into the following system of equations

$$\begin{cases} \dot{\varphi}_1 = 0 \\ \dot{\varphi}_2 = \dfrac{\rho_1 - \rho_2}{K^2} + (\varphi_1 - \varphi_2), & \text{if A or B} \\ \dot{\varphi}_2 = \dfrac{\rho_2 - \rho_1}{K^2} + (\varphi_1 - \varphi_2), & \text{if C or D} \end{cases} \tag{11}$$

where $\mathrm{A} = (\varphi_1^0 > \varphi_2^0, \rho_2'(\varphi_2^0) < 0)$, $\mathrm{B} = (\varphi_1^0 < \varphi_2^0, \rho_2'(\varphi_2^0) > 0)$, $\mathrm{C} = (\varphi_1^0 > \varphi_2^0, \rho_2'(\varphi_2^0) > 0)$, $\mathrm{D} = (\varphi_1^0 < \varphi_2^0, \rho_2'(\varphi_2^0) < 0)$.

The Eq. (11) is simulated by using the Algorithm 2 below. We used Runge-Kutta 4th order method (RK4) to calculate first order derivative in our computer simulation.

Algorithm 2. Classic HT for Line Detection

```
Data: $(x, y)$ locations of $i = 1, \dots, N$ edges in image space
Result: $(\rho, \theta)$ values that represent a line
Initialization. Calculate $\varphi_i^0$, $i = 1, \dots, N$
for $iter = 1$ to $maxL$ do
    for $rk = 1$ to 4 do
        if A or B then
            Calculate $\dot{\varphi}_i$ with $+$ sign
        else
            Calculate $\dot{\varphi}_i$ with $-$ sign
        endif
        RK4 approximation
    endfor
endfor
```

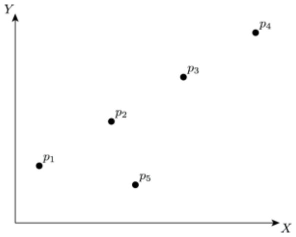


Fig. 4 Image space of five points

5. Results of Computer Experiments

In this section, we present the results of our proposed algorithm on multiple setups and provide an evaluation of correct detection as well as error.

Example 6 shows how the oscillators converge if a noncollinear edge exists in the image.

Example 6. $N = 5,\ p_1 = (1,3),\ p_2 = (4,5.33),\ p_3 = (7,7.76),\ p_4 = (10,10),\ p_5 = (5,2)$

In Example 6, five points are given, and the first point is the attractor, as shown in Figure 4. In the first part of the example p_1 is the attractor. Since, p_1 is part of the line through p_1 and p_4, oscillators converge into two groups. However, if the attractor is the outlier, such as p_5, oscillators converge into four groups. Notice that oscillators converged along the attractor's Hough line (dashed line in Figure 5).

We have also tested our method in the synthetic image with noise in two different scenarios, the same line with different noise and a different line with the same noise. The image has 100 collinear points and 1000 noise points in both scenarios with 20 samples each.

Table 1 and Table 2 shows the ratio of detected line points and errors. The threshold $\varepsilon = [10^{-3}, 10^{-4}, 10^{-5}]$ has been used to determine ratios. The distance D between (ρ, φ) values should be less than the threshold ε and was calculated by the following formula

$$D = \sqrt{\left(\varphi_i - \varphi_j\right)^2 + \left(\rho_i - \rho_j\right)^2} < \varepsilon, \tag{12}$$

where φ is the phase, ρ is the Hough value at the phase φ for the ith and jth oscillators, respectively.

The detected line points or cover ratio equals

$$CR = \frac{|LP| - |\mathrm{NofCover}|}{|LP|}, \tag{13}$$

where $LP = \{1,2,\dots,100\}$ set of collinear points, $\mathrm{NofCover} = |LP| - |DP|$ set difference of collinear points and all detected points.

The error ratio is calculated as follows

$$ER = \frac{|DP - LP|}{|pD|}, \tag{14}$$

where pD is set of all points, including collinear and noise.

From Table 1 and Table 2 we can see that the highest cover ratio CR and lowest error ratio ER occurs, when $\varepsilon < 10^{-4}$ at the iteration 125. Note that if the accuracy requirement increases, the time it takes to converge also increases ($\varepsilon = 10^{-5}, iter = 500$).

Figure 6 shows the convergence process in the parameter space for Table 1. Furthermore, Figure 7 shows the example of detected points in parameter and image spaces, respectively.

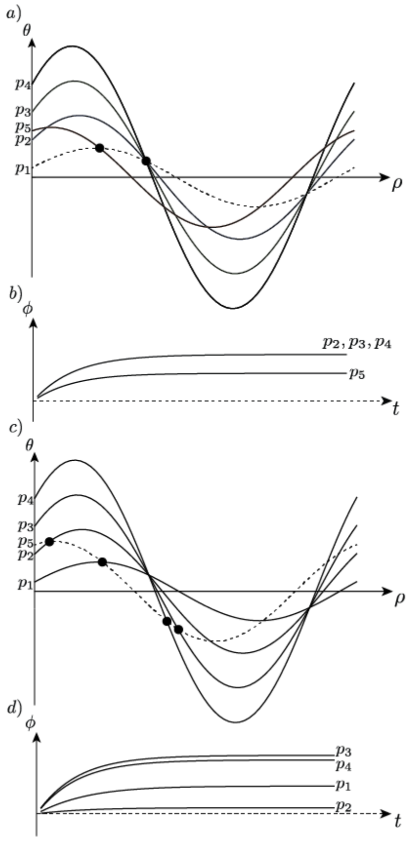


Fig. 5 Convergence of points with different attractors $\boldsymbol{p_1}$ – (a, b) and $\boldsymbol{p_5}$ – (c, d).

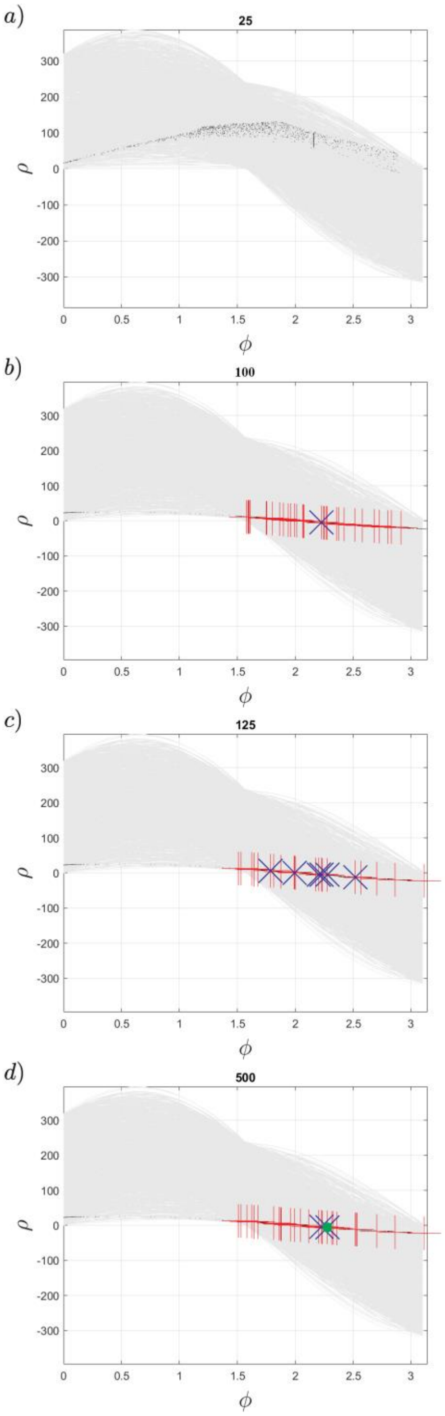


Fig. 6 A convergence process of moving points on projected lines in a) 25, b) 100, c) 125, d) 500 iteration periods. Colored marks represent $(\boldsymbol{\rho}, \boldsymbol{\theta})$ convergence points, when threshold value $\boldsymbol{\varepsilon = 10^{-3}}$(red "+"), $\boldsymbol{10^{-4}}$ (blue "X"), $\boldsymbol{10^{-5}}$ (green "dot").

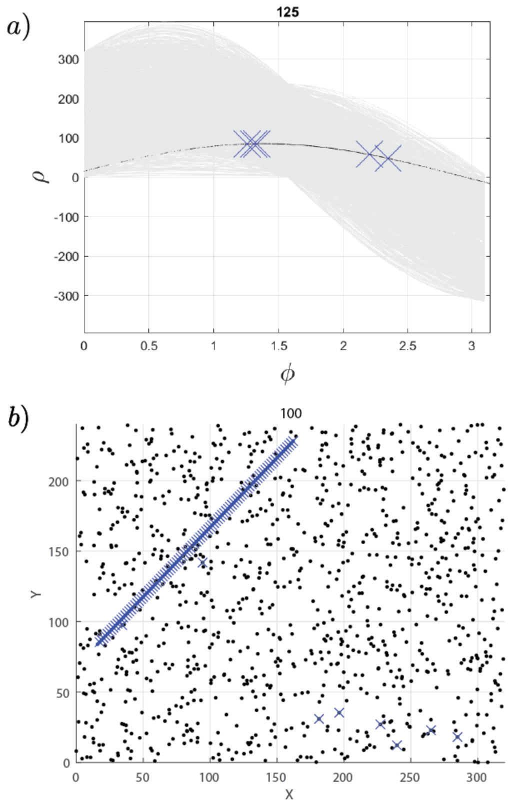


Fig. 7 Example of detected collinear points for same line different noise with $\boldsymbol{\varepsilon = 10^{-4}}$ and $\boldsymbol{iteration = 125}$. a) clustered points in parameter space b) detected points in image space (detected points are plotted with blue "X" sign)

6. Conclusion

Our method demonstrated the elimination of the use of the voting scheme of the classic HT by using coupled nonlinear oscillator dynamics.

The advantage of this method is that the user does not need to have any experience using the method since creating a specific-sized accumulation array is unnecessary. However, this method selects one point at a time, checks for colinear points, and clusters them into groups, which will have a slower implementation than the classic HT.

Further improvement of the method will improve HT's infinite line problem by clustering points into groups yielding finite lines.

Acknowledgements

This work was supported in part by JSPS KAKENHI (16H01616, 17H06383), Project on Regional Revitalization Through Advanced Robotics (Kyushu Institute of Technology/Kitakyushu city, Japan) and the New Energy and Industrial Technology Development Organization (NEDO).

Table 1. Average ratio of detected line points and errors for different line same noise at iterations (%).

ε	25		50		75		100		125		500	
	CR	ER	CR	ER	CR	ER	CR	ER	CR	ER	CR	ER
10^{-3}	0	0	0	0.09	5.20	1.07	**99.0**	**3.02**	99.00	3.70	99.00	3.72
$\mathbf{10^{-4}}$	0	0	0	0	0	0.01	4.95	0.11	**99.00**	**0.34**	99.00	0.45
10^{-5}	0	0	0	0	0	0	0	0	19.80	0.02	**99.00**	**0.05**

Table 2. Average ratio of detected line points and errors for same line different noise at iterations (%).

ε	25		50		75		100		125		500	
	CR	ER	CR	ER	CR	ER	CR	ER	CR	ER	CR	ER
10^{-3}	0	0	0.05	0.02	0.35	0.68	**99.00**	**3.43**	99.00	6.12	99.00	6.18
$\mathbf{10^{-4}}$	0	0	0.05	0	0	0	0.10	0.04	**99.00**	**0.52**	99.00	0.61
10^{-5}	0	0	0	0	0	0	0	0	0.05	0.02	**99.00**	**0.04**

References

1. B. K. P. Horn, "Robot Vision", MIT Press, 1986.
2. W. Chen, W. Wang, K. Wang, Z. Li, H. Li, S. Liu, "Lane departure warning systems and lane line detection methods based on image processing and semantic segmentation: A review", JTTE (English Edition), Vol. 7, No. 6, pp. 748-774, 2020.
3. Q. Luo, X. Fang, L. Liu, C. Yang, Y. Sun, "Automated Visual Defect Detection for Flat Steel Surface: A Survey", IEEE Transactions on Instrumentation and Measurement, Vol. 69, No. 3, pp. 626-644, 2020.
4. Q. -R. Wei, D. -Z. Feng, W. Zeng, J. -B Zheng, "Rapid line-extraction method for SAR images based on edge-field features", IEEE Geoscience and Remote Sensing Letters", Vol. 14, No. 10, pp. 1865-1869, 2017.
5. P. V. C. Hough, "Method and means for recognizing patterns", US Patent 3069654A, 1962.
6. T. K. Greeshma, S. Priya, C. Chaithanya, "A Survey on Line Detection Techniques using Different Types of Digital Images", IJARCET, Vol. 8, No. 5, pp. 162-168, 2019.
7. F. Tschopp, C. von Einem, A. Cramariuc, D. Hug, A. W. Palmer, R. Siegwart, M. Chli, J. Nieto, "Hough^{2}Map – Iterative Event-Based Hough Transform for High-Speed Railway Mapping", IEEE Robotics and Automation Letters, Vol. 6, No. 2, pp. 2745-2752, 2021.
8. C. Dalitz, T. Schramke, M. Jeltsch, "Iterative Hough Transform for Line Detection in 3D Point Clouds", Image Processing On Line, Vol. 7, pp. 184-196, 2017.
9. R. O. Duda, P. E. Hart, "Use of the Hough Transformation to Detect Lines and Curves in Pictures", Communications of the ACM, Vol. 15, No. 1, pp. 11-15, 1972.
10. D. H. Ballard, "Generalizing the Hough transform to detect arbitrary shapes, Pattern Recognition, Vol. 13, No. 2, pp. 111-122, 1981.
11. N. Kiryati, Y. Eldar, A. M. Bruckstein, "A Probabilistic Hough Transform", Pattern Recognition, Vol. 24, No. 4, pp. 303-316, 1991.
12. L. A. F. Fernandes, M. M. Oliveira, "Real-time line detection through an improved Hough Transform voting scheme", Pattern Recognition, Vol. 41, No. 1, pp. 299-314, 2008.
13. H. F. Li, D. Pao, R. Jayakumar, "Improvements and Systolic Implementation of the Hough Transformation for straight line detection", Pattern Recognition, Vol. 22, No.6, pp. 697-706, 1989.
14. S. Du, B. J. van Wyk, C. Tu, X. Zhang, "An improved Hough transform neighborhood map for straight line segments", IEEE Transactions on Image Processing, Vol. 19, No.3, pp. 573-585, 2010.
15. N. Aggarwal, W. C. Karl, "Line detection in images through regularized Hough transform", IEEE Transactions on Image Processing, Vol. 15, No. 3, pp. 582-591, 2006.
16. E. Hajjouji, S. Mars, Z. Asrih, A. E. Mourabit, "A novel FPGA implementation of Hough Transform for straight lane detection", JESTECH, Vol. 23, No. 2, pp. 274-280, 2020.
17. M. W. Spratling, "A neural implementation of the Hough transform and the advantages of explaining away", Image and Vision Computing, Vol. 52, No. 1, pp. 15-24, 2016.
18. J. A. Acebron, L. L. Bonilla, C. J. P. Vicente, F. Ritort, R. Spigler, "The Kuramoto model: A simple paradigm for synchronization phenomena", Reviews of Modern Physics, Vol. 77, No. 1, pp. 137-185, 2005.
19. M. Ignatov, M. Ziegler, M. Hansen, H. Kohlstedt, "Memristive stochastic plasticity enables mimicking of neural synchrony: Memristive circuit emulates an optical illusion", Science Advances, Vol. 3, No. 10, pp. e1700849, 2017.
20. M. Ignatov, M. Hansen, M. Ziegler, H. Kohlstedt, "Synchronization of two memristively coupled van der Pol oscillators", Applied Physics Letters, Vol. 108, No. 8, pp. 084105, 2016.
21. H. Mizuhara, Y. Yamaguchi, "Human cortical circuits for central executive function emerge by theta phase synchronization", NeuroImage, Vol. 36, No. 1, pp. 232-244, 2007.
22. T. Cattai, S. Colonnese, M. C. Corsi, D. S. Bassett, G. Scarano, F. D. V. Fallani, "Phase/amplitude synchronization of brain signals during motor imagery BCI tasks", arXiv:1912.02745 [q-bio.NC], 2019.
23. V. Rohr, R. Berner, E. L. Lameu, O. V. Popovych, S. Yanchuk, "Frequency cluster formation and slow oscillations in neural populations with plasticity", PLoS ONE, Vol. 14, No. 11, pp. e0225094, 2019.
24. P. A. Arinushkin, T. E. Vadivasova, "Nonlinear damping effects in a simplified power grid model based on coupled Kuramoto-like oscillators with inertia", Chaos, Solitons & Fractals, Vol. 152, No. 1, pp. 111343, 2021.
25. Y. Guo, D. Zhang, Z. Li, Q. Wang, D. Yu, "Overviews on the applications of the Kuramoto model in modern power system analysis, International Journal of Electrical Power & Energy Systems, Vol. 129, pp. 106804, 2021.
26. A. Argyris, D. Syvridis, L. Larger, V. A. Lodi, P. Colet, I. Fischer, J. G. Ojalvo, C. R. Mirasso, L. Pesquera, K. A. Shore, "Chaos-based communications at high bit rates using commercial fibre-optic links", Nature, Vol. 438, No. 7066, pp. 343-346, 2005.
27. P. Feketa, A. Schaum, T. Meurer, D. Michaelis, K. Ochs, "Synchronization of nonlinearly coupled networks of Chua oscillators", IFAC-PapersOnLine, Vol. 52, No. 16, pp. 628-633, 2019.
28. X. Lu, Y. Qiao, X. Chen, J. Miao, L. Duan, "Color Image Segmentation Based on Modified Kuramoto Model", Procedia Computer Science, Vol. 88, No. 1, pp. 245-258, 2016.

Authors Introduction

Mr. Amarbold Purev

He received his Master's degree from the School of Mechanical Engineering and Transportation, Mongolian University of Science and Technology, Mongolia in 2019. He is currently a Doctoral course student in Kyushu Institute of Technology, Japan.

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 Professor of the Department of Human Intelligence Systems. His research interests include nonlinear dynamics and robotics. He is a member of IEEE.

Survey on Harness Design for CubeSats: Understanding the Constraints of CubeSats Design and Toward an Optical Wireless Bus for CubeSats

Masahiro Tokumitsu
Division of Computer Science and Engineering, Department of Integrated Engineering, National Institute of Technology,
Yonago College, Japan, 4448 Hikonacho, Yonago, Tottori 683-0854, Japan

Masatoshi Tsuji
Department of Electrical and Computer Engineering, National Institute of Technology,
Kagawa College, Japan, 355 Chokushicho, Takamatsu, Kagawa 761-8058, Japan

Jun Nakaya
Department of Electronic Robot Engineering, Faculty of Engineering, Aichi University of Technology,
50-2 Manori, Nishihasamacho, Gamagori 443-0047, Japan
E-mail: tokumitsu@yonago-k.ac.jp, mtsuji@t.kagawa-nct.ac.jp, nakaya-jun@aut.ac.jp
www.yonago-k.ac.jp

Abstract

CubeSats of nanosatellites have attracted the attention of space scientists and engineers seeking to observe the space environment and develop innovative technologies in space engineering. The CubeSat is a class of miniaturized satellites with a form factor based on a 10 cm cube. However, the dimensional constraints of CubeSats restrict the embedding of relatively large mission devices, such as attitude control systems, into the satellites. Moreover, the harness used to transfer the data and supply power to the mission devices also occupies physical space to embed the mission devices. Therefore, this research surveys earlier studies on nanosatellite harness design. In addition, we consider the possibility of an optical wireless harness for the satellite bus system to achieve a more effective and reliable design for the CubeSats.

Keywords:

1. Introduction

Space developments recently have been gathering attention to provide services such as worldwide wireless communication, launch vehicles, and data analysis of remote sensing by artificial intelligence. The satellites are categorized by mass of the satellites, such as large satellites, microsatellites, and nanosatellites [1]. The category for microsatellites and nanosatellites is chosen actively for providing satellite-based services to construct their service systems because both kinds of satellites contribute to the development costs of the satellites and launch costs to orbits. Especially the small satellites (~180 kg) and their onboard devices have attracted space engineers, space researchers, and the space market because the micro/nanosatellites reduce the development costs and development period [2].

Moreover, universities and venture companies develop and operate CubeSats of nanosatellites for innovative technology demonstration and space or earth observation. The CubeSat is a miniaturized satellite with a 10cm cube (The basic unit size (1 Unit) of the CubeSat is a cubical shape with 10 cm × 10 cm × 10 cm. The physical size of the CubeSats depends on the mission requirements, and the total unit size varies from 1U to 12U. Fig. 1 shows the typical configurations of the CubeSats. The 1U, 2U, and 3U CubeSat are implemented as a single string of 1U

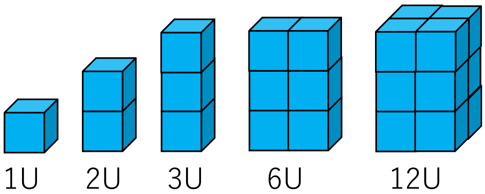


Fig. 1 Typical configurations of CubeSats

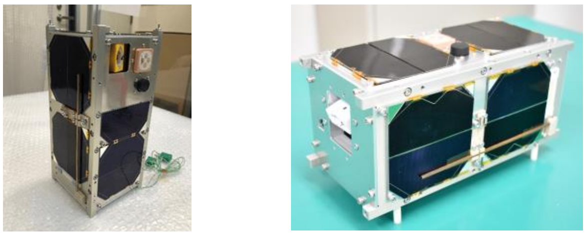

(a) KOSEN-1 (b) KOSEN-2

Fig. 2 Flight models of KOSEN-1 and KOSEN-2

CubeSat. On the other hand, 6U CubeSat is implemented as double lines of 3U Cubesat.

The physical size of the CubeSats is small compared with conventional satellites such as communication satellites and weather satellites [2]. However, many universities, research institutions, and venture companies attempt new and challenging technologies such as attitude control systems, inter-satellite communication, and space tether. Although launch services into space are necessary for the satellites, the major launch vehicles such as Epsilon Launch Vehicle (Japan Aerospace Exploration Agency, JAXA) [3], Polar Satellite Launch Vehicle (the Indian Space Research Organization, ISRO) [4], and Vega [5] and Ariane [6] (European Space Agency, ESA). The CubeSats usually are onboard as piggyback satellites, and the multiple CubeSats are onboard in a single launch. The PLSV-C37 by ISRO successfully launched 96 CubeSats in a single launch (The flight of PLSV-C37 carried 104 satellites in total in a single launch) [7], [8].

We belong to the KOSEN space collaborative group and have attempted CubeSat development twice: KOSEN-1 and KOSEN-2. Unfortunately, we faced various problems caused by the wire harnesses through two satellite developments and the function tests. The wire harnesses are used to connect electrical components to supply power and transfer control signals and mission data. Therefore, we considered that we needed to devise a different approach for harnessing to assemble the satellites.

In this paper, we survey on the satellite harness designs to consider the new different approaches to building the CubeSat for achieving highly reliable satellites. We also propose a new approach to constructing CubeSats' computer system using optical wireless communication technology. We consider some issues for implementing the optical wireless bus for the CubeSats. The remainder of this paper is as follows: Section 2 explains lessons learned from satellite developments of KOSEN-1 and KOSEN-2; Section 3 describes reviews of constraints of satellite design on CubeSats and safety regulations for satellite onboard; In Section 5, we explain the results of the survey on wired and wireless harness for satellites; In Section 5, we propose the optical wireless bus for the CubeSats and explain the basic concept of the proposed bus system; In Section 6, we conclude the survey on the CubeSats' harness and introduction to the optical wireless bus.

2. Lessons learned from satellite developments of KOSEN-1 and KOSEN-2

We have attempted to develop the two CubeSats: KOSEN-1 [9] and KOSEN-2 [10]. Fig. 2 shows flight models for launch vehicles onboard. The developments and technology demonstrations in space for KOSEN-1 and KOSEN-2 are selected for Innovative Satellite Technology Demonstration-2 (KOSEN-1) [9] and the same Demonstration-3 (KOSEN-2) [10] by JAXA. KOSEN-1 aimed to demonstrate an antenna expansion technology for Jupiter’s radio observation. KOSEN-1 was launched by JAXA’s Epsilon Launch Vehicle No. 5 in November 2021. KOSEN-2 aimed to demonstrate data transmission for sea crustal data from a marine buoy to KOSEN-2 at low orbit. KOSEN-2 was attempted to launch into space by JAXA's Epsilon Launch Vehicle No. 6 in October 2022. However, the launch of the Epsilon Launch Vehicle No. 6 failed the flight by a second motor trouble during the flight [10].

We faced various problems caused by harnesses through satellite developments and function tests. The satellites consist of onboard computers, batteries, chassis, harnesses, and mission devices. The main computers and sub-microcontrollers are interconnected and control the mission devices through electronic signals and electrical wires. The onboard computer boards, mission devices, and batteries are installed into the chassis of the CubeSats.

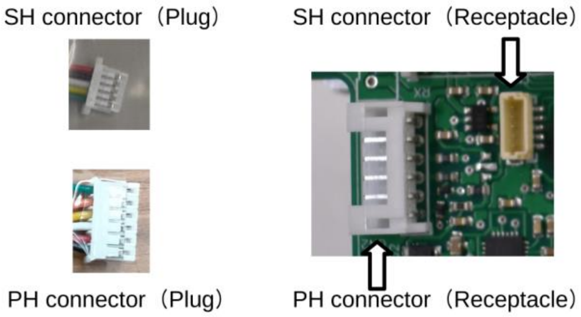
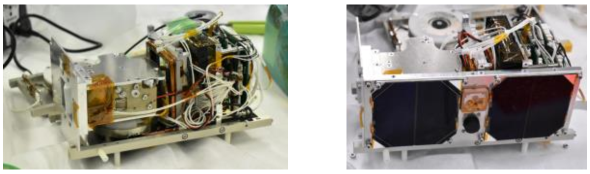


Fig. 3 Typical example of connectors and receptacles (J.S.T.Mfg.Co.,Ltd.)


(a) Before the body structure panel installation

(b) After the body structure panel installation

Fig. 4 The harness installation of KOSEN-2 flight model.

Fig. 3 shows a typical example of connectors used to interconnect the satellite mission devices.

The mission devices, electronic circuit boards, and batteries are installed into the chassis. The various kinds of harnesses are used for interconnecting the mission devices. Fig. 4(a) shows the actual harness installation of the KOSEN-2 flight model before the body structure panel installation. The workers need to be careful about damaging harnesses while installing the body structure panels and mounting the mission devices. Fig. 4(b) shows KOSEN-2 after the body structure panel installation. Eventually, the six panels are mounted around the chassis of the satellite.

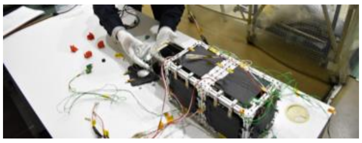
We recognized that some electrical mission devices of KOSEN-1 and KOSEN-2 did not work properly during the functional tests of the satellites. We found that the connectors’ engagement between plugs and receptacles made the contact failures. Eventually, the mission devices worked adequately after replacing the connectors with new ones. However, we lost the effort to solve the contact failures of the harnesses during the satellite production. Therefore, we conceived that we needed to devise a different approach for harnessing to assemble the CubeSats more effectively and accurately.

3. Constraints of satellite design for CubeSats

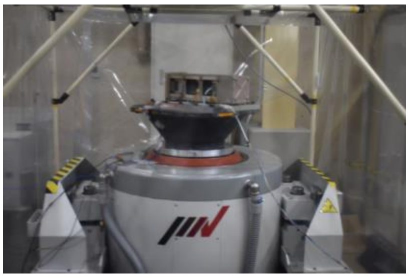
The satellites for flight models onboard to launch vehicles must satisfy various requirements such as the budget for the satellite developments, mission requirements, and safety regulations for the launch vehicles. Epsilon Launch Vehicles User’s Manual [11] is an example document on safety requirements for onboard satellites. The design standard document for the CubeSats is available [12] and many user manuals of the launch vehicles follow the CubeSat Design Specification. For the CubeSats, the vacant space for the main mission devices and their related devices occupies the internal space for the physical size. However, the physical size of the CubeSats and the various kinds of mission devices are embedded into the CubeSats in high density. The essential components of the satellites, such as onboard-computer boards, batteries, and wire harnesses, are essential devices for satellite implementation.

Although the physical size and the mass of the CubeSats are small, the CubeSats need to pass a safety regulation for launch vehicles. The launch vehicle service providers require safety inspections for the satellites. The safety inspections of the launch vehicles aim to prevent CubeSats' malfunction during the rockets' flights. The typical example of the CubeSats' malfunction during the flight is accidental radio emissions because the erroneous radio emissions interfere with monitoring and controlling the rocket's flights.

The designs and test results of the CubeSats need to satisfy the safety inspections for onboard launch vehicles. The tests for flight models for the launch vehicles involve various tests such as mass property tests, vibration tests, shock tests, and thermal vacuum tests. Fig. 5 shows the typical example of proto flight tests by KOSEN-2. Interface control documents define the test conditions for boarding the launch vehicles. Fig. 5(a) and Fig. 5(b) show the thermal vacuum test. The thermal vacuum test mainly aims to inspect the functions of the satellite under the vacuum state and the heat environment in space. Fig. 5(a) shows the setup for the thermal vacuum test, and the satellite is installed into the special enclosure that heats the satellite with manual control or programmed control. The satellites are installed and tested in the thermal vacuum chamber (Fig. 5(b)). Fig. 5(c) shows the vibration tests for the CubeSats. The vibration tests inspect the


(a) Thermal vacuum test (Setup)

(b) Thermal vacuum test (Testing)


(c) Vibration test

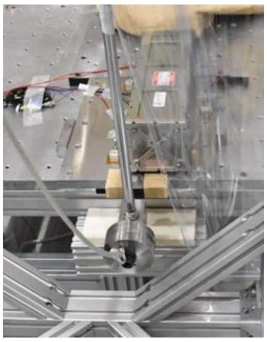

(d) Shock test

Fig. 5 Typical example of the proto flight test (KOSEN-2)

satellites under the vibration environments during the flights of the launch vehicles. The vibration tests simulate environments that assume various kinds of vibration caused by launch vehicles during the flights. Fig. 5(d) shows the impact tests for the satellites. The impact test aims to inspect whether the satellites endure the shocks for a fairing separation during the rocket's flight. In satellite developments, test conditions are defined in the rocket interface control documents. Therefore, the satellite development teams also need to consider the CubeSats' designs to pass the safety inspection regulations and their missions [13], [14].

4. Related work on satellite design for CubeSats

Space engineers and space scientists have studied extensively and deeply the design approaches of the CubeSats. However, the satellite developers need to devise the designs, assembly, and tests of the CubeSats because of the many design constraints require for the satellites.

The dimensions of the CubeSats are standardized by the design specification documents [12]. Therefore, the components such as electronic circuit boards also have modularity and compatibility with other components. The various components are available in the website shop [15]. Electronic circuit boards, such as onboard computers [16] and electrical power systems [17], are carefully considered for interconnectivity between the boards and stackable structures. The stackable modularity for the electronic circuit boards is a practical approach for the CubeSat design for high modularity. The PC/104 standards are used for stacking the electrical boards for the CubeSats [18]. Although the PC/104 standards are practical for interconnection among the electrical boards, the PC/104 socket is significant for the size of the CubeSats. The pins for the PC/104 sockets increase as the number of mission devices increases installed into the satellite. Using CPLDs for interconnecting the mission devices successfully reduced the wiring among electronic circuit boards [19]. The study [19] used CPLDs for routing the data and control signals for the mission devices. Using the CPLD approach and stackable electrical boards reduces the wire harness’s mass and space.

For the reduction of the satellite’s mass, the optical fibers are used for the signal transfer instead of the electrical wires [20]. The study [20] considered the optical fiber connectors and cables for spacecraft. Optical fibers overcome the data rate limitation and wire harness’s mass. The proposed optical fiber was examined in the tests of the engineering model, and the test was passed without performance degradation and failures of the fibers.

Using PC/104 standards and optical fibers for the CubeSats would be the candidate solutions for overcoming the constraints of the CubeSat design. However, satellite engineers and researchers have also studied the other approach to the wireless harness. Although the related studies aim to wireless harness technologies for ICT equipment [21]. The study [21] investigates microwave radio propagation in the ICT equipment to design the wireless harness. The study is the first step in constructing the ICT equipment by connecting the internal devices with the wireless harness. The study analyzed microwave radio propagation for ISM 2.4 GHz band. The communication for ICT equipment is up to a couple of meters at most. The study reported that the proposed modeling technique using a frequency-

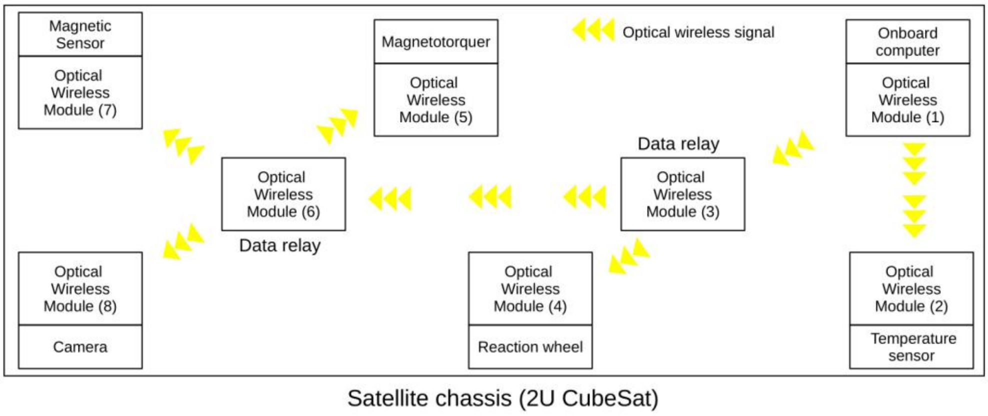

Magnetic Sensor
Optical Wireless Module (7)
Magnetotorquer
Optical Wireless Module (5)
Optical wireless signal
Onboard computer
Optical Wireless Module (1)
Data relay
Optical Wireless Module (6)
Data relay
Optical Wireless Module (3)
Optical Wireless Module (8)
Camera
Optical Wireless Module (4)
Reaction wheel
Optical Wireless Module (2)
Temperature sensor
Satellite chassis (2U CubeSat)

Fig. 6 Overview of the optical wireless bus. The numbers inside of the parentheses indicates module identifiers.

dependent path loss exponent expressing the near- and far-field propagation succeeded in extracting attenuation factors for the frequency and the propagation distance from the measured data.

The related study [22] also attempted constructing the wireless harness with ZigBee technology. The ZigBee is a technology and standard for constructing wireless sensor networks, and the study applied the ZigBee to a vehicle. The study reported viable technology for implementing an intra-car wireless sensor network. Two related studies reported that the wireless harness with IEEE 802.15.4 or ZigBee reduced the power requirements [23] and the satellites' power consumption [24].

The related study [25], [26], [27], [28] for the 3U CubeSat also attempted to construct the wireless harness called SKITH (SKIpTheHarness) of the InnoCube project. In the InnoCube project, Skith adopted the 2.4 GHz band for communication among the front-end modules. The Skith supports a data rate of up to 1 Mbps with GFSK modulation and transmission power up to 19 dBm.

Other related studies [29], [30] introduced the Ultra Wide Band (UWB) technology for substituting the wire harness to the wireless harness. The UWB is a personal wireless network technology for short-range communication (≤ 10 m). In the UWB technology, the modes communicate with very high-speed data rate and using wideband frequency (≥ 500MHz), compared with other related technologies such as ZigBee and Bluetooth. Furthermore, for the communication among the modes, the advantages of the UWB are low power, robustness for multipath fading, and reliability for communication errors. Therefore, earlier studies have investigated the applicability of spacecraft and vehicles for intra-interconnection for electrical devices.

The studies [31], [32] adopted a different approach to implementing the wireless harness for the satellites. The studies [31], [32] attempted to construct the wireless harness by optical wireless communication. There are various kinds of advantages of optical wireless communication for the wireless harness against radio wave communication. The optical wireless harness can avoid Electromagnetic Compatibility (EMC) in communication intra-spacecrafts and implement simplified circuits compared with RF circuits. Furthermore, the satellite development need not have RF licenses for optical wireless communication. The RF emissions during the launch vehicle's flight affect to monitor and control of the rocket. The study [31] described the development history of ten years on the optical wireless harness for spacecraft. However, in-orbit experiments on the optical wireless harness [31], [32] have still been challenged.

5. Optical wireless bus for CubeSats

5.1. *System overview of optical wireless bus*

In the previous section, we reviewed various papers on wireless harnesses for spacecraft and IoT equipment. Some studies on the wireless harness technology by UWB or ISM 2.4 GHz band demonstrated the feasibility of constructing the satellite bus system. However, the use of the wireless harness by radio waves usually are restricted by RF licenses and rocket interface control documents. The optical wireless communication for the mission devices would be the candidate for constructing the internal bus system for spacecraft. Optical wireless communication is available for spacecraft without RF licenses. Furthermore, optical wireless communication can avoid the restriction on erroneous RF emissions in launch vehicles' flights.

We propose the optical wireless bus based on the optical wireless harness to construct the bus system for the CubeSats based on the reviews of the earlier studies. The entire bus of the satellite consists of sensors and mission devices distributed in the satellite chassis. Fig. 6 shows the overview of the proposed optical wireless bus. The proposed optical wireless bus aims to interconnect the mission devices through optical communication, and each device transmits the mission and control data mutually.

Each mission device or sensor connects directly with an optical wireless module (OWM), and the OWMs control the devices. The transmission and reception of the optical signals of the OWMs would be realized with LEDs (Transmission) and photodiodes (Reception). Each OWM exchanges the mission and control data through optical wireless signals. The OWMs for the data relay act as router nodes for relaying the data. The OWMs ((3) and (6)) for the data relay receive the data from other OWMs, and the OWMs ((3) and (6)) relay the data to other OWMs ((7) and (8)).

5.2. *Implementation of optical wireless bus*

There are some technical issues with achieving the optical wireless bus. For the first issue, we need to design communication links among the OWMs for applying optical wireless communication to the CubeSats. However, the lighting LEDs of the OWMs consume the power of CubeSats for transmitting the signals. Furthermore, we need to consider the energy balance for power generation and power consumption for the operation of the CubeSats. Therefore, we investigate the communication links among the OWMs to realize the wireless optical bus.

For the second issue, we need to design the OWMs and implement the OWMSs as a module that can install into the CubeSats. We can implement the OWMs as printed circuit boards (PCB), including the transmission and reception circuits. Furthermore, the PCBs of the OWMs need to include microcontrollers for communication and controlling the mission devices or sensors. The transmission and reception circuits would be more straightforward than RF circuits because the optical wireless circuits for the proposed bus would not require high-frequency oscillation circuits. However, we need to consider the microcontrollers for the OWMs carefully because of the processing speed, storage space for programs, and input pins and output pins for analog circuits or digital circuits for mission devices and sensors. Therefore, selecting the microcontrollers for the OWMs depends on these requirements.

For the third issue, we need to design and standardize the communication protocol for the OWMs. The interconnectivity of the OWMs plays an essential role in the development of CubeSats. We need to consider various perspectives for the communication protocol, such as a physical layer, the data link layer, and the network layer. We need to define the modulation methods and data speed rate for the physical layer. We need to define the data frame, including error correction and header fields. In the end, we need to implement the communication protocol as a program for the microcontroller.

5.3. *In-orbit experiment plan and satellite development plan*

We develop the optical wireless bus and need to demonstrate the proposed bus in orbit. We have already obtained the opportunity for the in-orbit experiment as a CubeSat of KOSEN-3 (2U). KOSEN-3 is selected as Innovative Satellite Technology Demonstration-4 by JAXA. The primary mission of KOSEN-3 aims to demonstrate the pulsed plasma thruster system to maintain the orbit. KOSEN-3 will be launched by Epsilon S Launch Vehicle in the 2024 fiscal year. For the in-orbit experiment, the KOSEN-3 project is developing the satellite for the flight model. We consider and develop the engineering (EM) model of KOSEN-3 to verify the satellite design that complies with the safety requirements of the JAXA's Epsilon Launch Vehicle.

The optical wireless bus is still under development in the KOSEN-3 project. The optical wireless bus would contribute to the satellite itself and the development process of the satellite. The optical wireless bus would reduce the wired harness to interconnect the mission devices and sensors. However, the satellites require the harness for the power supply at the current development progress. The contactless power transmission technology can be the candidate to reduce the wired harness for power supply to the internal devices of the satellites. Other advantages of the optical wireless bus also can reduce the mass of the satellite by reduction of the wired harness.

Furthermore, the connection reliability of the connectors for interconnecting the mission devices or sensors would be improved because the mission devices or sensors exchange only the signals of the mission data or control data by wireless communication. Finally, the optical wireless bus enables us to change the design of the fitting out because of the mount position of the mission devices or sensors without changing the harnesses. The satellite development team can attempt various designs to fit out of the internal devices in the development process. The optical wireless bus would contribute to the rapid development of the satellite compared with the conventional development by the wired harness.

6. Conclusions

We reviewed the earlier studies on the harness for the CubeSats and discussed the constraints of the CubeSats' designs. Therefore, we proposed the optical wireless bus for the satellite bus system. The optical wireless bus comprises the optical wireless modules that communicate with other modules. The optical wireless modules are connected directly to the mission devices or sensors to transmit the mission data or control data for the satellite operation. In addition, we considered the possibility of an optical wireless bus for the satellite bus system to achieve a more effective and reliable design for the CubeSats. We explained that the optical wireless modules could be implemented with simplified circuits and microcontrollers compared with high-radio frequency circuits. For future work, we will implement the optical wireless bus for KOSEN-3 and demonstrate the usefulness of the optical wireless bus in orbit.

Acknowledgments

We acknowledge JAXA for their support in developing KOSEN-1 and KOSEN-2 CubeSats and sending KOSEN-1 to space. This work was supported by the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT) through the FY 2022 Aerospace Science and Technology Promotion Fund, Aerospace Human Resource Development Program, "Next Generation KOSEN Space Human Resource Development through Continuous Nano-Satellite Development and Operation" (Principal Investigator: National Institute of Technology, Niihama College). This work was partially supported by JSPS KAKENHI Grant Number 20K11965.

References

1. E. Mabrouk, "What are SmallSats and CubeSats?", August 7, 2017.
2. SpaceWorks Enterprise, Inc., "2020 Nano/Microsaatellite market forecat, 10th edition", 2020.
3. The Japan Aerospace Exploration Agency, "Epsilon Launch Vehicle".
4. The Indian Space Research Organization, "PLSV".
5. The European Space Agency, "Vega".
6. The European Space Agency, "Ariane".
7. The Indian Space Research Organization, "PSLV-C37 Successfully Launches 104 Satellites in a Single Flight", 2017.
8. S. Mathewson, "India Launches Record-Breaking 104 Satellites on Single Rocket", 2017.
9. The Japan Aerospace Exploration Agency, "Innovative Satellite Technology Demonstration Program 2".
10. The Japan Aerospace Exploration Agency, "Innovative Satellite Technology Demonstration Program 3".
11. The Japan Aerospace Exploration Agency, "Epsilon Launch Vehicle User's Manual".
12. The CubeSat Program, Cal Poly SLO, "CubeSat Design Standard Rev. 14.1".
13. National Aeronautics and Space Administration, "CubeSat 101: Basic Concepts and Processes for First-Time CubeSat Developers".
14. University Space Engineering Consortium (UNISEC), "Mission Assurance Handbook for the University-built Lean Satellite".
15. CubeSatShop, "CubeSatShop.com - One-stop webshop for CubeSats & Nanosats".
16. ISISPACE, "ISIS On Board Computer".
17. ISISPACE, "Modular Electrical Power System".
18. PC/104 Consortium, "Application of General Industrial Standards for Building Spaceborne Computer Systems".
19. S. Marloun, T. Yamauchi, N. C. Orger, Y. Otani, and M. Cho, "Scalable and Configurable Electrical Interface Board for Bus System Development of Different CubeSat Platforms", Applied Sciences 12, no. 18: 8964.
20. J. Blasco, C. Vernich, S. D. L. Rosa, M. Á. Esteban, M. J. Méndez, "Optoelectronic link for optical satellite harnessing substitution in space communications", 2018,
21. M. Ohira, T. Umaba, S. Kitazawa, H. Ban and M. Ueba, "Experimental Characterization of Microwave Radio Propagation in ICT Equipment for Wireless Harness Communications", in IEEE Transactions on Antennas and Propagation, vol. 59, no. 12, pp. 4757-4765, Dec. 2011.
22. H. -M. Tsai, O. K. Tonguz, C. Saraydar, T. Talty, M. Ames and A. Macdonald, "Zigbee-based intra-car wireless sensor

networks: a case study,” in IEEE Wireless Communications, vol. 14, no. 6, pp. 67-77, 2007.

23. N. Herscovici, C. Christodoulou, V. Lappas, G. Prassinos, A. Baker and R. Magnuss, “Wireless Sensor Motes for Small Satellite Applications”, in IEEE Antennas and Propagation Magazine, vol. 48, no. 5, pp. 175-179, Oct. 2006,
24. M. Hara, “A case study analysis of the CO2 emission reduction potentiality of the wireless harness within ICT equipment”, In: M. Matsumoto, Y. Umeda, K. Masui, S. Fukushige (eds) Design for Innovative Value Towards a Sustainable Society. Springer, Dordrecht, 2012.
25. A. Matsubara, T. Ichikawa, A. Tomiki, T. Toda and T. Kobayashi, “Measurements and characterization of ultra wideband propagation within spacecrafts”, 2009 Loughborough Antennas & Propagation Conference, 2009, pp. 565-568.
26. B. Grzesik, T. Baumann T, T. Walter T, F. Flederer F, F. Sittner F, E. Dilger E, S. Gläsner S, J.-L. Kirchler, M. Tedsen M, S. Montenegro S, E. Stoll, “InnoCube—A Wireless Satellite Platform to Demonstrate Innovative Technologies”, Aerospace, 2021, Vol. 8, No. 5, 127,
27. B. Grzesik, S. Gläsner, T. Walter, T. Baumann, F. Sittner, F. Flederer, E. Dilger, S. Montenegro, E. Stoll, “InnoCube - Technology Demonstration of a Wireless Satellite Bus and An Experimental Solid-State Battery”, in the proceedings of the 4S Symposium, 2022.
28. T. Baumann, B. Grzesik, T. Walter, F. Flederer, F. Sittner, E. Dilger, S. Gläsner, Jan.-L. Kirchler, M.Tedsen, S. Montenegro, E. Stoll, “Development and experiences of a fully wireless satellite bus for the InnoCube CubeSat mission”, in the proceedings of the 4S Symposium, 2022.
29. A. Matsubara, T. Ichikawa, A. Tomiki, T. Toda and T. Kobayashi, “Measurements and characterization of ultra wideband propagation within spacecrafts”, 2009 Loughborough Antennas & Propagation Conference, 2009, pp. 565-568.
30. W. Niu, J. Li, S. Liu and T. Talty, “Intra-Vehicle Ultra-Wideband Communication Testbed”, MILCOM 2007 - IEEE Military Communications Conference, 2007, pp. 1-6.

31.“Optical Wireless Links for Intra-Satellite Communications (OWLS): The Merger of Optoelectronic and Micro/Nano-Technologies”, NanoTech 2002 – “At the Edge of Revolution”, 2002.

31. I. Arruego et al., “OWLS: a ten-year history in optical wireless links for intra-satellite communications,” in IEEE Journal on Selected Areas in Communications, vol. 27, no. 9, pp. 1599-1611, December 2009.
32. J. Blasco, C. Vernich, S. D.-L. Rosa, M. Á. Esteban et al., “Optical Wireless Links for Intra-Satellite Communications (OWLS): The Merger of Optoelectronic and Micro/Nano-Technologies”, NanoTech 2002 – “At the Edge of Revolution”, 2002.

Authors Introduction

Dr. Masahiro Tokumitsu

He received his Ph.D. degree from Toyohashi University of Technology in 2012. He is an Associate Professor at Division of Computer Science and Systems Engineering, Department of Integrated Engineering, National Institute of Technology, Yonago College in Japan.

Dr. Masatoshi Tsuji

He received his Ph.D. degree from Ritsumeikan University in 2006. He is a Professor at Department of Electrical and Computer Engineering, National Institute of Technology, Kagawa College in Japan.

Dr. Jun Nakaya

He received his Ph.D. degree from Muroran Institute of Technology in 2004. He is an Associate Professor at Department of Electronic Robot Engineering, Faculty of Engineering, Aichi University of Technology in Japan.

Magnetic Anomaly-Matched Trajectory and Dead Reckoning Fusion Mobile Robot Navigation

Yong Hun Kim
Department of Intelligent Mechatronics Engineering and Department of Convergence Engineering for Intelligent Drone,
Sejong University, 05006, Seoul, Republic of Korea

Bo Sung Ko
Department of Intelligent Mechatronics Engineering and Department of Convergence Engineering for Intelligent Drone,
Sejong University, 05006, Seoul, Republic of Korea

Jin Woo Song
Department of intelligent Mechatronics Engineering and Department of Convergence Engineering for Intelligent Drone,
Sejong University, 05006, Seoul, Republic of Korea

E-mail: jwsong@sejong.ac.kr
https://sites.google.com/view/incsl

Abstract

Environments with varying magnetic field distortion cannot be navigated stably with magnetic anomaly based navigation algorithms. In this study, we propose a stable navigation solution for various indoor environments by fusing magnetic anomaly matched trajectories and mobile robot inertial trajectories. The proposed method uses dead reckoning as the primary navigation system and compensates for the navigation sensor error with a feedback structure through the optimization of the anomaly matching trajectory and dead reckoning trajectory. In addition, by determining the trajectory key-frame, the extended Kalman filter measurement update is performed using only the localization results with high accuracy. An open dataset was used to verify the performance of the algorithm, which was compared with existing algorithms. The proposed method is cost-effective, because the proposed method uses only an odometer, gyroscope, and magnetometer for indoor navigation.

Keywords: Mobile robot, Navigation, Magnetic sensor, Trajectory matching, Hidden Markov model

1. Introduction

Positioning is essential However, in indoor conditions where global navigation satellite system signals cannot be received, fusing sensors are required to calculate the position [1-4] without divergence [5,6]. However, recent studies have mostly focused on suppressing the divergence of position to achieve acceptable navigation accuracy, and do not consider the cost. When the cost is considered, the magnetic anomaly matching localization method can be a superior solution, and several studies have established it as an acceptable indoor positioning system [7,8].

However, owing to the similar steel structure inside the building, there are multiple distorted magnetic fields with similarities, which results in position errors. When the magnetic anomaly localization results are merged using the extended Kalman filter (EKF), this position error propagates and causes navigation errors. Magnetic fields are distorted due to several reasons, and thus, the measurement covariance, which is the localization accuracy, cannot be defined with a constant value. This may result in large navigation errors or failures.

In this study, to solve this problem, the position result calculated through magnetic anomaly localization is configured as a trajectory and used for navigation. The calculated trajectory is compared with dead reckoning (DR), which calculates the position and altitude using mobile robot wheel spin and gyroscope. In the process of optimizing the trajectory, the navigation sensor error was calculated, and the navigation performance was improved by compensating it with a feedback structure. In addition, a stable and accurate indoor navigation algorithm was implemented by performing position measurement updates by detecting a straight section with a high localization accuracy of the magnetic anomaly-matched trajectory as a key-frame. The implemented algorithm was verified using an open dataset.

The remainder of this paper is organized as follows. In Section 2, magnetic anomaly matching and trajectory optimization techniques are described. Section 3 describes the evaluation of the algorithm quantitatively using an open dataset, and Chapter 4 presents the conclusions.

2. Mobile robot navigation using magnetic anomaly-matched trajectory

2.1. *Magnetic anomaly-matched trajectory*

Magnetic anomaly based localization is divided into methods that use a cost function and those that use a probability model.

In this study, each technique must be independently localized, because we consider that the magnetic anomaly and DR trajectory are fused. The cost function-based method has low accuracy when only the magnetic anomaly matches; therefore, in this study, a magnetic anomaly-matched trajectory was constructed using a probability-based hidden Markov model (HMM).

Fig. 1 shows the HMM in which each parameter is connected at the current time t and previous time t-1. Each hidden parameter of the HMM was selected as a magnetic map grid point within 3 m based on the previous location, and the transition probability a_{Nj} was modeled as the distance of the grid point based on the previously determined location.

$$a_{Nj} = \frac{1}{\sqrt{2\pi}\sigma_d} exp\left(-\frac{(\|\boldsymbol{Map}(N)-\boldsymbol{Map}(j)\|^2)}{2\sigma_d^2}\right) \qquad (1)$$

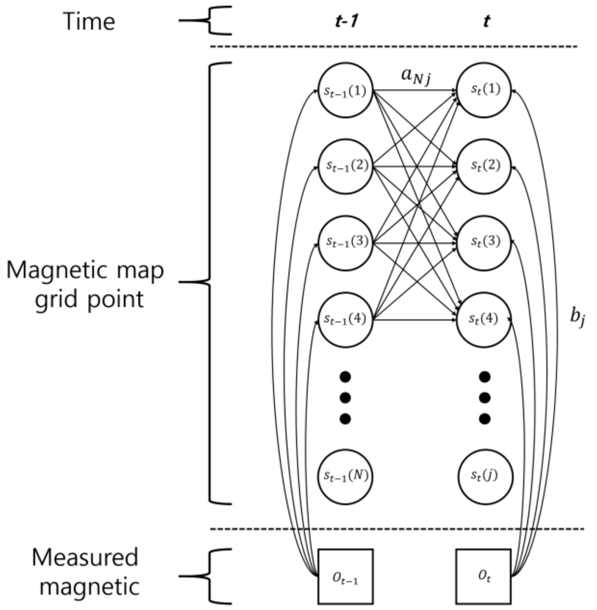


Fig. 1. HMM model for magnetic anomaly-matching.

Eq. (1) gives the transition probability, where $\mathbf{Map}(\cdot)$ means a matrix that stores the positional coordinates of the preconfigured magnetic field grid points, $\boldsymbol{N,j}$ is an index representing a row, and σ_d is the variance value according to the distance, selected as $\sqrt{3}$, which is the area setting value.

The emission probability b_j is modeled by the difference between the measured magnetic field and the constructed magnetic field map.

$$b_j = \frac{1}{\sqrt{2\pi}\sigma_m}\exp\left(-\frac{(\mathbf{Mag}(j)-mag)^2}{2\sigma_m^2}\right) \qquad (2)$$

In Eq. (2), **Mag** is the magnetic field stored at a pre-configured grid point, and $\boldsymbol{mag}$ is the magnetic field measured at the current time. Here, σ_m is $var(\mathbf{Mag})$, the variance of the total magnetic field recorded in the magnetic field map.

Using Eqs. (1) and (2), the forward probability at each time point can be calculated, as shown in Eq. (3).

$$\mathrm{s}_t(j) = \Sigma_{i=1}^{N}\mathrm{s}_{t-1}(i)a_{ij}b_j \qquad (3)$$

After calculating the parameter probability, and decoding using the Viterbi algorithm to form the trajectory. Therefore, a trajectory is constructed by collecting the positions of the grid points with the maximum probability among the forward probabilities. The trajectory was smoothed using spline fitting, and the trajectory arrangement interval was 1 second.

2.2. *Key-frame detection and error compensation techniques*

In this study, to minimize the error of the magnetic anomaly matched trajectory and DR trajectory, stable navigation is performed by compensating for the odometry scale factor error, which is the wheel rotation error and gyroscope bias.

In the case of the turning section, wheel slip occurs significantly, and magnetic anomaly matching results become inaccurate. Thus, the odometry scale factor error and gyro bias are compensated through trajectory information, not position error compensation.

$$e(o_s, bg) = \underset{o_s,b_g}{argmin}\|\mathbf{M}_{tr} - \mathbf{D}_{tr}\|^2 \qquad (4)$$

In Eq. (4), $\mathbf{M}_{tr}$ is the position of the magnetic anomaly-matched trajectory, and $\mathbf{D}_{tr}$ is the position of the trajectory calculated by DR. Stable navigation is performed by calculating the odometry scale factor $\boldsymbol{o_s}$ and gyro bias $\boldsymbol{b_g}$ that minimize the position errors of mutual trajectories and then compensating for them.

Even if the odometry scale factor error and gyro bias are compensated, the position eventually diverges over time if the position is not compensated. Therefore, in a straight section with a relatively high magnetic anomaly matching accuracy, the position error is compensated by using the matched position as the position measurement of the EKF.

The turning and straight sections were detected according to the gradient change in the magnetic anomaly-matched trajectory. When the detection result was straight, EKF measurement was performed to compensate for the position error.

Fig. 2 shows the trajectories of the detected turning and straight sections. After calculating the trajectory, the odometry scale factor and gyro bias that minimize the error between dead reckoning and magnetic anomaly matching trajectory were calculated and compensated.

The EKF navigation model and process are described in [9], and the algorithm structure of the mobile robot navigation system proposed in this study is illustrated in Fig. 3.

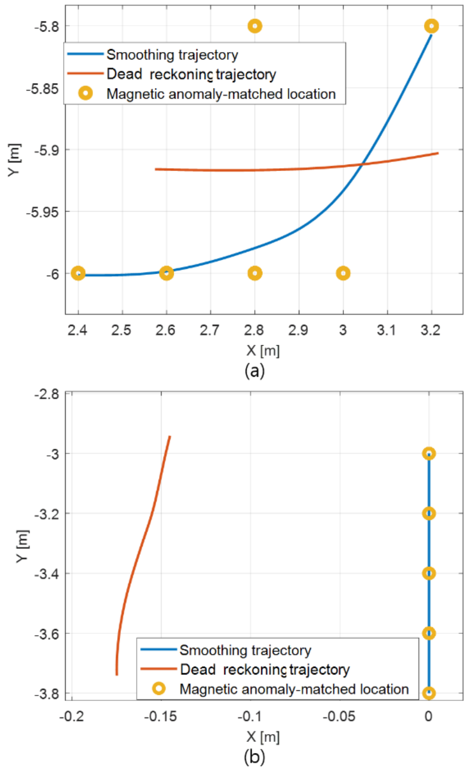


Fig. 2. Magnetic anomaly-matching result and smoothed trajectory (a) detected turning section (b) detected straight section.

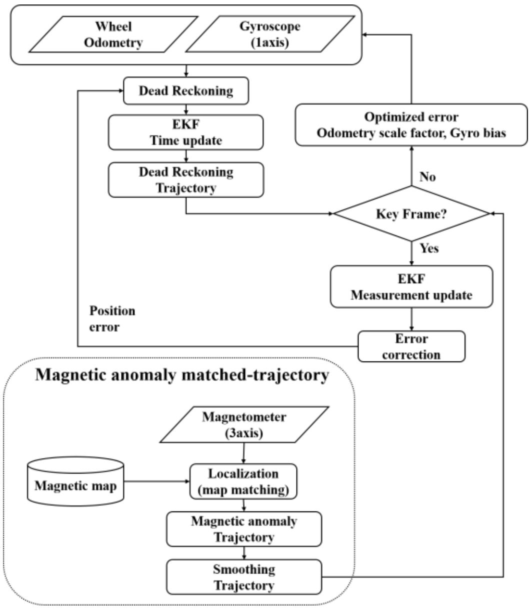

Wheel Odometry
Gyroscope (1axis)
Dead Reckoning
EKF Time update
Dead Reckoning Trajectory
Optimized error Odometry scale factor, Gyro bias
No
Key Frame?
Yes
EKF Measurement update
Position error
Error correction
Magnetic anomaly matched-trajectory
Magnetometer (3axis)
Magnetic map
Localization (map matching)
Magnetic anomaly Trajectory
Smoothing Trajectory

Fig. 3 Full block diagram of the proposed navigation algorithm

3. Verification using the open dataset

The proposed algorithm was verified using the MagPIE dataset [10]. The algorithm accuracy was analyzed for three trajectories out of the entire dataset, and the magnetic map grid size was selected as 0.2 m.

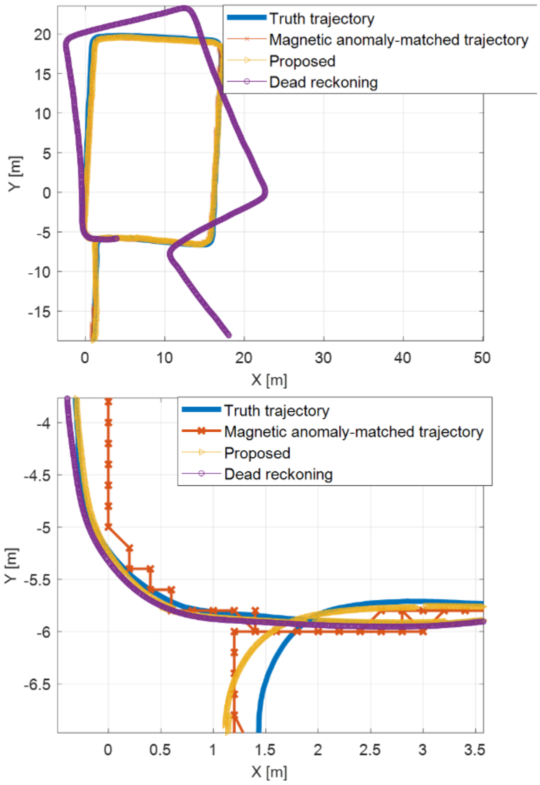


Fig. 4. Algorithm result for the entire section and turning section.

As shown in Fig. 4, it can be confirmed that stable navigation is realized. In addition, because the position is not used directly, a structure that compensates for the errors is used. Therefore, it is not dependent on magnetic anomaly matching results. The algorithm performances for the three experiments are presented in Table 1.

Table 1. Position RMSE (MM is magnetic anomaly matching)

	DR (m)	Proposed (m)
Dataset1	7.4037	0.2976
Dataset2	7.0465	0.3572
Dataset3	18.708	0.5583

The proposed technique enables navigation with very high accuracy compared to DR. In addition, the magnetic anomaly matching technique must determine the region corresponding to the candidate group based on the previous location. When using the proposed technique, the DR performance and localization accuracy of the magnetic anomaly matching technique are improved.

4. Conclusion

This study aimed to improve the inaccuracy of the magnetic anomaly matching technique, which can be used to construct a low cost indoor navigation system. The proposed method used the localization result as trajectory information instead of the measurement of the EKF, and demonstrated that it is possible to perform navigation accurately and stably by compensating for the error with an optimization method in a turning section with less accuracy. In the future, we plan to conduct more rigorous verification by conducting our own experiments.

Acknowledgments

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2022-2018-0-01423) supervised by the IITP(Institute for Information & Communications Technology Planning & Evaluation) and supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No.2020R1A6A1A03038540).

References

1. Gan, X., Yu, B., Wang, X., Yang, Y., Jia, R., Zhang, H., ... & Wang, B. (2019). A new array pseudolites technology for high precision indoor positioning. IEEE Access, 7, 153269-153277.
2. Kim, Y. H., Choi, M. J., Kim, E. J., & Song, J. W. (2019). Magnetic-map-matching-aided pedestrian navigation using outlier mitigation based on multiple sensors and roughness weighting. Sensors, 19(21), 4782.
3. Kwon, Y. S., Kim, Y. H., Do, H. V., Kim, H. J., & Song, J. W. (2021, October). Radar Velocity Measurements Aided Navigation System for UAVs. In 2021 21st International Conference on Control, Automation and Systems (ICCAS) (pp. 472-476). IEEE.
4. Kim, Y. H., Kwon, Y. S., & Song, J. W. (2020, October). Magnetic-map-matching Navigation For Odometry Using Low-cost Mobile Robot. In 2020

IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE) (pp. 392-395). IEEE.

5. Yudanto, R. G., & Petre, F. (2015, October). Sensor fusion for indoor navigation and tracking of automated guided vehicles. In 2015 International Conference on Indoor Positioning and Indoor Navigation (IPIN) (pp. 1-8). IEEE.
6. Alatise, M. B., & Hancke, G. P. (2020). A review on challenges of autonomous mobile robot and sensor fusion methods. IEEE Access, 8, 39830-39846.
7. Gozick, B., Subbu, K. P., Dantu, R., & Maeshiro, T. (2011). Magnetic maps for indoor navigation. IEEE Transactions on Instrumentation and Measurement, 60(12), 3883-3891.
8. Wu, F., Liang, Y., Fu, Y., & Ji, X. (2016, April). A robust indoor positioning system based on encoded magnetic field and low-cost IMU. In 2016 IEEE/ION Position, Location and Navigation Symposium (PLANS) (pp. 204-212). IEEE.
9. Cho, B. S., Moon, W. S., Seo, W. J., & Baek, K. R. (2011). A dead reckoning localization system for mobile robotswmd using inertial sensors and wheel revolution encoding. Journal of mechanical science and technology, 25(11), 2907-2917.
10. Hanley, D., Faustino, A. B., Zelman, S. D., Degenhardt, D. A., & Bretl, T. MagPIE: A dataset for indoor positioning with magnetic anomalies. In 2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN) (pp. 1-8). IEEE.

Authors Introduction

Mr. Yong Hun Kim

He received the B.S. degree in Robotics Engineering, from Hoseo University in 2018, and the M.S. degree in Software Convergence, in 2020, from Sejong University, Seoul, South Korea, where he is currently working toward the Ph.D. degree in Department of Intelligent Mechatronics Engineering and Convergence Engineering for Intelligent Drone. His research interests include indoor navigation, sensor fusion, INS/GNSS integration and nonlinear filtering.

Mr. Bo Sung Ko

He received the B.S. degree in Mechanical and Aerospace Engineering from Sejong University, Seoul, South Korea, in 2022, where he is currently pursuing the M.S. degree with the Department of Intelligent Mechatronics Engineering and Convergence Engineering for Intelligent Drone. His research interests include precision navigation, SLAM, sensor fusion

Dr. Jin Woo Song

He received the B.S. and M.S. degrees in control and instrumentation engineering and the Ph.D. degree in electrical, electronic and computer engineering from Seoul National University, Seoul, South Korea, in 1995, 1997, and 2002, respectively. From 2003 to 2014, he was a CTO with Microinfinity Company, Ltd. He is currently an Associate Professor with the school of Intelligent Mechatronics Engineering and Convergence Engineering for Intelligent Drone, Sejong University, Seoul, South Korea. His research interests include GPS/INS integration, robust and optimal control, MEMS inertial sensors and systems, indoor navigation, and robot localization.